Children and Youth Services Review 177 (2025) 108502

Contents lists available at ScienceDirect

Children and Youth Services Review

journal homepage: www.elsevier.com/locate/childyouth

Maximising existing assets: the potential of Bayesian hierarchical approaches to improve the risk assessment process in youth justice

Helen Ruth Hodges

CASCADE, Cardiff University, SPARK Building, Maindy Road, Cardiff CF24 4HQ, United Kingdom

ARTICLE INFO

Keywords:
Bayesian hierarchical modelling
Youth justice
Risk assessment
Data linkage

ABSTRACT

Bayesian approaches provide a fresh lens to consider the interaction between various characteristics and circumstances of children who have offended, and their likelihood of committing further offences over time. This has been done here by revisiting data captured as part of the risk assessment process used until the mid-2010s across England and Wales to consider how the probability of further offending behaviour differs depending upon whether the child has a prior history of offending or not.

Whilst Asset has now been replaced by AssetPlus, the structure of the earlier risk assessment tool lends itself to mimicking rapid changes in the lives of some children and the evolving nature of youth offending. This has been done by treating the assessments conducted with 87 children in the formal youth justice system in a single Welsh local authority area whose supervision started in either 2012/13 or 2023/14 as a longitudinal dataset. The likelihood of further offending behaviour at different time points has then been estimated using additive binary logistical regression models based upon practitioner ratings for 12 domains of 'risk', along with a range of time-varying and non-time varying variables reflecting facets of their criminal career.

This study demonstrates the utility of conducting the analysis in a Bayesian framework. Notably it highlights the potential to incorporate new ideas be they emerging theoretical perspectives, interventions or additional variables into existing models to increase understandings of the complex relationship between 'risk' and 'protective' factors for different subgroups.

1. Introduction

For nearly twenty-five years, the dominant paradigm for understanding and addressing youth crime in many jurisdictions, including England and Wales, has been provided by the concept of risk factors, with the risk factor prevention paradigm (RFPP) – transplanted from medicine and public health (Farrington, 2000b) – proving to be attractive to policy makers, practitioners and researchers interested in youth offending (Armstrong, 2006; Garside, 2009). Consequently, risk factor research within youth justice policy and practice has grown exponentially (Case & Haines, 2009). Whilst this has led to advances in our understandings of the aetiology of children's offending behaviours which have been translated into practice, there remains concern around the completeness and focus of the evidence base. This has ramifications for the implementation of more holistic and children's rights-based approaches. Notably Case and Haines (2009) highlight how factorisation and reductionism have over-simplified the context and operation of risk whilst the aggregation of findings has limited the potential to understand some of the complex relationships that exist within youth offending. These authors were also critical of the promotion of prevention as a dichotomy of risk.

This concern around oversimplification also applies to the individual, with dimensional identity (Schwalbe et al., 2006) tending to be underexplored, leaving significant gaps in the evidence base with respect to minority groups within the youth justice system (YJS). Notably in England and Wales there have been reviews around the over-representation of care-experienced children i.e. those with current or prior experience of out of home care (Prison Reform Trust, 2016) and those from minority ethnic groups (Lammy, 2017). The tendency to treat these as homogenous cohorts rather than acknowledge their diversity means that not only has little been done to investigate potential differences on the basis of age, gender and heritage, but also that less is known about minority subgroups such as those of Muslim faith; care-experienced girls; children with disabilities, learning difficulties and speech, language and communication needs; and trafficked and foreign national children including asylum seekers (Lammy, 2017; Prison Reform Trust, 2016).

At a time when we are seeing increased use of diversion from the

E-mail address: hodgesh@cardiff.ac.uk.

https://doi.org/10.1016/j.childyouth.2025.108502
Received 17 September 2024; Received in revised form 24 July 2025; Accepted 25 July 2025
Available online 26 July 2025

formal YJS (Kelly & Armitage, 2015), the statutory case load is acknowledged to be becoming increasingly complex (Youth Justice Board & Ministry of Justice, 2020, 2021; Youth Justice Board, 2016a). For researchers wishing to untangle the complex web of relationships between risk, need and recidivist behaviour, this represents both a significant mathematical and conceptual challenge especially in jurisdictions where there are now a significantly smaller number of potential data subjects within the cohort – in the year to March 2023, there were 13,743 children[1] cautioned or sentenced across England and Wales compared to 49,222 in the equivalent period 10 years ago. Of these, 3,694 (27.9 %) were identified as being minority ethnic children, 1,769 (13.8 %) were known to be female and 579 (4.2 %) were pre-teens. Just 104 (08 %) were female pre-teens whilst 373 (2.7 %) were minority ethnic females (Youth Justice Board, 2024).

Given the number of permutations of offence type, sentence and individual level characteristics, it is tempting to aggregate groups. However, if we are to truly understand the aetiology of offending behaviour and devise tailored interventions to promote desistance, then we require methods which allow us to use an intersectional approach drawing upon more efficiently upon the data available to us. Notably, the smaller numbers in the youth justice system mean that we are arguably now at the limits of what can be achieved using classical statistical techniques which tend to require large sample sizes. This paper therefore presents the case using Bayesian hierarchical modelling as an alternative.

Whilst rarely used in criminology (Stander et al., 2022), Bayesian approaches to data analysis and parameter estimation are based on Bayes' theorem whereby all observed and unobserved parameters in a statistical model are given a joint probability distribution, termed the prior and data distribution. The resulting posterior distribution reflects updated knowledge, balancing prior knowledge with observed data, and is used to conduct inferences (van de Schoot et al., 2021). Rather than p-values being reported, the significance of tests is reported using Bayes factors. Mathematically, these are defined as the ratio of two marginal likelihoods: the likelihood of the data under the null hypothesis (H_0) and the likelihood of the data under the alternative hypothesis (H_1). Bayes factors therefore can be used to quantify the evidence in favour of one statistical model compared to another (Kass & Raftery, 1995). Further information about how to interpret Bayes factors can be found in Appendix 1.

The known difference in the proven reoffending rates of first-time entrants and those with a prior history of offending (Youth Justice Board, 2024) provides an opportunity to demonstrate how Bayesian models can be adapted and new variables added as and when new information emerges. Focusing specifically on the period after the index assessment has been completed, this study considers how the probability of further offending behaviour differs depending upon first-time entrant (FTE) status. In addition to considering initial differences, temporal changes in the likelihood of further offending behaviour for both groups are considered since it is hypothesized that as practitioners work with the child, the likelihood of further offending behaviour will decrease at different rates. Notably whilst the child is being supervised by the youth offending team (YOT) there are opportunities for intervention and the child's circumstances may change. The efficacy of these interventions and potential impact of these changes are beyond the scope of this paper. However, the intention is to illustrate how the approach could be built upon to enhance understandings of the complex relationship between risk and protective factors which can then be translated into practice.

[1] The age of criminal responsibility in England and Wales is 10. The age of criminal majority is 18. Age is calculated at the time of the first hearing. Since children are supported through the trial process by the YOT even if they turn 18 before the sentence is passed, there is a small number of 18-year-olds within the YJS.

2. Literature review

2.1. The opportunity for a change in approach

Criminologists have a history of drawing upon 'sophisticated and cutting-edge approaches from other fields' and have 'given significant attention to the ways such approaches must be adapted to fit criminological problems' (Bushway & Weisburd, 2006, p. 1). The adoption of novel statistical approaches such as data visualisation, data linkage and Bayesian statistical inference represents a continuation of this tradition, adapting techniques which have been successfully employed elsewhere to advance knowledge and understanding around the aetiology of offending behaviour, and informing what works in terms of societal responses. For example, data visualisation techniques including geospatial Bayesian applications have been used to consider crime trends (Law et al., 2014), and more recently the age-crime curve (Stander et al., 2022). Additionally, the UK Ministry of Justice (MoJ) (2022) has made anonymised criminal court, prison, probation and family court data available to accredited researchers opening up the opportunities for data linkage.

The abundance of routine data from different policy areas provides an opportunity for researchers to link subject-level data to compile a comprehensive picture of individuals' lives, with information being available about groups typically under-represented in research studies and/or about rare phenomena. Revisiting the evidence base in youth justice using a fresh lens and the Asset data provide an opportunity to address the lack of sensitivity of risk factor research and the risk assessment tools which it has informed. Not only do the risk assessment tools and other routine data from the wider justice system provide a rich picture of children's lives, but their standardised structure lends itself to treating the assessments as a longitudinal dataset to consider change over time. Findings can then be used to inform the development of a more holistic approach which takes into account individual needs and circumstances, offering the potential for a further shift away from the spectre of risk-focused youth justice towards one which is better aligned with Child First principles.[2]

2.2. Risk assessment in youth justice

As with the wider criminal justice system, youth justice has become increasingly reliant upon standardised actuarial risk assessment tools, not just when considering sentencing and release, but also to make decisions around assignment to interventions treatment (Bonta & Wormith, 2013). As with many of the risk assessment tools which have emerged largely since the 1990s, Asset was grounded in the statistical association between risk and repeat offending (Schwalbe, 2007). Focusing on prediction and classification, under the 'Scaled Approach' introduced in 2009 (Youth Justice Board, 2010), the tool was characterised by its predictive role in informing intervention planning. As such it combined fixed or 'static' risk factors like offence history with an array of 'dynamic' risk factors which it may be possible to change as a result of intervention (Farrington, 2000a) – see Fig. 1. A description is also provided in Appendix 2. It has since been superseded by AssetPlus which purports to support desistance approaches (Hampson, 2018; Youth

[2] Child First is a pro-children's rights approach which prioritizes treating children and young people involved in the youth justice system as "children first, offenders second". Taking its origins from Welsh policy, the Child First framework, adopted by the Youth Justice Board for England and Wales, is built on four key tenets: seeing children as children, building a pro-social identity, collaborating with children, and diverting children from stigma. As such it emphasizes the importance of addressing the child's needs and well-being above their actions as an offender, aiming to prevent future offending and promote positive development. (See for example Case & Browning, 2021; Youth Justice Board, 2019).

Fig. 1. Components of asset – static and dynamic risk factors. Adapted from Baker et al. (2003) (p 99).

Justice Board, 2014) but as highlighted previously, is also considered to have significant limitations.

With the shift towards more sophisticated tools such as AssetPlus which require information to be gathered not only about criminogenic needs and responsiveness, but also to actively gather information to facilitate planning, case management, supervision and service delivery (Vaswani & Merone, 2013), it is now even more important that we need to know *why* it is some children come into conflict with the law and subsequently commit further offences; *why* certain interventions work for some and not for others, and *why* some may be more responsive to specific interventions than their peers. Without this, 'risk is hidden beneath a plethora of correlations that in themselves tell us little about the socio-historical nature, meaning and significance of crime and its discourses in these time in which we are now living' (Armstrong, 2004, p. 113). If we are to appropriately respond to youth offending behaviours, it becomes all the more important to appropriately conceptualise and operationalise both the outcome and the predictor variables.

The developmental focus of risk factor research has historically meant that researchers have sought early childhood psychosocial factors that are statistically related with the onset of teenage offending (Case & Haines, 2009). In doing this, risk factor research has tended to utilise broad (factorised) measures of risk factors, relating them statistically to broad categories of offending (i.e. a single offence of any type is counted as 'offending' and any three offences are taken to be 'serious offending') (Case & Haines, 2010). As a result, studies of the risk factor-offending relationship for children have been overly superficial, generalised and insensitive. It has also been suggested that there is a 'psycho-social bias' (France & Homel, 2016) which has resulted in 'an artificial restriction in the range of factors that have been explored' (Case & Haines, 2009, p. 22). The reality is that 'many thousands of factors may place young people 'at risk' of offending, including, at different ages, 'biological, individual, family, peer, school, neighbourhood, and situational factors' (Brown, 2005, p. 100). This leads to problems as many children who are technically 'at risk' lead 'successful lives', thus there is also a need to understand factors which in many various combinations act in a protective way, mitigating risk.

Attempts have been made by the Youth Justice Board (YJB) to provide clarification around the definition of risk being used and to reflect emerging research, policy and practice to reduce the psychosocial bias (Baker, 2014; Youth Justice Board, 2014). Alongside the more recent interest in promoting desistance, there has been a growing awareness of the extent to which those in the YJS have been exposed to childhood adversity which has led to the adoption of trauma informed practice, and the need for a Child First approach. Whilst the increased recognition that involvement with the formal YJS can be criminogenic and potentially inherently damaging for children (McAra & McVie, 2010) has contributed to changes in police reporting criteria and increased diversion (Bateman, 2020), these practice shifts have increased the burden placed on practitioners to collate information both from the individual and partner agencies in order to build a picture of the child's circumstances and hence determine appropriate interventions based on their needs and vulnerabilities.

In the setting out the rationale for AssetPlus it is suggested that whilst developing the new tool there was an opportunity to 'build on previous assessment tools used in youth justice settings … incorporating both the lessons learned from their use since 2000 and new insights from research and the academic literature' (Baker, 2014: p3). It was further noted that AssetPlus was 'designed to reflect the changing context for practice in which greater emphasis is now being placed on flexibility and the importance of professional discretion' (Baker, 2014: p3). Whilst proponents welcomed the shift in assessment practices away from risk-focus and towards a more holistic and dynamic assessment and intervention planning framework (Case, 2021), it is felt that AssetPlus has many limitations. These include its unnecessary length and complexity, its child-unfriendliness and its lack of empirical validation. Consequently, it has been argued that a full review and redesign is overdue (Drew, 2023) and resource should be allocated to developing a framework that is easier to use and better aligned to the Child First framework (Smith & Paddock, 2025).

2.3. Asset verses AssetPlus

At the time of the study, data had been collected in a consistent manner using Asset for more than 15 years with the tool being the subject of three evaluations (Baker et al., 2003, 2005; Wilson & Hinks, 2011). As a result, its relative strengths and limitations were well known (see for example Case & Haines, 2009, Stephenson et al., 2011). Namely that in addition to the conceptual and methodological issues associated with the RFRP underpinning Asset, there were also concerns about its predictive utility, how well it aligned with policy and the evidence base, as well as how efficient it was given resource constraints.

Whilst it would be desirable to utilise more recent data – AssetPlus was rolled out across all youth justice services between September 2015 and November 2017, and across the secure estate in the first half of 2018, it is not without its limitations. As Smith & Paddock highlight in their critical evaluation of the framework, AssetPlus 'does not lend itself to statistical testing due to it being a holistic assessment where the outcome is a robust intervention plan, rather than a single score which can be easily analysed' (2025: p6). A key advantage that Asset therefore has over the AssetPlus is the way in which the dataset is structured. Notably, the practitioner ratings for each of the 12 domains of risk were dropped as part of the transition and it is these which enable changes in risk to be modelled over time in response to key events and changes in circumstances.

Whereas Asset has been found to be an accurate predictor of reoffending (Wilson & Hinks, 2011), the validity of AssetPlus has never been statistically assessed since there are no high-quality data available to explore its ability to predict outcomes such as reoffending – something which in itself does not devalue AssetPlus as a tool since it has a wider purpose. It is intended to include all information of clinical relevance to practitioners for intervention planning. However, in the context of this study this is pertinent since it is the relationship with further offending that is being explored.

3. Methods

3.1. Contextualisation

This study was conducted using standardised risk assessment data collected using the Asset tool completed by practitioners from a single Welsh local authority YOT. Permission was granted by the data owners to extract historic individual-level records from the YOT's case management system (CMS) prior to the roll out of its successor, AssetPlus in the mid-2010s. Ethical approval for this was granted by Swansea University and the data owner, and was permitted on condition that the data were anonymised for the purposes of this study. As a result, all preparatory work was undertaken under the supervision of the YOT's Information Manager within their secure working environment. This included the removal of all potentially identifiable information and generation of a randomly generated unique identifier. Despite its historical nature, every effort was taken to ensure that it remained compliant with all applicable data retention and data governance policies and practice. Notably, the General Data Protection Regulation (GDPR) had not yet come into force at the time of that the research was conducted. However, local authority specific guidelines for the retention of personal and sensitive data were followed – these had been updated to be consistent with the underpinning GDPR principles.

The CMS contains individual level data stored in a number of different formats including subjective rating scales from the Assets, free text, dates and postcodes. Whilst it was possible to do bulk-download of the Assets completed between specified dates from the live system and address histories, it was necessary to go into individual case histories to download the data required around offences and court appearances. The presence of a unique system identifier meant that it was possible to link individual's records across the various tables. The 'date stamping' of activity provides a means of establishing temporal ordering for example, highlighting the timing of offences and court proceedings, and changes in their risk scores over time. Having the individual's offending history meant that FTE status could be determined enabling the sample to be divided into first-time entrants and those with a prior offending history at the time of their index Asset (Time 0).

The temporal converge of the records was constrained by the introduction of a new software system in 2012 which had resulted in a purge of older records, and the transition to AssetPlus at which point practitioner ratings ceased to be part of the assessment. During this period, local YOTs were provided with 'reoffending spreadsheets' as part of the YJB's Reducing Reoffending Programme (Youth Justice Board, 2016b) to enable them to identify areas for improvement and targeting resources; reconciling and identify gaps between the local and police national computer (in particular for 17-year olds and pre-court disposals); and compare performance both over time and with different geographies (Youth Justice Board, 2017). Pre-populated with data from the police national computer, these spreadsheets provided data on reoffending for each member of their reoffending cohort. Table 1 summarises the initial profile of those with Assets that the YOT had engaged

Table 1
Inclusion criteria: initial profile.

Cohort	Period of interest	Unique individuals	No. subject to a full risk assessment	Reoffending	
				n	Rate
2012/13	Assets dated from the outcome date of their primary offence in 2012/13 to 31st March 2014	134	63	35	55.6 %
2013/14	Assets dated from the outcome date of their primary offence in 2013/14 to 31st March 2015.	131	61	16	26.2 %

with as part of their statutory caseload.

3.2. The sample and its representativeness

Across the two cohorts, there were a total of 100 unique individuals. However, those Assets which were blank, duplicates (based on start date and domain scores) or which pre-dated the individual's entry to the cohort were excluded as were those individuals with only a single Asset. This reduced the size of the cohort to 87. Between them they had 545 Assets covering the period from 1st April 2012 to 31st March 2015. Their demographic characteristics are summarised in Table 2 along with the respective proportions that went on to commit further offending during the period of interest.

Gender and ethnicity were taken from the YOT's CMS. Age at the time of the first offence were calculated using the child's date of birth. As can be seen in Table 2, there were a low number of females and ethnic minority children in the cohort. This is broadly in line with what would be expected given that across England and Wales only 19 % (7720 out of 41,569) of those cautioned or sentenced in 2013/14 were girls (Youth Justice Board & Ministry of Justice, 2015), and the relative lack of diversity within the local population (Office for National Statistics, 2012). The low number of minority ethnic children (6 out of 87, 7 %) limits the amount of analysis that can be undertaken around ethnicity, particularly if cross-referenced by gender.

The non-White group included individuals from Black, Asian and mixed backgrounds. However, they made up just 6.9 % of the cohort. All were male. The youngest children were aged 10 whilst the older children were approaching their 18th birthday. The average age at index offence was 13 years, 9 months with their being no difference in the mean age by gender. Almost six out of ten were aged 10 to 14 at the time of their index offence.

Overall, 37 out of the 87 went on to commit a further offence – equivalent to a further offending rate of 42.5 %. Rates were lower amongst females, those who were non-White and those who were older when they committed their index offence. The respective rates do not give cause for concern in terms of model development. However, the small numbers of females and minority ethnic children mean that there is likely to be greater uncertainty around the estimates of the model coefficients, if the models did indeed converge. Had the overall sample size been higher, it would have been desirable to incorporate dimensional identity eg age, gender and ethnicity alongside FTE status in the modelling when addressing the research question. These variables have however, been considered when looking for explanations for the observed differences over time.

3.3. Data and measures

3.3.1. Outcome variable – further offending

An event flag to denote whether, in the period prior to the assessment, the child has committed one or more further offences was added to the dataset based on information held in their YOT offence record. As a result, at Time 0 (the initial assessment), every child is reflected as having committed one or more offences. Regardless of the ultimate disposal, the initial Asset coinciding with arrest/charge and subsequent sets of risk scores provide a picture of the child's circumstances, how they change whist supervised by the YOT and if there is any further identified/detected offending behaviour. The measure of identified further offending was therefore preferable to the administrative measure of re-offending/reconviction which is contingent upon the offence being proven. It is also consistent with the model's hierarchical structure.

Table 2
Further offending rate, by gender, ethnicity and age at index offence.

Comparator groups			n	Further offending	
				n	Rate
Gender	1	Male	79	35	44.3 %
	2	Female	8	2	25.0 %
Ethnicity	1	White	81	35	43.2 %
	2	Non-White	6	2	33.3 %
Age	1	10–14 years	52	25	48.1 %
	2	15–17 years	35	12	34.3 %
Total			87	37	42.5 %

Note: Since some were in both the 2012/13 and 2013/14 reoffending cohorts, the child's age group reflects their age at the time of their 'index' offence i.e. the offence which led to their inclusion within the reoffending cohort.

3.3.2. Predictor variables

Reflecting the Asset Core Profile under the Scaled Approach (see Fig. 1 for a diagrammatic representation of the relationship between the static and dynamic risk factors and Appendix 2 for a description), the decision was made to replicate the structure of the risk assessment tool using an additive model. Predictor variables when considering how the probability of further offending differs depending upon FTE status therefore include:

- The 12 'dynamic' domains of risk against which practitioners rated the likelihood of reoffending on a scale from 0 – no association to 4 – very strongly associated.
- FTE status as an example of a 'static' factors associated with the individuals' criminal career.
- Time – reflected by the number of Assets completed.

The number of Assets is depend upon the duration of a child's order, the complexity of their personal circumstances and offending behaviours – individuals under the supervision of the YOT typically have an 'Start' and 'Finish' Asset but may also have 'Review' Assets since the National Standards (Youth Justice Board, 2013) recommended that assessments were reviewed every three months or where there had been a significant change in the child's circumstances. In this instance 61 % (53 out of 87) of the individuals had five or fewer records with the maximum being 19. Since measurement occasions are not fixed and individuals can have differing numbers of assessments, the data are considered to be unbalanced.

The subjective nature of the ratings means that the domain scores are ordinal (from 0 to 4). However, in keeping with the arguments made by Robitzsch (2020), the ratings have been treated as being continuous. The distributions of each of the domain scores suggests that it is appropriate to assume that each set of measurements is independent and that each potential rating has an equal probability of being assigned. In this instance, the zero is meaningful therefore it has not been necessary to centre or standardise the domain scores.

3.4. Statistical analysis

The longitudinal nature of the repeated Assets lends itself to analysis by way of a hierarchical or multilevel where Level 1 is generally associated with a single measurement in time and Level 2 refers to an individual subject. In this way it is possible to maximise the advantages associated with the flexibility and power of such models. Notably, it is advocated that modelling longitudinal data in a multilevel framework enables the simultaneous modelling of both intra-individual and inter-individual change (Finch et al., 2014).

Multilevel models also allow for more complex data structures to be explored and can be considered to be a powerful extension to linear and generalised linear regression modelling (Gill & Womack, 2013). Through use of a nested data structure, it is possible to avoid the unaccounted-for heterogeneity and correlation which are common in conventional, flat modelling. This has made hierarchical linear models the main type of application in biological and medical sciences (Snijders & Bosker, 2012). Whilst examples can be found in the social sciences, particularly in education and psychology, they are less commonly used in criminology. Notably it is easy to incorporate both time-varying 'Level 1′ predictors and time invariant 'Level 2′ or individual level characteristics. In this way, temporal changes associated with both the domain scores and individual characteristics can be explored. The example given here relates to the addition of a Level 2 predictor.

Having initially compiled a series of preparatory models as recommended by Robson & Pevalin (2016), three models were then compiled to demonstrate the ease with which additional predictor variables can be added and to enable the research question to be considered:

- Model 1 – the initial *basic model* involving all 12 domain scores plus time without any interactions. Conceptually it was felt that none of the domains could be dropped to reduce complexity
- Model 2 – the *basic dynamic model* reflecting the time-varying nature of the risk assessment process by including interactions between each of the 12 domains and time
- Model 3 – the *basic dynamic model with interaction terms* to enable sub-cohorts to be compared.

In the case of Model 2, this enables the domain scores to vary by time and in response to changes in circumstances meaning that it more closely mimics real life. For Model 3, first-time entrant (FTE) status has been used as a non-time varying predictor. The models were compiled in RStudio using the MCMCglmm package (Hadfield, 2010). For a technical explanation of the model development see Appendix 3.

Throughout, the goodness of fit of the generalised linear models with binary distribution has been evaluated using the Deviance Information Criterion (Spiegelhalter, Best, & Carlin, 2002). The Deviance Information Criterion (DIC) is a Bayesian information criterion that quantifies the information in the fitted model by measuring how well the model reduced uncertainty of future predictions. Adding more parameters often improves the fit of the model (resulting in a lower DIC), but a penalty can be incurred as complexity increases. The DIC simultaneously accounts for model complexity (number of parameters) and model fit, by penalizing based on the number of (effective) parameters. It is calculated based on the sum of the effective number of parameters and the posterior mean of the deviance, with deviance defined as −2 times the log of likelihood function. The DICs for the preparatory models can be found in Table 3.

When comparing the rates of further offending by sub-cohort, significant tests were completed under a Bayesian framework using the test for Bayesian Contingency Tables within JASP version 0.19.1.

4. Results

When considering how the probability of further offending differs depending upon FTE status, three models have been compiled to examine how the probability changes over time. Given that it is hypothesized that there will be a difference in the respective rates at which the probability of further offending behaviour reduces in response to activities that take place during their supervision, these models have been constructed to also enable this to be tested. Possible explanations for these differences linked to their circumstances at Time 0 and the composition of the respective cohorts are also presented here.

Table 3
The DICs for the preparatory models.

Description	DIC
Empty or 'Null' model – a random intercept model for further offending	661.9
Random intercept model for further offending with a single predictor (Time)	671.9
Random intercept model for further offending including the 12 domains and Time with the individual as a random effect (but not Time)	**608.5**
Random intercept and varying slop model for further offending in which the random effects reflect both the individual and Time.	487.6

4.1. The basic and basic dynamic models

The *Basic Model* (Model 1, summarised in Table 4) represents the repeated measurements from the Asset's 12 domains and is a random intercept with varying slope model for further offending including the 12 domains and Time. The DIC of 476.2 compares favourably to that for the equivalent model without time as a random coefficient (Table 3) – that model had a DIC of 608.5 suggesting that despite the additional predictors, Model 1 represents an improved fit for the data. Time is both significant as a fixed and random effect i.e. 0 is not in the interval. Notably, the adjusted posterior mean estimate for Time as a fixed effect i.e. Exp(β) and its 95 % credible interval (CI) are negative suggesting that in addition to the random effect of both time and individual, the probability of further offending decreases as Time increases. This is what would be expected if the individual is working with a practitioner from the YOT.

Model 2 extends the Basic Model to create a 'dynamic' model by enabling each of the 12 domain predictors to vary by time. With a DIC of 257.8, the *Basic Dynamic Model (*Model 2) represents a further improvement upon the preparatory model without random effects and upon Model 1 despite its increased complexity. Under Model 2, the estimates for the fixed effect of the Lifestyle and the Thinking and Behaviour domains are flagged as being significant along with those for the interactions between Time and Perception of Self, and Time and the Thinking and Behaviours domain i.e. 0 is in the interval. These results suggest that the ratings for these two domains are statistically significantly related to further offending. The significant interactions suggest that how a child's ratings for these domains change over time is also related to their likelihood of further offending.

4.2. Incorporating a non-time varying measure reflecting the individual's criminal career

As acknowledged in the design of the Asset, the static (non-time varying) factors also have a role to play. Without access to the police national computer, it was not possible to obtain an accurate measure of the individual's previous offending, relying instead on information from the offence and court records held within the YOT's CMS. This information was used to determine whether or not the individual was a first-time entrant (FTE) to the YJS at the time of their index offence. This measure has been added to the Basic Dynamic Model (Model 2) to illustrate how easy it is to build on this to explore the impact of additional factors. Table 4 summarises the adjusted odds ratios and 95 % credible intervals for this model.

With a DIC of 459.3, the 'penalty' associated with the additional complexity is apparent especially when compared to Model 2. However, it also compares favourably with both Model 1 and the preparatory models, and crucially enables us to model differences in the probability of further offending behaviour based on changes in domain scores over time depending upon FTE status.

4.3. Model fit

As a test of the model fit, the coefficients have been used to estimate the probability of further offending from Time 0 (the initial assessment) to Time 10. Whilst it is somewhat artificial, the domain scores were fixed to establish whether the resulting decay curves are consistent with what would be expected for those with differing initial perceptions of risk working with the YOT over a period of time i.e. that after the index offence, the probability of offending declines over time, tending towards zero. This confirmed that those who are judged to have the highest risk of further offending after primary offence i.e. having the maximum rating of 4 for each of the 12 domains, have the highest estimated probability of further offending at Time 0, whilst those considered to have the lowest risk have the lowest initial estimate of probability.

The estimates of the probability of further offending for those with a

Table 4
The adjusted odds ratios and 95 % credible intervals for the three models.

Fixed effect:	Model 1 the basic model			Model 2 the basic dynamic model			Model 3 the basic dynamic model involving FTE status		
	Exp(Post. Mean)	95 % CI		Exp(Post. Mean)	95 % CI		Exp(Post. Mean)	95 % CI	
Individual (intercept)	−1.168	[−2.379, 0.129]		0.333	[0.171, 0.617]		0.214	[0.021, 2.080]	
Time	−0.153	[−0.283, −0.153]	#	0.992	[0.897, 1.092]		0.849	[0.593, 1.227]	
Living arrangements (live)	0.033	[−0.216, 0.033]		1.316	[0.623, 2.719]		0.954	[0.542, 1.775]	
Family and personal relationships (relation)	0.275	[−0.026, 0.275]		1.376	[0.498, 3.783]		1.896	[0.982, 3.833]	
Education, training & employment (ETE)	0.094	[−0.152, 0.094]		0.596	[0.271, 1.315]		0.740	[0.437, 1.179]	
Neighbourhood (where)	0.044	[−0.166, 0.044]		0.775	[0.373, 1.601]		1.218	[0.728, 2.091]	
Lifestyle (life)	0.024	[−0.316, 0.024]		4.280	[1.267, 14.768]	#	1.437	[0.600, 3.791]	
Substance use (drugs)	0.158	[−0.087, 0.158]		0.726	[0.353, 1.439]		1.230	[0.758, 2.154]	
Physical health (physical)	−0.114	[−0.394, −0.114]		0.557	[0.141, 2.374]		0.663	[0.345, 1.271]	
Emotional and mental health (emotion)	−0.003	[−0.249, −0.003]		0.683	[0.269, 1.676]		1.006	[0.619, 1.734]	
Perception of self and others (self)	−0.138	[−0.443, −0.138]		2.717	[0.884, 8.008]		0.594	[0.256, 1.372]	
Thinking and Behaviour (Think)	−0.160	[−0.508, −0.160]		0.330	[0.105, 0.908]	#	0.997	[0.461, 2.258]	
Attitude to offending (attitude)	0.043	[−0.298, 0.043]		1.014	[0.319, 3.053]		1.566	[0.715, 3.758]	
Motivation to change (change)	0.231	[−0.095, 0.231]		2.785	[0.855, 9.372]		0.917	[0.404, 2.041]	
First time entrant (no = ref) (FTE)							2.705	[0.189, 38.879]	
Time: live				0.959	[0.848, 1.102]		0.465	[0.216, 0.955]	#
Time: relation				0.939	[0.763, 1.163]		3.001	[0.725, 12.775]	
Time: ETE				1.083	[0.926, 1.245]		0.208	[0.046, 0.812]	#
Time: where				1.074	[0.951, 1.213]		1.305	[0.388, 4.880]	
Time: life				0.816	[0.659, 1.008]		0.228	[0.069, 0.745]	#
Time: drugs				1.093	[0.947, 1.267]		2.696	[0.521, 14.889]	
Time: physical				1.033	[0.840, 1.265]		0.498	[0.139, 1.713]	
Time: emotion				1.110	[0.921, 1.352]		0.666	[0.171, 2.762]	
Time: self				0.767	[0.614, 0.961]	#	0.591	[0.184, 1.730]	
Time: think				1.334	[1.052, 1.704]	#	10.937	[2.468, 50.930]	#
Time: attitude				1.071	[0.837, 1.362]		0.617	[0.145, 2.690]	
Time: change				0.851	[0.662, 1.083]		0.287	[0.068, 1.259]	
FTE: time: live							3.298	[0.671, 15.392]	
FTE: time: relation							1.019	[0.901, 1.163]	
FTE: time: ETE							0.931	[0.802, 1.084]	
FTE: time: where							1.055	[0.944, 1.172]	
FTE: time: life							0.995	[0.894, 1.102]	
FTE: time: drugs							0.993	[0.836, 1.200]	
FTE: time: physical							0.974	[0.871, 1.094]	
FTE: time: emotion							1.122	[0.949, 1.316]	
FTE: Time: Self							1.008	[0.900, 1.119]	
FTE: time: think							1.088	[0.921, 1.271]	
FTE: time: attitude							0.940	[0.783, 1.120]	
FTE: time: change							0.904	[0.752, 1.068]	
Random effect:	**Exp(Post. Mean)**	**95 % CI**		**Exp (Post. Mean)**	**95 % CI**		**Exp(Post. Mean)**	**95 % CI**	
Individual (intercept)	0.101	[1.99E−04, 0.366]	#	1.197	[1.000, 2.007]	#	2.039	[1.000, 5.703]	#
Time	1.267	[0.338, 2.605]	#	1.134	[1.000, 1.672]	#	23.012	[1.686, 1024.54]	#
DIC	476.20			256.77			458.28		

Notes: # denotes where the adjusted posterior mean estimate of β is statistically significant ie the 95 % credible interval (CI) does not include 1. Where the interval is less 1, this suggests that as the factor decreases the likelihood of further offending behaviour. Where CI > 1, it suggests the factor increases the likelihood of further offending.

prior history of offending (the reference group) and FTEs is shown in Fig. 2 – the domain scores have been artificially fixed at 2 to reflect the average domain score across the dataset. As can be seen, the initial probability of further offending, when all other factors are equal, is not that dissimilar. However, the downward trajectory of the estimates for subsequent time points suggest that FTEs tend towards 0 at a faster rate than those who have a prior history of offending. This is broadly consistent with what would be expected. Based upon proven reoffending rates, it was expected that those with a history of prior offending would be more likely to go on to commit further offences. For example, in 2013/14, whist the overall national proven reoffending rate was 42.9 %, amongst those with no previous offences it was 24.2 %. This compares to 55.2 % who had one or more previous offences, with the published data also showing that the rate increased as the number of previous offences increases (Youth Justice Board & Ministry of Justice, 2015).

4.4. *Understanding the differences between the two sub-groups*

Despite the small numbers involved, there is moderate evidence in favour of the alternative hypothesis that the rate of further offending is higher amongst those with previous offending than FTEs (BF_{10} for the one-sided test = 5.535).

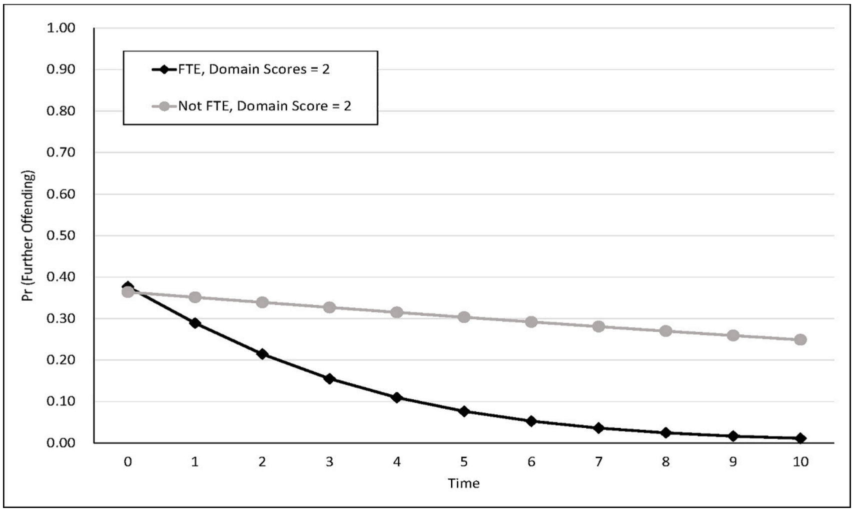

Fig. 2. Changes in the probability of further offending over time based on the basic dynamic model involving FTE status (Model 3).

As can be seen from Table 5, whilst the rates of further offending within the data are not directly comparable to the more artificial measure of proven reoffending cited above, the trends are consistent with those observed within the official figures. With fewer FTEs having gone on to commit further offences, the estimated probability of further offending over time tends towards zero quicker than for those who have a history of prior offending behaviour.

In generating the estimates of the probability of further offending summarised in Fig. 2, the domain scores were artificially set at 2. However, it is necessary to consider whether the two sub-groups would typically have the same initial domain scores representing the perceived likelihood of reoffending (Fig. 3).

Certainly T-tests comparing the average domain scores at Time 0 for the two groups suggest that whilst there is moderate evidence that there is no difference with respect to Education, Training and Employment ($BF_{10} = 0.263$ in favour of H_0: FTE = Prior), there is moderate evidence that those with a prior history of offending have higher domain scores at time of their initial assessment for:

- Living Arrangements ($BF_{10} = 12.930$ in favour of H_1: FTE < Prior).
- Family and personal relationships ($BF_{10} = 7.135$).
- Perception of Self and Others ($BF_{10} = 6.884$).

Average initial assessment scores are unlikely to be the only difference, it is also likely that other factors will also be different. Looking more closely at the composition of the two sub-groups (Table 6) highlights a key limitation of this study especially with respect to gender and ethnicity which was also alluded to in the methods section.

The small numbers of females relative to males, and of non-White relative to White children (labelled as Group 2 in Table 6) contribute to the respective Bayes factors for the two-sided tests to establish whether the respective proportion of the cohort who are FTEs is the same for each group being inconclusive.

When interpreting the Bayes factors for the one-sided tests, the inverse of 0.177 has been calculated since this can be read more intuitively. The BF_{01} (the Bayes factor for the null hypothesis) suggests that the data are 5.5 times ($BF_{01} = 1/0.177$) more likely under the null hypothesis (H_0) than the alternative (H_1). Therefore, despite the apparent difference in the rates in Table 6, there is moderate evidence to suggest that there is not a difference in the likelihood of males being FTEs when compared to females. Similarly, the BF_{01} of 2.4 (1/0.419) with respect to ethnicity suggests there is only anecdotal evidence in favour of the null that there is not a difference in the likelihood of Whites being FTEs when compared to non-Whites. Had the number of females and non-White children respectively been higher, then it potentially could have provided greater evidence in favour of the apparent trend in Table 6.

Table 5
Rates of further offending across the two cohorts, by offending history.

	Comparator group	n	Further offences n	Further offences Rate	Bayes factor (BF_{10}) (H_1: Group 1 ≠ Group 2)	Bayes factor (BF_{10}) (H_1: Group 1 < Group 2)
1	FTE	33	11	33.3 %	2.808	5.535
2	Previous Offending	54	31	57.4 %		

Notes: Bayes factors were calculated using the test for Bayesian Contingency Tables within JASP version 0.19.1. The two sets of Bayes factors represent the results of (1) a two-sided alternative hypothesis that the respective proportions of first-time entrants are equal (Alternative Hypothesis: Group 1 ≠ Group 2), and (2) a one-sided alternative hypothesis that the rates for Group 2 are larger than Group 1. Bayes factors quantify the evidence for the alternative hypothesis relative to the null hypothesis and are interpreted using the categories suggested by Jeffreys (1961) and adjusted by Lee and Wagenmakers (2013) – see Appendix 1 for further detail.

5. Discussion

The approaches presented here represent cutting edge methodological developments pertinent to both criminology and public administration researchers, providing the opportunity to challenge (or confirm) pre-exiting ideas through the use of a fresh lens. Notably, the efficiencies afforded by using the approach described here mean that smaller sample sizes are required than if employing traditional techniques. In the context of youth justice this is particularly important since it means that not only is it possible to make maximum use of the data that we already have, but moving forwards, as caseloads decrease in size but increase in complexity, we will still be able to utilise routinely collected data to learn more about the aetiology of youth offending and adapt models as new evidence emerges.

Significantly the approach described in this article demonstrates how even with a small person-period dataset (545 assessments relating to 87 individuals covering the period from 1st April 2012 to 31st March 2015) drawn from the statutory caseload of a single YOT, it is possible to gain fresh insights to inform the evidence base. Notably, the addition of further Level 2 predictors such as age, gender and ethnicity into the Basic Dynamic Model, would enable a more nuanced understanding of the impact of having a prior criminal career to be considered for

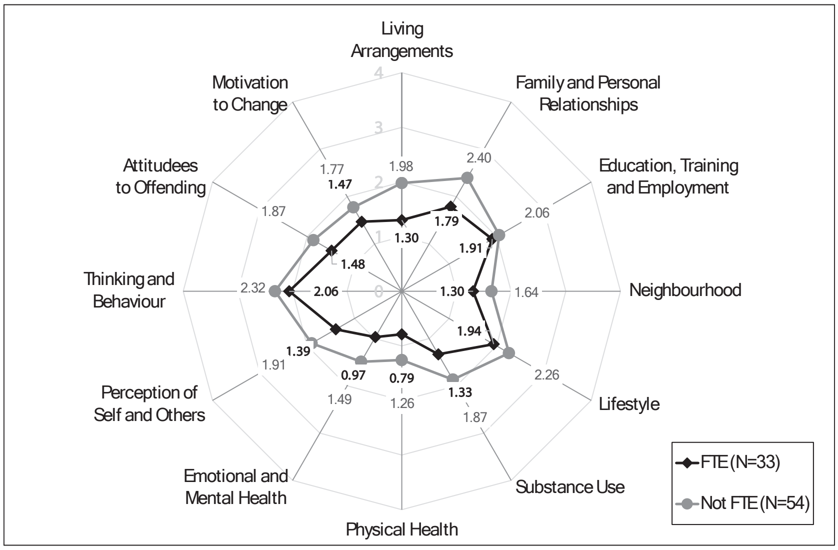

Fig. 3. Mean domain scores at time 0, by FTE status.

Table 6
The respective proportions of FTEs by gender and ethnicity.

Comparator groups			n	FTE n	FTE Rate	Bayes factor (BF_{10}) (H_1: Group 1 ≠ Group 2)	Bayes factor (BF_{10}) (H_1: Group 1 > Group 2)
Gender	1	Male	79	32	40.5 %	1.222	0.177
	2	Female	8	1	12.5 %		
Ethnicity	1	White	81	31	38.3 %	0.464	0.419
	2	Non-White	6	2	33.3 %		
Total			87	33	37.9 %		

Notes: Bayes factors were calculated using the test for Bayesian Contingency Tables within JASP version 0.19.1. The two sets of Bayes factors represent the results of (1) a two-sided alternative hypothesis that the respective proportions of first-time entrants are equal (Alternative Hypothesis: Group 1 ≠ Group 2), and (2) a one-sided alternative hypothesis that the rates for Group 1 are larger than Group 2.

different sub-groups, including – assuming the relevant data were available and systematically collected, the minority groups identified by Lammy (2017) and Laming (Prison Reform Trust, 2016). Similarly, there is scope to add alternative (or additional) non-time varying predictors which reflect the index offence or specific patterns of offending; a particular shared characteristics such as being neurodivergent, or care experienced; or draw on recently developed typologies which reflect school performance (for example, Wickersham, 2024).

Additional time-varying predictors can also be added to look at the impact of key events. These might be linked to the administration of justice such as breach proceedings being instigated in response to non-compliance with the requirements of the sentence, returning to court or spending time in the secure estate, or significant life events. By adding these as Level 1 predictors, there is scope to look at the whether the timing matters. This could be particularly pertinent about making decisions around when best to make referrals to other agencies. Since not all life events are planned, this approach may well provide new insights into why some children in conflict with the law respond positively say to starting at a new school whereas others can respond very negatively.

Whilst the findings included in this paper confirm what is already known i.e. that FTEs are generally less likely to go on to commit further offences, unpicking this to try to understand why there may be differences for example as a result of considering differences in the demographics of the two groups and in their initial domain scores, highlights the extent to which those with an offending history are significantly more likely to have living arrangements, and family and personal relationships which place them at higher risk of reoffending. Their perception of self and others is also more likely to be problematic.

A further advantage afforded by utilising hierarchical modelling under a Bayesian framework is that this approach can more effectively handle small datasets. By incorporating prior information, it is possible for rare events to be considered, such as less frequently occurring offence types as well as different permutations of individual characteristics. However, as highlighted here there can be issues with convergence and potentially undue weight being placed on a limited number of cases if numbers are particularly low. A thorough understanding of the data is therefore required it appreciate what can and cannot be achieved.

Looking forward, there is the potential to further maximise the richness of the risk assessment data through linking to other sources such as health, education, social services and court data. This would enable a more detailed model of the role different factors and their relationship with offending behaviour to be developed. Employing data linkage techniques, especially reflecting more recent offending and data from AssetPlus, would also assist criminologists in their endeavours to operationalise key concepts of interest. This will not only enable more robust analysis to be undertaken, but it will also support the investigation of new lines of enquiry which are not captured by Asset or its successor. Notably Bayes factors can be used to compare the strength of evidence for and against the use of new or alternative constructs developed to measure potential factors, with administrative data also assisting in determining temporal precedence. Whilst not possible to do here, having access to multiple administrative datasets will enable some gaps and anomalies to be investigated through the use of different sources of information.

5.1. Implications for practice

A key feature of Bayesian hierarchical approaches is that models can be updated as and when new information becomes available. As a result, they lend themselves not only to situations such as dynamic models of risk where there is a need to update the assessment as more is learnt about the individual, but also where there is scope to continue to add new variables to reflect emerging crime types and increased awareness of issues as a result of research. These can sequentially be added to the model as illustrated here rather than having to rebuild the model each time. In this way the modelling can continue to evolve to reflect both emerging research and policy concerns.

Whilst the example of FTE status has been discussed here as has the scope to use this approach for a range of different demographic characteristics, conduct, aspects of the criminal career (the static factors), and different permutations of this. Potential 'new' variables include those linked to promoting desistance and acknowledging different forms of vulnerability/need. Further, additional levels can be added to consider geographic/practice variation; the nature and/or severity of the index offence, or specific interventions. The ability to do this is limited only by the availability and completeness of the information required to operationalise the concept/metric. As part of this feasibility study, attempts were made to reflect offence type. However, the way in which the index offence was recorded in the CMS made this free text field too onerous to standardise and there were insufficient cases to produce robust models using the 1–8 gravity score.

As Child First (Case & Hazel, 2023) becomes more firmly embedded into policy and practice across the YJS in England and Wales, there is also an opportunity to move away from the existing risk-orientated model of assessment and place more emphasis on the welfare of the child. For example, the Bayesian hierarchical models could also be used to increase understandings of how individuals who share a specific set of characteristics typically respond to "events" relative to their peers. These insights can then be used to pre-empt negative outcomes through the provision of timely, appropriate support. Specifically, understanding where differences exist such as those between demographic groups or as result of prior contact with the justice system enables more sensitive decision making to be employed. Thus, it facilitates the development of more holistic assessments and targeted interventions, accelerating the shift away from the one-size-fits-all approach which has been pervasive in youth justice practice in England and Wales for too long.

On paper, the richness of the data captured within AssetPlus offers promise as a source of data to utilise when modelling relationships associated with offending behaviour and/or desistence. Not only have data been collected in a systematic basis for approaching 10 years not, but it proports to contain all the clinically relevant data required by practitioners to inform intervention planning along with screening tools for physical and mental health; speech, language and communication needs; and alcohol use (Youth Justice Board, 2014). However, one of the key challenges in adapting the approach outlined here is in determining how key concepts could be operationalised especially given the different purpose of the tool. This is something which has already been identified as an issue when it has come to evaluating AssetPlus with Smith & Paddock (2025) highlighting the challenge faced by authors of a recent YJB commissioned review of AssetPlus – the review was intended to measure 18 outcomes identified by the YJB including youth justice targets such as reducing reoffending. However, they found that 10 of these were immeasurable due to the nature of the data available. Echoing arguments previously set out by Case & Haines (2009) and others, Smith and Paddock were similarly critical of the over-simplification of concepts and use of proxy measures.

Assuming it is considered appropriate to predict further offending behaviour, care would need to be taken if attempting to replicate this approach using AssetPlus not least because the predictive utility of the tool has not been tested. Given the increasing emphasis on desistence and Child First, it may be more meaningful to seek to model the factors that increases the likelihood of desistence in order to increase the evidence base around this especially given that much of the research to date has been adult offenders.

5.2. Limitations

The primary limitation of this study is the small sample size and reliance upon data from a single YOT. As highlighted in Table 2, it was not possible to undertake robust intersectional analysis due to the low number of children from minority ethnic groups who were on the statutory caseload – this is in part a reflection of the characteristics of the population in the local area. The solution adopted here was not include predictors reflecting dimensional identity in the models and to aggregate those from minority ethnic groups in the subsequent analysis. However, this is far from ideal especially given the findings from Hunter (2023) which highlight the 'double whammy' of disadvantage of being care experienced and from certain racially minoritised groups in England.

It is possible to produce separate models which include gender/sex and ethnicity as Level 2 predictors in the same way as FTE status has been done here. However, attempts to consider the interaction between gender/sex and ethnicity as well as each of the domains and time struggle to converge. Notably where there is a class imbalance, there is also the risk that too much weight will be placed on a limited number of cases effectively introducing bias. Consideration therefore also needs to be given to the characteristics of those with a larger number of observations since there are fewer of these. For this reason, the probability of further offending behaviour has not been calculated indefinitely, rather it stops at Time = 10. Identifying these issues was only possible because of the comparatively small size of the dataset and therefore serve as key learning points for others. Notably, they highlight the need to have a thorough understanding of the strengths and limitations of the underlying data and the implications that this might have on the resulting modelling. Should the opportunity arise to replicate this approach using a larger dataset it is advocated that sufficient time be allowed for this preparatory work to be completed.

Unlike many studies conducted using administrative data, this study was conducted outside a trusted research environment with data being extracted from the YOT's CMS. Whilst having access to identifiable records and the live system afforded the opportunity to investigate some inconsistencies and populate fields which could not be pulled back through structured queries / bulk downloads, it was only feasible to do this due to the comparatively small number of cases that met the inclusion criteria. Two quirks of the CMS had a particular impact on the data: the first was that to aid practitioners, it was possible to clone the child's previous Asset. Whilst this was intended to reduce the administrative burden on the practitioner, it meant that it was necessary to remove incomplete and duplicate records. The second was that the Asset score under the scaled approach was not always accurately calculated as it was necessary to indicate if the child had committed a specified offence. If this study was to be repeated across a larger number of YOTs the impact of these quirks would significantly add to the time taken to prepare the data for analysis.

Should the opportunity arise to apply the approaches demonstrated in the paper on a significantly larger dataset, this would mean that there were sufficient cases to reserve a proportion to use as a 'test' dataset to determine the degree to which the models accurately classify those who committed further offences. However, perhaps more fundamentally there would be the opportunity to consider a greater range of interactions between characteristics and events. Ideally the analysis would be undertaken in a trusted research environment where there is the potential to link to data from across different policy areas. For this to be possible, personal identifiers would need to be accurately recorded so that a data linkage key such as an ALF can confidently be assigned.

6. Conclusion

The potential of using Bayesian approaches to mimic the rapidly changing nature of the lives of children who have come into conflict with the law and the evolving nature of youth offending more accurately has been demonstrated through this feasibility study. However, access to a larger dataset is required in order for the potential to be fully realised. The recent work by Smith and Paddock (2025) highlights some of the challenges that would be faced should attempts be made to build upon this work using AssetPlus and therefore contribute to the review and redesign of assessment processes across the youth justice system in England and Wales. These are not insurmountable given the wealth of routine data now available in trusted research environments and the wealth of experience that has been accumulated by practitioners. However, care would need to be taken when operationalising measures to avoid the oversimplification and homogenisation of trends that has previously occurred.

Declaration of competing interest

The author has declared the following financial interests/personal relationships which may be considered as potential competing interests: Helen Ruth Hodges reports financial support was provided by ESRC Wales Doctoral Training Centre.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.childyouth.2025.108502.

Data availability

The data that has been used is confidential.

References

Armstrong, D. (2004). A risky business? Research, policy, governmentality and youth offending. *Youth Justice, 4*(2), 100–116. https://doi.org/10.1177/147322540400400203

Armstrong, D. (2006). Becoming criminal: The cultural politics of risk. *International Journal of Inclusive Education, 10*(2–3), 265–278. https://doi.org/10.1080/13603110500296711

Baker, K. (2014). *AssetPlus Rationale*. London: Youth Justice Board.

Baker, K., Jones, S., Merrington, S., & Roberts, C. (2005). *Further development of Asset*. London: Youth Justice Board.

Baker, K., Jones, S., Roberts, C., & Merrington, S. (2003). The evaluation of the validity and reliability of the youth justice board's assessment for young offenders. *Oxford: Probation studies unit, centre for criminological research*. University of Oxford.

Bateman, T. (2020). The state of youth justice 2020: An overview of trends and developments. *National Association for Youth Justice*.

Bonta, J., & Wormith, J. S. (2013). Applying the risk-need-responsivity principles to offender assessment. In *What works in offender rehabilitation* (pp. 69–93). John Wiley & Sons. https://doi.org/10.1002/9781118320655.ch4.

Brown, S. (2005). *Understanding youth crime: Listening to youth?* (2nd ed.). Open Univeristy Press.

Bushway, S., & Weisburd, D. (2006). Acknowledging the centrality of quantitative criminology in criminology and criminal justice. *The Criminologist, 31*(4), 1–4. http://www.asc41.com/Criminologist/2006/July-August%202006.htm.

Case, S. (2021). Challenging the reductionism of "evidence-based" youth justice. *Sustainability, 13*(4), 1735. https://doi.org/10.3390/su13041735

Case, S., & Browning, A. (2021). *Child First Justice: the research evidence-base*. https://repository.lboro.ac.uk/articles/report/Child_First_Justice_the_research_evidence-base_Full_report_/14152040.

Case, S., & Haines, K. (2009). *Understanding youth offending: Risk factor research, policy and practice*. Willan Publishing.

Case, S., & Haines, K. (2010). Risky business? The risk in risk factor research. *Criminal Justice Matters, 80*(1), 20–22. https://doi.org/10.1080/09627251.2010.482234

Case, S., & Hazel, N. (Eds.). (2023). *Child first. Developing a new youth justice system*. Palgrave Macmillan.

Drew, J. (2023). Developing child first youth justice policy in England and wales—A view from inside the YJB and Westminster. In S. Case, & N. Hazel (Eds.), *Child first*. Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-031-19272-2_6.

Farrington, D. P. (2000a). Developmental criminology and risk focused prevention. In M. Maguire, R. Morgan, & R. Reiner (Eds.), *The oxford handbook of criminology* (3rd ed.). Oxford University Press.

Farrington, D. P. (2000b). Explaining and preventing crime: the globalization of knowledge-the American society of criminology 1999 presidential address. *Criminology, 38*(1), 1–24. https://doi.org/10.1111/j.1745-9125.2000.tb00881.x

Finch, W. H., Bolin, J. E., & Kelley, K. (2014). *Multilevel modeling using R*. CRC Press.

France, A., & Homel, R. (2016). Societal access routes and developmental pathways: Putting social structure and young people's voice into the analysis of pathways into and out of crime. *Australian & New Zealand Journal of Criminology, 39*(3), 295–309. https://doi.org/10.1375/acri.39.3.295

Garside, R. (2009). Risky individuals, risky families or risky societies? *Criminal Justice Matters, 78*(1), 42–43. https://doi.org/10.1080/09627250903385305

Gill, J., & Womack, A. (2013). In *The multilevel model framework*. SAGE Publications Ltd.. https://doi.org/10.4135/9781446247600

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: TheMCMCglmmRPackage. *Journal of Statistical Software, 33*(2), 22. https://doi.org/10.18637/jss.v033.i02

Hampson, K. S. (2018). Desistance approaches in youth justice – The next passing fad or a sea-change for the positive? *Youth Justice, 18*(1), 18–33. https://doi.org/10.1177/1473225417741224

Hunter, K. (2023). *Policy briefing: Care experience, ethnicity and youth justice involvement – key trends and policy implications*. ADR UK. https://www.adruk.org/fileadmin/uploads/adruk/Documents/Policy_Briefings/Policy-briefing-Katie-Hunter.pdf.

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association, 90*(430), 773–795.

Kelly, L., & Armitage, V. (2015). Diverse diversions: Youth justice reform, localized practices, and a 'new interventionist diversion'? *Youth Justice, 15*(2), 117–133. https://doi.org/10.1177/1473225414558331

Lammy, D. (2017). *The Lammy review: An independent review into the treatment of, and outcomes for, black, Asian and minority ethnic individuals in the criminal justice system*. London: Gov.UK.

Law, J., Quick, M., & Chan, P. (2014). Bayesian spatio-temporal modeling for analysing local patterns of crime over time at the small-area level. *Journal of Quantitative Criminology, 30*(1), 57–78. https://doi.org/10.1007/s10940-013-9194-1

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

McAra, L., & McVie, S. (2010). Youth crime and justice: Key messages from the Edinburgh Study of Youth Transitions and Crime. *Criminology & Criminal Justice, 10*(2), 179–209. https://doi.org/10.1177/1748895809360971

Ministry of Justice. (2022). *Ministry of Justice: Data First*. Gov.UK. https://www.gov.uk/guidance/ministry-of-justice-data-first.

Office for National Statistics. (2012). *DC2101EW – Ethnic group by sex by age*. NOMIS Web. https://www.nomisweb.co.uk/census/2011/dc2101ew.

Prison Reform Trust. (2016). *In Care, Out of Trouble: How the life chances of children in care can be transformed by protecting them from unnecessary involvement in the criminal justice system. Prison Reform Trust*. http://www.prisonreformtrust.org.uk/wp-content/uploads/old_files/Documents/In%20care%20out%20of%20trouble%20summary.pdf.

Robitzsch, A. (2020). Why ordinal variables can (almost) always be treated as continuous variables: Clarifying assumptions of robust continuous and ordinal factor analysis estimation methods [perspective]. *Frontiers in Education, 5*. https://doi.org/10.3389/feduc.2020.589965

Robson, K., & Pevalin, D. (2016). *Multilevel modeling in plain language*. London: Sage.

Schwalbe, C. S. (2007). Risk assessment for juvenile justice: A meta-analysis. *Law Hum Behav, 31*(5), 449–462. https://doi.org/10.1007/s10979-006-9071-7

Schwalbe, C. S., Fraser, M. W., Day, S. H., & Cooley, V. (2006). Classifying juvenile offenders according to risk of recidivism: Predictive validity, race/ethnicity, and gender. *Criminal Justice and Behavior, 33*(3), 305–324. https://doi.org/10.1177/0093854806286451

Smith, H., & Paddock, E. (2025). Does AssetPlus facilitate effective assessment of children within the youth justice system?: A critical evaluation. *Probation Journal*. https://doi.org/10.1177/02645505241300687

Snijders, T. A., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). SAGE.

Spiegelhalter, DJ, Best, NG, Carlin, BP, et al. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 64*(4), 583–616.

Stander, J., Farrington, D. P., & Lubert, C. (2022). Understanding how offending prevalence and frequency change with age in the Cambridge study in delinquent development using Bayesian statistical models. *Journal of Quantitative Criminology*. https://doi.org/10.1007/s10940-022-09544-x

Stephenson, M., Henri, G., & Brown, S. (2011). *Effective practice in youth justice*. Abingdon, Oxon: Routledge.

van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., & Yau, C. (2021). Bayesian statistics and modelling. *Nature Reviews Methods Primers, 1*(1), 1. https://doi.org/10.1038/s43586-020-00001-2

Vaswani, N., & Merone, L. (2013). Are there risks with risk assessment? A study of the predictive accuracy of the youth level of service-case management inventory with young offenders in Scotland. *British Journal of Social Work, 44*(8), 2163–2181. https://doi.org/10.1093/bjsw/bct059

Wickersham, A. (2024). *Data Insight: Changes in school performance and involvement in the criminal justice system*. https://www.adruk.org/news-publications/publications-reports/data-insight-changes-in-school-performance-and-involvement-in-the-criminal-justice-system/.

Wilson, E., & Hinks, S. (2011). *Assessing the predictive validity of the Asset youth risk assessment tool using the Juvenile Cohort Study (JCS)*. London: Ministry of Justice.

Youth Justice Board & Ministry of Justice. (2015). *Youth Justice Statistics 2013/14 – England and Wales (Supplementary Tables)*. https://www.gov.uk/government/statistics/youth-justice-statistics.

Youth Justice Board & Ministry of Justice. (2020). *Assessing the needs of sentenced children in the Youth Justice System. 2018 to 2019 AssetPlus experimental statistics*. https://www.gov.uk/government/statistics/assessing-the-needs-of-sentenced-children-in-the-youth-justice-system.

Youth Justice Board & Ministry of Justice. (2021). *Assessing the needs of sentenced children in the Youth Justice System 2019/20*. https://assets.publishing.service.gov.uk/media/604a3ee28fa8f540179c6ab7/experimental-statistics-assessing-needs-sentenced-children-youth-justice-system-2019-20.pdf.

Youth Justice Board. (2010). *Youth justice: the scaled approach – A framework for assessment and interventions*. http://yjbpublications.justice.gov.uk/Resources/Downloads/Youth%20Justice%20the%20Scaled%20Approach%20-%20A%20framework%20for%20assessment%20and%20interventions.pdf.

Youth Justice Board. (2013). *National Standards for Youth Justice Services*. Youth Justice Board. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/296274/national-standards-youth-justice-services.pdf.

Youth Justice Board. (2014). *AssetPlus Model Document*. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/364092/AssetPlus_Model_Document_1_1_October_2014.pdf.

Youth Justice Board. (2016a). Understanding and Improving Reoffending Performance: A summary of learning from the YJBs Reoffending Programme with implications for practice. Youth Justice Board. https://yjresourcehub.uk/yjb-effective-practice/youth-justice-kits/item/download/563_c323b3ab846cf1babaf74c698a02ed12.html.

Youth Justice Board. (2016b). Understanding and Improving Reoffending Performance: Annex B – What does ASSET data tell us about changes in the youth justice cohort over time? Youth Justice Board. https://yjresourcehub.uk/yjb-effective-practice/youth-justice-kits/item/download/567_5f1df82e058308c67a2ec3cdb35b294e.html.

Youth Justice Board. (2017). *Youth Justice Resource Hub: How to Reduce Reoffending by Children and Young People*. Youth Justice Board. https://yjresourcehub.uk/yjb-effective-practice/youth-justice-kits/item/469-how-to-reduce-reoffending-by-children-and-young-people.html.

Youth Justice Board. (2019). *Standards for children in the youth justice system*. Youth Justice Board. https://assets.publishing.service.gov.uk/media/6363d2328fa8f50570e54222/Standards_for_children_in_youth_justice_services_2019.doc.pdf.

Youth Justice Board. (2024). *Youth Justice Statistics: 2022 to 2023 supplementary tables*. Youth Justice Board. https://assets.publishing.service.gov.uk/media/65ba67f2c75d30000dca0ff0/Supplementary_Tables.zip.