Journal of Quantitative Criminology
https://doi.org/10.1007/s10940-022-09566-5

ORIGINAL PAPER

Trajectories of Change in Acute Dynamic Risk Ratings and Associated Risk for Recidivism in Paroled New Zealanders: A Joint Latent Class Modelling Approach

Ariel G. Stone[1] · Caleb D. Lloyd[1] · Benjamin L. Spivak[1] · Nina L. Papalia[1] · Ralph C. Serin[2]

Accepted: 23 December 2022
© The Author(s) 2023, corrected publication 2023

Abstract

Objectives Prior studies indicate risk for recidivism declines with time spent in the community post-incarceration. The current study tested whether declines in risk scores occurred uniformly for all individuals in a community corrections sample or whether distinct groups could be identified on the basis of similar trajectories of change in acute risk and time to recidivism. We additionally tested whether accounting for group heterogeneity improved prospective prediction of recidivism.

Methods This study used longitudinal, multiple-reassessment data gathered from 3,421 individuals supervised on parole in New Zealand ($N = 92{,}104$ assessments of theoretically dynamic risk factors conducted by community corrections supervision officers). We applied joint latent class modelling (JLCM) to model group trajectories of change in acute risk following re-entry while accounting for data missing due to recidivism (i.e., missing not at random). We compared accuracy of dynamic predictions based on the selected joint latent class model to an equivalent joint model with no latent class structure.

Results We identified four trajectory groups of acute dynamic risk. Groups were consistently estimated across a split sample. Trajectories differed in direction and degree of change but using the latent class structure did not improve discrimination when predicting recidivism.

Conclusions There may be significant heterogeneity in how individuals' assessed level of acute risk changes following re-entry, but determining risk for recidivism should not be based on probable group membership. JLCM revealed heterogeneity in early re-entry unlikely to be observed using traditional analytic approaches.

Keywords Joint latent class models · Re-entry · Dynamic risk factors · Recidivism prediction · Group trajectory modelling

✉ Ariel G. Stone
astone@swin.edu.au

[1] Centre for Forensic Behavioural Science, Swinburne University of Technology, Level 1, 582 Heidelberg Road, Alphington, Melbourne, VIC 3078, Australia

[2] Department of Psychology, Carleton University, Ottawa, ON, Canada

Published online: 16 January 2023

Introduction

Mandated community supervision following release from incarceration (i.e., parole) is a common correctional practice in many countries, including Australia (Australian Bureau of Statistics 2018), Canada (Public Safety Canada 2020), New Zealand (New Zealand Department of Corrections 2020), the United Kingdom (Ministry of Justice 2018) and the United States (Kaeble and Alper 2020). Although community corrections practices vary across jurisdictions, both within the United States and internationally (Cullen et al. 2017), community supervision supports successful re-entry when supervision officers focus their time with clients on addressing risk factors and barriers to desistance (Chadwick et al. 2015; Dowden and Andrews 2004). Assessment of risk for recidivism is standard practice in many jurisdictions, but it is unclear how risk assessments might support successful rehabilitative efforts (Viljoen et al. 2018). In particular, it is unclear how supervision officers should conceptualize changes in risk and respond when observing these changes. This study applies a novel analytic approach to help address this knowledge gap.

In this study, we used two types of information about individuals re-entering the community following incarceration. The first was repeated assessment of dynamic (changeable) risk factors. Theoretically, when dynamic risk is reduced, there is associated reduction in the likelihood of a new offense (Andrews and Bonta 2010). Statistically, this information is longitudinal within-individual data as each participant has multiple, time-linked assessment values. The second type of information was recidivism data, detailing if and when an individual committed a new criminal offense. Statistically, this information is single-instance time-to-event data, as each participant has one end time (time of recidivism or end of observation) and one recidivism outcome (dichotomously coded). These features yield two study objectives: describe how risk scores change over time and predict recidivism.

Joint modelling is an innovative approach that allows researchers to directly study the association between change in a longitudinal variable and time to a survival outcome. The joint modelling approach comprises a model of change in a time-dependent variable (usually a mixed model) and a model of time-to-event (usually a proportional hazard model), linked via a shared latent structure (Proust-Lima et al. 2014). In joint latent class modelling (JLCM, see Proust-Lima et al. 2014, 2017) trajectories of time-dependent variables are linked to survival outcomes via a shared latent class structure. This approach assumes that within a heterogeneous population, there are homogeneous subgroups that share both the same trajectory of a time-dependent variable and the same time-to-event. Essentially, it combines survival analysis with latent group trajectory analysis. This approach integrates latent classes into a prediction model and controls for systematically missing data when conducting trajectory analyses. Beyond statistical sophistication, the JLCM approach adds conceptual value for community corrections research. It is well-established that risk changes following re-entry (Papalia et al. 2020), but it is unclear how change varies across individuals. Whereas prior analytic approaches implicitly assumed a single-group trajectory during re-entry (e.g., Babchishin and Hanson 2020; Davies et al. 2021; Lloyd et al. 2020), this is an unlikely assumption and JLCM allows researchers to better examine individualized change while conserving data excluded in traditional group trajectory analysis.

Overview: Community Corrections, Core Correctional Practice, and Acute Risk Factors

Community corrections is often viewed as merely an alternative to incarceration or a strategy for shortening custodial sentences (Cullen et al. 2017). This reflects both growing recognition that mass incarceration is unsustainable (Petersilia and Cullen 2015) and the persistence of "nothing works" ideology that view rehabilitation efforts as ineffective (see Martinson 1974; Martinson and Wilks 1978). The latter is of particular interest for both its deleterious impact on corrections policy (Petersilia 2008; Sarre 2001) and because it is unfounded (Andrews and Bonta 2010; Gendreau and Ross 1979; Lipsey 1989, 1995).

Within the framework of a "nothing works" ideology, community supervision functions as a form of containment, with surveillance and control as the primary objectives. Indeed, without rehabilitation-focused initiatives, Bonta et al. (2008) found that supervision officers were unlikely to use time with clients to address barriers to re-entry, focusing instead on adherence to parole conditions. Control-based approaches to supervision are generally ineffective for reducing recidivism (Georgiou 2014; Hyatt and Barnes 2017; Lane et al. 2005; Petersilia & Turner 1993) and may result in increased absconding and reincarceration (Hyatt & Barnes 2017).

The Risk Needs Responsivity model (RNR, Andrews & Bonta 2010; Bonta & Andrews 2017), drawn from social learning, social control, and differential association theories, offers an alternate approach. It posits that the degree of rehabilitation services must match the empirically assessed level of risk (i.e., individuals more likely to recidivate receive higher intensity interventions), that rehabilitation should target factors that are empirically related to recidivism and appropriate for intervention, and that program delivery should be consistent with evidence-based practices and the learning/motivational style of the individual. Use of the RNR model has proliferated, and RNR-based practices have been adopted by jurisdictions internationally (Cullen 2013).

Core Correctional Practice (CCP, Dowden & Andrews 2004) is the application of the RNR model. In the context of community corrections, supervision officers implementing CCP act as facilitators of clients' process of change, or as agents of change. Within the context of a collaborative working alliance with a fair but firm approach, officers model and reinforce prosocial attitudes, facilitate concrete problem solving skills, and assist clients to access and effectively use community resources to overcome barriers to successful re-entry. In short, they identify and address risk factors. The CCP approach is associated with improved outcomes (Bonta et al. 2008; Dowden & Andrews 2004; Kennealy et al. 2012), and training supervision officers to implement these practices is associated with substantial reductions in recidivism compared to standard practice (Bonta et al. 2011; Chadwick et al. 2015).

Risk factors are personal, social, contextual, and community factors associated with a higher likelihood of recidivism. In this study, we focused on risk factors that predict recidivism for individuals already involved in the criminal justice system. Conceptually, risk factors may be divided into static and dynamic risk factors (Andrews & Bonta 2010). Static risk factors cannot be changed with intervention and are typically limited to demographic and criminal history factors, such as age of first involvement with the justice system, number of prior convictions, or severity of past convictions. Static risk factors are robust predictors of future recidivism, but by definition cannot be targets for intervention (Douglas & Skeem 2005). In contrast, dynamic factors can theoretically change over time and are promising targets for interventions.

Although all dynamic risk factors must be able to change, different risk factors change at different rates. Stable dynamic risk factors, such as antisocial attitudes, impulsivity, and poor problem solving are unlikely to change on a day-to-day basis, but may shift over months or years (Hanson & Harris 2000). In contrast, acute risk factors change rapidly. These can be internal states, such as anger/hostility (Hanson & Harris 2000; Pettersen et al. 2015; Skeem et al. 2006) and substance misuse (Horney et al. 1995; Kazemian et al. 2009; Serin et al. 2013), or situational factors, such as employment status (Berg & Huebner 2011; Hanson & Harris 2000; Kazemian et al. 2009) and relationship conflict (Mann et al. 2010; Serin et al. 2013). Acute risk factors may change over the course of days or even minutes.

Acute risk factors may act as destabilizers, interfering with adjustment and prosocial success outside of prison. Zamble and Quinsey (2001) found that 22% of recently reincarcerated men self-reported experiencing a destabilizing event in the month preceding their most recent recidivism event and that men who recidivated had higher rates of acute risk factors such as anger, substance misuse, and interpersonal conflict, relative to a sample of men who had not recidivated. In large-scale prospective studies controlling for static risk, acute dynamic risk factors predict recidivism better than stable dynamic risk factors (Davies et al. 2021; Lloyd et al. 2020).

These findings are consistent with phenomenological explanations of criminal behavior that deemphasize the role of general propensities and abstract categorizations. Rather, they posit that individuals choose actions that align with their perceived options, goals, and roles in the present situation (Lloyd 2015). Some theorists use the metaphor of a "game" (Szasz 1961), where individuals subjectively determine what outcomes constitute "winning". As such, the situational antecedents of crime behaviors (e.g., opportunities, motivations, social standing) are integral to understanding criminal actions, because antecedents set and constrain the rules of the subjective "game". Predicting recidivism in the short-term should therefore be improved by better understanding situational antecedents in general and acute risk factors in particular. In other words, because propensity-focused frameworks cannot explain periods of non-criminal behavior, frameworks focused on the immediate interaction between antecedents and subjective goals are better poised to explain the timing of recidivism (see Douglas & Skeem 2005). Methodologically, this approach requires assessing individualized experiences of potentially destabilizing antecedents and their association with criminal behavior across time. Individuals' personal understandings and perceptions of their situation are especially relevant.

Community Studies of Risk Reassessment

Longitudinal studies of dynamic risk factors (e.g., ratings of internal and external situational antecedents) in community settings are the best approach to understand how dynamic risk changes. The community setting ensures that each participant has similar opportunity to commit a new crime. Further, the greater number of reassessments included in the study, the higher the resolution for describing the change process. Community studies with frequent reassessments of dynamic risk remain relatively rare. Some notable exceptions include studies by Hanson et al. (2007); see also Babchishin & Hanson (2020), Hanson et al. (2021), Howard and Dixon (2013), Davies et al. (2021), Lloyd et al. (2020), and Stone et al. (2021). These studies have tended to focus primarily on prediction, with descriptions of risk change limited to the average trend for the total sample or for retrospectively identified subgroups (e.g., based on later recidivism status). These researchers have indirectly addressed change in risk using Cox regression survival analysis with time-varying

predictors. In this approach, recidivism is framed as the survival event in a time-to-event survival analysis with reassessments of dynamic risk framed as a time-dependent covariate. Further, these researchers have applied different specifications of dynamic risk scores (e.g., total averages across time, recent averages of scores, etc.) to evaluate model fit and effect sizes.

These studies support the conclusion that change in risk levels is important to understanding risk of recidivism by indicating updated risk assessments incrementally predict recidivism including all outcomes, violence, and sexual recidivism (Howard & Dixon 2013; Davies et al. 2021; Babchishin & Hanson 2020; Lloyd et al. 2020) with aggregated assessments particularly predicting violent recidivism (Babchishin & Hanson 2020; Stone et al. 2021) over initial time-at-release risk scores. However, it may be beneficial to more closely examine how individuals' assessed risk levels change over time. On average, levels of risk factors tend to decrease in the community whereas protective factors increase (Lloyd et al. 2020; Stone et al. 2021). Although there is clearly variability in how much each individual's level of assessed risk changes following community re-entry, it remains unclear whether individuals' risk score trajectories differ primarily in degree, or if there are qualitatively distinct patterns of change.

Group Trajectory Models

Group trajectory methods are useful for describing different patterns of change within a population. Group trajectory approaches assume that persistent unobserved heterogeneity has a discrete rather than continuous distribution (Piquero 2008). The method may be used to identify groups within a population that share similar trajectories of a variable (e.g., rate of criminal charges, levels of specific risk factors) over time. These methods are popular in the field of criminology (Piquero 2008), especially to describe patterns of criminal activity across the life course.

Research generally has substantiated that subpopulations of individuals involved in the criminal justice system follow different trajectories of criminal activity over time (Jennings & Reingle 2012; Piquero 2008). It would be theoretically consistent for trajectories of risk factor ratings (i.e., risk-related psychosocial factors) to be similarly discretely distributed in these populations. For instance, among a group of individuals paroled into the community, those with desisting trajectories would be expected to have faster declining levels of risk factors compared to those still persisting in criminal involvement, whereas those in the early stages of their criminal careers may be expected to have concurrently increasing levels of risk factors. Risk factors are not evenly distributed among individuals on supervision (Taxman & Caudy 2015) and different profiles of risk are associated with different recidivism outcomes. Zamble and Quinsey (2001) demonstrated among recently reincarcerated men that recent charges of assault, robbery, or property crimes were associated with different constellations of acute risk factors. For example, although each group of men reported alcohol misuse, alcohol was most strongly associated with assault, whereas men with recent robbery convictions reported more cocaine use and men with recent property crime convictions reported more heroin use. Affective factors especially differentiated. Men with new assault convictions reported high levels of anger in the months preceding their arrest, whereas men convicted of robbery reported general dysphoria (e.g., hopelessness, anxiety) without anger. Men convicted of property crimes reported increasing levels of frustration prior to arrest, also without high levels of anger. Importantly, Zamble and

Quinsey observed these different patterns over time, suggesting that trajectories of change in acute risk may inform recidivism and type of new charge.

However, researchers have rarely used group trajectory methods to describe patterns of change in ratings of risk factors. Two studies used samples of forensic patients (Billen et al. 2019; Quinsey et al. 2006a, b) and another two examined samples of justice-involved juveniles (Baglivio et al. 2017; Clark 2015). Each study found evidence of distinct groups with different patterns of change in various risk factors, revealing heterogeneity in trajectories of overall risk scores (Baglivio et al. 2017; Clark 2015; Quinsey et al. 2006a, b) and impulsivity and self-control (Billen et al. 2019), although the number of groups identified varied across studies. Additionally, post-hoc analyses of outcomes (e.g., ANOVA tests of differences in recidivism rates between participants designated to trajectory groups by posterior probabilities) indicated that group trajectories may be useful to understand the relationship between change in risk factors and recidivism or related outcomes (e.g., elopement, violation of conditions, violent and aggressive incidents in Quinsey et al. 2006a, b). For example, Billen et al. (2019) found that likelihood of recidivism was different across the three groups identified from trajectories of change in impulsivity ratings. Together, these studies indicated that trajectories helped to describe risk for recidivism in the context of community re-entry following either residential placement (Baglivio et al. 2017) or discharge from forensic psychiatric care (Billen et al. 2019; Quinsey et al. 2006a, b).

To our knowledge, however, no published studies have applied group trajectory methods to community-based samples of adults with criminal histories. One likely reason is non-randomly missing data. In samples of paroled individuals, recidivism often results in a return to incarceration. Once returned to incarceration, community-based risk assessment data ceases. Because higher risk for recidivism is associated with higher likelihood of return to incarceration, this non-randomly censored data introduces a high degree of bias within traditional group trajectory analyses that require uniform assessment schedules, thus over-representing individuals with lower risk levels in the data. Developing a nuanced understanding of how risk levels change across re-entry requires methods that allow retaining data from individuals who recidivate prior to study completion.

Joint Latent Class Modelling (JLCM)

JLCM presents a direct means of modelling group trajectories while accounting for non-randomly missing data by combining time-to-event survival analysis with latent class modelling. To our knowledge, this approach has not been used in community corrections research, but has been gaining traction in the field of biostatistics. For example, Smagula et al. (2019) used JCLM to control for attrition and death while modelling trajectories of daytime sleepiness in patients at risk for dementia. Proust-Lima et al. (2017) used JLCM to model trajectories of cognitive decline with time to dementia diagnosis, to create individualized dynamic predictions of when dementia diagnosis might occur. This approach can also be expanded to include multiple competing outcomes; Brilleman et al. (2019) examined trajectories of change in body mass index as these changes may have preceded either kidney transplantation or death.

JLCM undoubtedly presents a more statistically valid approach to modelling group trajectories in samples where non-randomly missing data are unavoidable. However, it is unclear whether this approach would improve understanding of community re-entry and changes in acute risk factors. This study aims to demonstrate the utility of JLCM for modelling corrections data and to test the value of JLCM as an approach to predicting

recidivism. Specifically, we aim to illustrate how latent groups describe the association between changes in assessed levels of acute risk factors and recidivism. Given prior research demonstrating average declines in acute risk (Davies et al. 2021; Lloyd et al. 2020) we expected that most participants would evidence decreasing trajectories of assessed risk, although a small proportion would evidence increasing risk. We expected that increasing acute risk trajectories would be associated with high recidivism and that decreasing trajectories would be associated with lower recidivism. However, in consideration of phenomenological explanations of crime, we expected that decreases in risk would not be uniform and would not be uniformly associated with reduced recidivism. Specifically, participants with improving risk profiles might recidivate after experiencing a destabilizing event. Such events might be reflected by subsequent increases or plateaus in the assessed risk trajectory, or might occur so close to the recidivism event as to be unobservable in participants' assessment histories. We expected a trajectory reflecting this pattern to be associated with more likely recidivism despite initially-decreasing acute risk.

Method

Participants

Beginning in April 2010, the New Zealand Department of Corrections mandated that supervision officers administer an assessment of dynamic risk (DRAOR, see 2.2.1) as part of routine practice with their caseloads of paroled individuals, re-scoring the tool at each contact of sufficient length to reassess the items. Coinciding with the implementation of this risk measure, this study used data recorded between April 1, 2010 and March 31, 2012. The data set included all individuals released into community supervision in New Zealand following incarceration during this two year period.

New Zealand law mandates that all individuals sentenced to two or more years in prison must report to a supervision officer (approximately weekly to fortnightly) for at least six months following release. Those who had served the entirety of their sentences in prison were thus supervised in the community for six months, whereas those released prior to sentence expiry were supervised for additional time, depending when parole was granted during their sentence. Although some records extended beyond 1 year, we used only up to 52 weeks of assessment data for each participant; during these 52 weeks, assessment records were largely complete up to time of censoring (i.e., weekly or fortnightly records), but assessments became less consistent as reporting requirements reduced after 52 weeks. To describe the association between recidivism and acute risk levels in the short-term, assessments must be adequately proximal to recidivism. After 52 weeks, this was rarely achievable, given sporadic assessment. The data set included 97,188 assessments completed for 3,498 participants Lloyd (2015).

Measures

Recidivism

New Zealand Department of Corrections provided official record recidivism outcomes from the beginning of data collection through July 18, 2012. Outcomes included technical violations, non-violent criminal offenses, or violent criminal offenses. Technical violations

were defined as crimes without identifiable victims (e.g., parole violations, substance use offenses). Non-violent criminal offenses included non-violent crimes against others (e.g., fraud, property destruction). Violent offenses included non-sexual violent crimes against others (e.g., assault, robbery) as well as sexual violence. In our data set, parole violations were formally prosecuted, convicted, and convictions almost always resulted in cancellation of parole (i.e., reincarceration). As such, these outcomes were not fully due to supervision officers' discretion.

Static Risk

In New Zealand Department of Corrections, a computer automatically scores static risk using the Risk of Reconviction*Risk of Re-Imprisonment (RoC*RoI) tool developed internally from the logistic regression model reported by Bakker et al. (1999). The RoC*RoI's 16 items include demographic information (gender and age) and criminal history information (e.g., criminal career length, crime frequency, and crime seriousness). We did not include age as a separate variable, as, in addition to being included in RoC*RoI's algorithm, it is strongly correlated with RoC*RoI scores in this sample ($r = -0.51$). RoC*RoI scores range from 0 to 1, with scores accurately calibrated to the likelihood of being reconvicted and reincarcerated within the next 5 years, with strong discrimination of future recidivists (AUC = 0.76; Nadesu 2007). Scores represent likelihood such that a score of 0.65 indicates a 65% likelihood of reincarceration within the next 5 years.

Dynamic Risk and Protective Factors

Supervision officers assessed dynamic factors using the Dynamic Risk Assessment for Offender Re-entry (DRAOR; Serin 2007), a 19-item case management tool conceptually organized into three subscales assessing the three domains of stable risk factors, acute risk factors, and protective factors. Assessors decide which risk factors are present and to what extent primarily based on interview information, supplemented with available file or collateral information. The DRAOR Stable subscale contains six relatively stable risk factors that are still theoretically amenable to change (i.e., attitudes, traits, and behavior patterns). The DRAOR Acute subscale comprises seven items that assess potentially rapid-changing situations (employment, living situation, interpersonal relationship conflict, and opportunity to recidivate), moods (negative mood and anger/hostility), and behaviors (substance use). Items from both DRAOR Stable and Acute subscales are scored as either *no problem* (0), *slight/possible problem* (1), or *definite problem* (2), for a maximum score of 12 and 14, respectively. Higher scores on DRAOR Stable and Acute subscales are associated with higher recidivism risk. The DRAOR Protect subscale consists of six items assessing prosocial perceptions of self and positive social connectedness. These are scored as either *not an asset* (0), s*light/possible asset* (1), or *definite asset* (2), for a maximum score of 12, with higher scores associated with lower recidivism risk.

When supervision officers use the in-session interview to assess the DRAOR Acute dynamic risk factors, they take a phenomenological approach. This is especially apparent in assessments of internal mood but is also relevant when recording situations and behaviors. For example, the DRAOR manual guides supervision officers to query individuals' perceptions of their substance use, employment, and relationships to assess personal engagement, rather than simply the presence / absence of the factor. For example, to determine the intensity of the problem, ratings are not simply based on drug tests, employment

status, or relationship status, but the individual's personal experiences pursuing sobriety, fulfilling employment, and relationship satisfaction.

Although scores on DRAOR subscales are correlated, the subscales are conceptually distinct by assessing risk of recidivism from different temporal and directional vantage points, and the tool developers currently do not recommend using a DRAOR composite score within research. For this reason, we chose not to examine composite scores. DRAOR has effectively discriminated people with future recidivism versus non-recidivism in multiple populations, including both men and women (Yesberg et al. 2015; Yesberg & Polaschek 2014, 2019). It discriminated general recidivism (Davies et al. 2021; Lloyd et al. 2020) as well as violent recidivism (Stone et al. 2021), incremental to static risk measures and initial dynamic scores assessed at time of re-entry.

Procedure

Ethics approval was obtained from institutional review boards at the following institutions: Carleton University, Swinburne University of Technology, and University of Texas at El Paso.

Supervision officers initially assessed and reassessed (approximately weekly to fortnightly) all 19 DRAOR items as part of routine practice. We excluded participants (less than 8% of the population of people supervised in the jurisdiction during the study time frame) when assessments could not be matched with a release date, if a static risk assessment was not recorded in the data set, or if the participant's initial dynamic assessment occurred more than four weeks following release from incarceration. During the study period, some participants experienced return to incarceration and re-release. Because all relevant information was updated in the data base with each re-release (i.e., age, static risk, and dynamic risk), we included subsequent re-entry sequences in our analyses as new trajectories toward recidivism. This was conceptually consistent with our definition of a trajectory as progressing from release from incarceration, reflecting a unique re-entry experience.

Given the approximate weekly assessment schedule, we chose weeks as a meaningful unit of time to measure dynamic risk trajectories. Although JLCM does not require imputation, it reduces model bias to have complete assessment schedules prior to censorship. To consider the maximum amount of data while ensuring that trajectories reflected actual scores, we imputed the most recent assessment into weeks missing an assessment, up to no more than six consecutive weeks. If seven consecutive weeks of data were missing, we truncated sequences at the last recorded assessment. Only 14.3% of sequences required imputation beyond three weeks, most missing weeks (65.7%) were single occurrences, and the majority (94.2%) of recidivism events occurred within two weeks of an assessment. Additionally, we lagged recidivism outcomes by one week to ensure that all DRAOR assessments were prospective to recidivism.

Analytic Approach

We conducted analyses on each of the three DRAOR subscales. However, we present analysis and results from DRAOR Acute only in this article because theoretically acute risk factors have unique value for predicting short-term recidivism. Further, prior studies

demonstrated that change in DRAOR Acute scores occurred faster than change in other subscales. We include results from DRAOR Stable and DRAOR Protect in appendices.

Prior to estimating latent classes for our joint model, we used mixed modelling to determine the appropriate specification for time and establish variability in individual slopes and intercepts of DRAOR Acute. We used the `lme4` package (Bates et al. 2015) for R (R Core Team 2015). For model selection, we randomly selected a calibration sample comprising approximately 60% of the full dataset ($N = 51{,}864$ observations of 2,050 re-entry sequences). Splitting the sample for model selection allowed us to test the parameters estimated and reduced computation time. We then applied the selected model to the remaining 40% of the sample. We compared proportions of participants likely to belong to each group across samples, and visually compared the predicted trajectories. Finally, we applied the final selected model to the full sample for descriptive analyses.

Modelling

We conducted joint latent class modelling using the `Jointlcmm` function of the `lcmm` package of R (Proust-Lima et al. 2017). This approach estimates a group structure that accounts for both the trajectory of a longitudinally reassessed variable (i.e., DRAOR Acute ratings) and the time to a survival event (i.e., recidivism). Individuals in the same group share similar trajectories of acute risk as well as similar time to recidivism. To be clear, individuals do not actually *belong* to any group identified by JLCM. Rather, the groups identified reflect points of support for unobserved heterogeneity in trajectories of acute risk and associated time to recidivism. We refer to participants as belonging to specific groups for the sake of parsimony, but as true latent class membership is unknown, when we refer to group membership we are actually referring to the probability that an individual can be characterized as belonging within a specific, unknowable latent class. The JLCM approach estimates these groups using (1) a class membership submodel, (2) a longitudinal submodel, and (3) a hazard submodel (see Proust-Lima et al. 2014, 2017).

Class Membership Submodel

The first regression submodel contained within the JLCM approach designates probabilistic class membership. We assume that each individual i ($i = 1, \ldots, N$) belongs to a single latent class characterized by dynamic risk trajectory and associated recidivism outcome. True latent class membership, c_i, cannot be known. We therefore model the probability that an individual i belongs to each of the possible latent classes through multinomial distribution where $P(c_i = g) = \pi_{ig}$ for latent classes g ($g = 1, \ldots, G$) (Proust-Lima et al. 2017). We included RoC*RoI scores as a covariate in the class membership model to account for demographic (including age) and criminal history differences associated with recidivism and to ensure that groups differed primarily in terms of dynamic risk.

Longitudinal Submodel

The second submodel within the JLCM approach describes the class-specific trajectories of longitudinal variables (i.e., DRAOR Acute scores). This model is identical to the standard linear mixed with the exception that it allows for class-specific distribution of both fixed

and random effects. Assuming a Gaussian outcome, the longitudinal trajectory of marker variable y for individual i at time t_{ij} ($j = 1,\ldots, n_i$) in class g is assumed to follow

$$y_{ij}|_{c_i=g} = X_{L1i}\left(t_{ij}\right)^T \beta + X_{L2i}\left(t_{ij}\right)^T v_g + Z_i\left(t_{ij}\right)^T u_{ig} + w_i\left(t_{ij}\right) + \epsilon_{ij} \tag{1}$$

(Proust-Lima et al. 2017, p.6), where X_{L1i} is the vector of covariates associated with class-common fixed effects, β; X_{L2i} is the vector of those same covariates, associated with class-specific effects, v_g; and Z_i is the vector of covariates associated with individual random effects, u_{ig}. The process $w_i\left(t_{ij}\right)$ is a zero-mean Gaussian stochastic process, and ϵ_{ij} represents normally distributed standard error.

For the longitudinal model, we included assessment occasions and intercepts as covariates, with class-common, class-specific, and random effects. We included RoC*RoI scores as class-common covariates.

Hazard Submodel

The final submodel within JLCM describes class-specific hazard for a survival event (i.e., recidivism). We used any recidivism (including reconvictions for technical violations) as a survival event.

Preliminary visual inspection indicated that our survival data followed a convex monotonically decreasing pattern, with rates of recidivism decreasing over time and cumulative hazard plots indicated relative good fit of a Weibull parametric model. We therefore chose a class-specific Weibull proportional hazard model with the RoC*RoI as a class-common covariate. In the class-specific survival model described in Eq. (2) (see Proust-Lima et al. 2017, p. 7), T_i^* denotes the time to recidivism, while $\widetilde{T}_i$ denotes censoring time (i.e., exit from the dataset in the absence of recidivism). The time of exit from the dataset is thus $T_i = min(T_i^*, \widetilde{T}_i)$, and time to event is $E_i = 1_{T_i^* \leq \widetilde{T}_i}$. Given the latent class g and survival up until time (t), the risk of recidivism can be described using the following proportional hazard model:

$$\lambda_i(t)|_{c_i=g} = \lambda_{0g}(t) e^{X_{Si1}^{\mathrm{T}} V + X_{Si2}^{\mathrm{T}} \delta_g}, \tag{2}$$

where λ_{0g} indicates the class-specific baseline hazard, X_{Si1} is the vector of covariates associated with parameters common over classes, v, and X_{Si2} is the vector of covariates associated with class-specific parameters, δ_g.

Model Selection

We compared models with varying numbers of classes, estimated from the 60% calibration sample, using a spline link function to account for higher order time effects. Our model was selected using several criteria. First, the Bayesian information criterion (BIC) was considered. BIC provides the relative marginal likelihood of competing models, penalizing free parameters (Schwarz 1978). The model with the smallest BIC is most supported by the data. Another indication of fit, relative entropy, was calculated using the methods described by Brilleman et al. (2019; equation in Appendix 1). Values of relative entropy range from 0

to 1, with 0 indicating random classification and 1 indicating perfect classification. Another consideration was the conditional independence test statistic. JLCM assumes that the latent group structure accounts for the entirety of the relationship between the trajectory variable (risk for recidivism) and the time to the survival event (Proust-Lima et al. 2014). We therefore granted preference to models in which this assumption was upheld. The proportion of participants designated to each class was also considered, in accordance with the suggestions of Nagin (2005). Solutions were not considered if any of the classes estimated contained less than 1% of the sample. Finally, also in accordance with guidelines provided by Nagin (2005), we considered mean posterior probabilities, or the likelihood that any given participant belonged to the group to which they were assigned. Solutions were only considered if the mean posterior probability for each class was above 70% (Nagin 2005).

Post Fit Analyses

Descriptive

After selecting a model, posterior probabilities were used to designate participants to the groups identified. Descriptive statistics were then compared across groups, including participant age, length of follow up time (i.e., number of weeks assessed), risk scores, and recidivism. Risk scores included mean baseline RoC*RoI and DRAOR scores, as well as mean total change and mean net change in DRAOR scores. Mean total change was calculated by subtracting final DRAOR scores from baseline scores for each participant. It describes the overall direction and amount of change in risk, such that a negative score indicates a decrease in risk or protective scores, whereas a positive score indicates an increase. In contrast, mean net changes were calculated to determine how much DRAOR scores changed on average from one week to the next, regardless of direction of change. This measure calculates the mean absolute difference between weekly scores, such that a higher number indicates greater instability from one week to the next. We also conducted paired-sample *t*-tests to assess whether participants' final DRAOR ratings were different from initial DRAOR ratings to a statistically significant degree. Frequency of recidivism, as defined in the previous text, was also compared across groups.

We conducted Kruskal–Wallis rank sum tests to check for statistically significant differences between groups in continuous variables. We used Pearson's Chi-squared test to assess differences in recidivism rates between groups.

Dynamic Prediction

Dynamic predictions were extracted from the final selected model. Dynamic predictions were computed from the `lcmm` program (see Proust-Lima et al. 2017). The dynamic prediction provides the probability that an event T_i occurs due to cause p during the prediction window $(s, s+t)$, given longitudinal data collected up until time s and a defined number of groups, θ_G. Given a single cause of event (i.e., recidivism), $\delta_i = p = 1$,

$$\mathrm{P}\left(T_i \leq s+t, \delta_i = p | T_i \geq s, Y_i^{(s)}, X_i^{(s)}, X_{Si}, X_{ci}; \theta_G\right)$$

$$= \frac{\sum_{g=1}^{G} \mathrm{P}\left(c_i = g|X_{ci};\theta_G\right)\mathrm{P}\left(T_i \in (s, s+t]|X_{Si}, c_i = g;\theta_G\right)f\left(Y_i^{(s)}|X_i^{(s)}, c_i = g;\theta_G\right)}{\sum_{g=1}^{G} \mathrm{P}\left(c_i = g|X_{ci};\theta_G\right)S_i\left(s|X_{Si}, c_i = g;\theta_G\right)f\left(Y_i^{(s)}|X_i^{(s)}, c_i = g;\theta_G\right)} \quad (3)$$

where the probability of class membership $\mathrm{P}\left(c_i = g|X_{ci};\theta_G\right)$, the class-specific cumulative incidence$\mathrm{P}\left(T_i \in (s, s+t]|X_{Si}, c_i = g;\theta_G\right)$, the density of longitudinal outcomes$f\left(Y_i^{(s)}|X_i^{(s)}, c_i = g;\theta_G\right)$, and the class-specific survival function $S_i\left(s|X_{Si}, c_i = g;\theta_G\right)$ are conditional on covariate scores at time s (Eq. (3) from Proust Lima et al., 2017, p. 24).

We selected prediction windows $(s, s+t)$ that reflected early milestones in re-entry (4–12 weeks, 12–20 weeks, and 24–32 weeks). We chose to use narrow prediction windows to mirror the timescale over which supervision decisions are generally made.

The dynamic predictions from the selected model were compared to predictions from an equivalent joint model with no latent group structure (see Appendix 2). These models were fit using `JM` (Rizopoulos 2010).

Using methods described by Blanche et al. (2015), we used the dynamic predictions to calculate dynamic AUCs and Brier scores for the specified prediction windows. Dynamic AUCs quantify a model's discriminative ability, such that a higher AUC indicates a greater ability to discriminate between participants who recidivate and participants who do not, during the specified prediction window. The Brier score reflects calibration, or how close the estimated risk score is to true underlying risk, such that a lower Brier score indicates a more accurate model. AUCs and Brier scores from the selected models were compared to those from the joint models with no latent class structure to test if the addition of a latent group structure improved prediction.

Results

Descriptive Statistic of the Full Sample

After eliminating assessments in the data set for the reasons described earlier (representing $N = 77$ individuals, see Lloyd (2015), 92,104 observations of 3,421 individuals remained. These participants were primarily male (92.8%). Ages ranged from 17 to 86 years, with a mean age of 35.1 ($SD = 11.7$). Follow up times for individual participants ranged from one to 52 weeks ($M = 25.2$, $SD = 17.0$), at which point we truncated the follow up.

Of these individuals, some were returned to incarceration before being re-released into community supervision. There were 207 participants with two assessment sequences and 13 with three assessment sequences. As a result, there were a total of 3,648 independent sequences for analysis. Forty-two percent ($N = 1{,}535$) of the sample had any recidivism recorded. With overlap, convictions for technical violations were recorded for 30.4% of the sample ($N = 1{,}108$), nonviolent criminal offenses were recorded for 11.2% of the sample ($N = 407$), and non-sexual violent or sexual recidivism was only observed in 7.6% of the sample ($N = 279$).

The mean RoC*RoI score was 0.51 ($SD = 0.24$), representing 51% likelihood of returning to incarceration within 5 years. We calculated averages of each participant's entire history of DRAOR reassessments and computed sample means of these averages. The mean average score for DRAOR Stable was 5.70 ($SD = 2.64$), representing the equivalent of almost three out of six stable risk factors rated as *definite problem*. The mean average score

for DRAOR Acute was 4.99 ($SD = 2.18$), representing two to three *definite problem* ratings out of seven acute risk factors or five factors rated as *possible problem*. The mean average score for DRAOR Protect was 6.78 ($SD = 2.43$), representing the definite or partial presence of three to six strength factors. Although participants' risk scores showed statistically significant change using each DRAOR subscale (DRAOR Stable: $t = 22.52$, $p < 0.01$; DRAOR Acute: $t = 31.97$, $p < 0.01$; DRAOR Protect $t = -26.98$, $p < 0.01$), DRAOR Acute scores tended to change more than other subscale scores during the follow up. On average, DRAOR Acute scores decreased by 1.32 ($SD = 2.50$) between first and final assessments, whereas DRAOR Protect scores increased by 1.04 ($SD = 2.32$), and DRAOR Stable scores decreased by only 0.89 ($SD = 2.38$). DRAOR Acute scores also demonstrated more week-to-week fluctuation, with mean net change of 0.38 ($SD = 0.50$). DRAOR Stable scores had a mean net change of 0.24 ($SD = 0.51$) and DRAOR Protect scores had a mean net change of 0.23 ($SD = 0.51$).

Assumption Checks

Mixed modelling indicated that there was statistically significant variation across individual slopes and intercepts. Thus, it was appropriate to proceed with group-based trajectory analysis. A full discussion of the mixed modelling process is presented in Appendix 3.

Model Selection

Joint latent class models fitted on the 60% calibration sample are presented in Table 1. See Appendices 4–5 for a discussion of selection of initial values for model estimation and results from DRAOR Stable and Protect subscales.

When modelling trajectories of DRAOR Acute, BIC values continued to decrease as classes were added to the model. This is common for group-based trajectory analyses, especially when sample sizes are large, and has been presented as a criticism of group-based trajectory analyses (Erosheva et al. 2014; Piquero 2008). Relative entropy was highest for the three-group solution, but the conditional independence test statistic was highly statistically significant. The four-group solution did not violate the assumption of conditional independence, and because the mean posterior probabilities fell below 0.70 with the addition of a fifth group, we selected the four-group model.

The solution indicated by the 60% calibration sample was then applied to the remaining 40% of the sample. Visual comparison indicated that intercepts and slopes of the classes identified remained consistent across samples, and class proportions were similar (Appendix Fig. 3). Importantly, groups that represented only a small proportion of participants were reliably estimated across samples, indicating that these were not merely artefacts of the latent group estimation process. The four-group solution for DRAOR Acute was therefore upheld and applied to the full sample. See Appendix 6 for a more complete discussion of this process.

Descriptive Statistics of the Selected Model

The four groups of trajectories of DRAOR Acute scores differed in the direction and degree of change and had different associated hazard for recidivism (see Fig. 1). Note that the JLCM approach explicitly includes survival time in the class structure. Therefore, the

Table 1 DRAOR acute model selection criteria derived from 60% calibration sample

Model	BIC	Max LL	CI test statistic	Relative entropy	Mean posterior probability within class (Percentage)				
					1	2	3	4	5
DRAOR acute									
2 classes	167,636.87	− 83,738.37	1.65	0.51	0.88 (50.59)	0.83 (49.41)			
3 classes	167,432.05	− 83,613.08	38.77*	0.68	0.90 (72.39)	0.70 (14.73)	0.73 (12.88)		
4 classes	167,360.2	− 83,554.28	3.51	0.64	0.74 (38.59)	0.74 (8.39)	0.83 (49.61)	0.82 (3.41)	
5 classes	167,298.92	− 83,500.76	4.45	0.66	0.66 (4.34)	0.83 (3.27)	0.74 (9.41)	0.85 (50.15)	0.70 (32.83)

Calibration sample comprised 51,633 observations of $N = 2050$ randomly selected participants. There were 787 recidivism events recorded in this sample

DRAOR Acute = Dynamic Risk Assessment for Offender Re-entry (Serin 2007), Acute subscale

BIC indicates Bayesian information criterion. Max LL indicates maximum log-likelihood. CI test statistic indicates conditional independence test statistic

*Indicates significant conditional independence test statistic

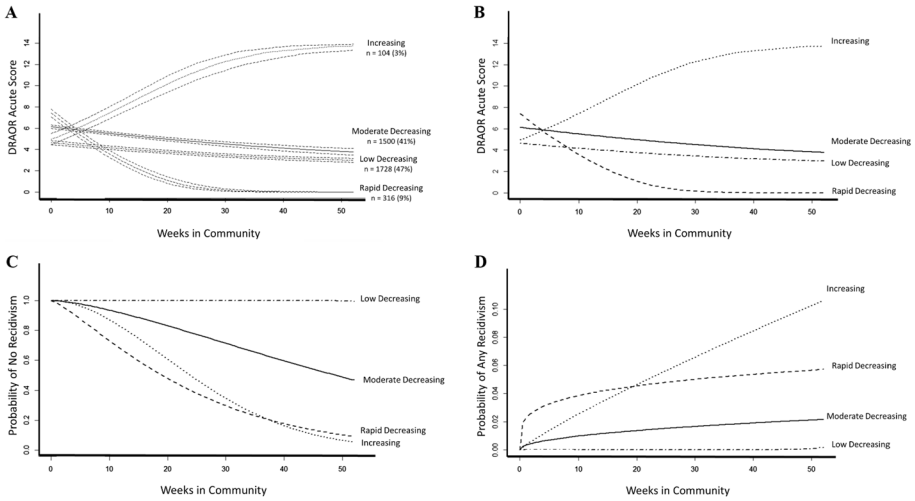


Fig. 1 Mean group predicted trajectories of DRAOR Acute plotted with (**A**) and without (**B**) Monte Carlo confidence intervals, mean group survival curves (**C**) and group-specific baseline hazard rates (**D**). Predicted values from `lcmm`'s `Jointlcmm` function (see Proust-Lima et al. 2017, p. 22). DRAOR Acute = Dynamic Risk Assessment for Offender Re-entry (Serin 2007), Acute subscale. Scores range from 0 to 14 with higher scores indicating higher risk

relationship between class membership and recidivism is descriptive rather than predictive in this context. The association between class membership and recidivism only becomes predictive when probability of class membership informs the baseline hazard of the survival model for prospective predictions as described in the Dynamic Prediction subsection of the Methods and applied in Dynamic Predictions and Comparison of Predictive Ability subsection of Results.

The first group was characterized by moderate initial DRAOR Acute scores (equivalent of three to four of seven risk factors rated as *definite problem*) that decreased over time (*moderate decreasing* group; about 41% of sample; see Fig. 2 for sample trajectories within groups). The second group of trajectories was characterized by low initial DRAOR Acute scores (equivalent of two to three *definite problem* ratings) that declined slightly over time (*low decreasing* group; about 47% of sample). Trajectories in the third group started with moderate/high DRAOR Acute scores (equivalent of four to five *definite problem* ratings) that declined rapidly (*rapid decreasing* group; about 9% of sample). Finally, a fourth group of trajectories was characterized by low initial DRAOR scores (equivalent of two to three factors rated as *definite problem*) that increased over time (*increasing* group; about 3% of sample).

Descriptive statistics for these groups are presented in Table 2. Tests indicated statistically significant group differences in each domain. Participants in the *low decreasing* group were older than participants in other groups, were assessed in the community longer, and had far lower rates of any recidivism than any of the other groups. They had low initial risk levels as indicated by RoC*RoI and baseline DRAOR ratings, and they improved across subscales while maintaining low mean net change.

The three remaining groups had similar levels of static risk indicated by RoC*RoI scores. The *moderate decreasing* and *rapid decreasing* groups started with higher initial

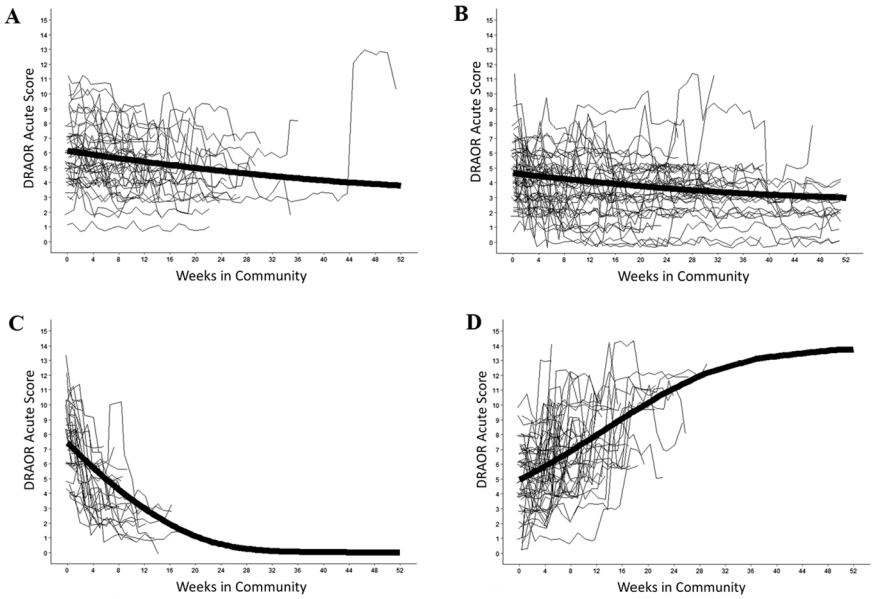


Fig. 2 Heavily weighted line represents predicted mean trajectories of DRAOR Acute for **A** *Moderate Decreasing* (41% of participants), **B** *Low Decreasing* (47% of participants), **C** *Rapid Decreasing* (9% of participants), and **D** *Increasing* (3% of participants). Unweighted lines represent 50 randomly selected individual sample trajectories within each group, jittered to reduce overlap. Jittering increases readability, facilitating illustration of within-group noise, but creates appearance of oscillation, where scores may actually be constant across measurement occasions. DRAOR Acute = Dynamic Risk Assessment for Offender Re-entry (Serin 2007), Acute subscale. Scores range from 0 to 14 with higher scores indicating higher risk

DRAOR ratings of risk relative to other groups. Although participants in the *rapid decreasing* group demonstrated more improvement in risk levels across DRAOR subscales relative to the *moderate decreasing* group (a decrease equivalent to two to three risk factor ratings decreasing from *slight/possible problem* to *no problem* compared to a decrease equivalent to one to two risk factor ratings decreasing from *slight/possible problem* to *no problem*), they had higher rates of recidivism and were assessed in the community for shorter periods of time. Of all groups, the *rapid decreasing* group had the highest rates of criminal recidivism and violent recidivism.

Participants in the *increasing* group had high recidivism rates, similar to the *rapid decreasing* group. Although DRAOR scores were moderate at baseline, RoC*RoI scores indicated that these participants were high risk at time of re-entry. This was the only group to demonstrate deterioration (i.e., increasing risk ratings) across DRAOR subscales over time (an increase equivalent to three risk factor ratings increasing from *no problem* to *slight/possible problem*). Descriptive statistics of model-selected DRAOR Stable and Protect groups are presented in Appendices 7–8.

Dynamic Predictions and Comparison of Predictive Ability

Although the joint latent class models demonstrated good predictive discrimination (AUCs ranging from 0.68 to 0.74), when compared to equivalent joint models with no latent class structure, the joint latent class model selected did not demonstrate any advantages in model

Table 2 Descriptive qualities of four DRAOR acute trajectories

	Moderate decreasing	Low decreasing	Rapid decreasing	Increasing	
	N (%)				
Number of Participants	1500 (41.12)	1728 (47.37)	316 (8.66)	104 (2.85)	
	Mean (SD)				Kruskal–Wallis χ^2
Age	31.63 (9.42)	38.57 (12.58)	30.96 (9.26)	30.35 (8.76)	315.89***
Weeks assessed prior to censoring or recidivism	17.23 (11.44)	36.07 (15.72)	7.29 (6.37)	14.17 (6.38)	1420.8***
RoC*RoI	0.64 (0.17)	0.37 (0.23)	0.62 (0.20)	0.62 (0.17)	1049.2***
Baseline stable	7.07 (2.39)	5.50 (2.36)	7.84 (2.45)	6.82 (2.56)	427.64***
Baseline acute	6.57 (2.52)	5.04 (2.12)	8.19 (2.35)	4.89 (2.24)	596.32***
Baseline protect	5.50 (2.28)	6.75 (2.25)	4.63 (2.51)	5.80 (2.35)	325.39***
Change stable	− 0.44 (2.26)	− 1.39 (2.39)	− 0.85 (2.17)	0.87 (2.32)	205.82***
Change acute	− 0.94 (2.28)	− 1.70 (2.37)	− 2.50 (2.50)	3.06 (1.91)	374.17***
Change protect	0.55 (2.21)	1.57 (2.30)	0.94 (2.23)	− 0.61 (2.17)	240.67***
Mean net change stable	0.32 (0.62)	0.12 (0.21)	0.35 (0.79)	0.46 (0.65)	83.33***
Mean net change acute	0.48 (0.59)	0.24 (0.24)	0.54 (0.71)	0.85 (0.80)	321.32***
Mean net change protect	0.32 (0.67)	0.13 (0.17)	0.34 (0.66)	0.46 (0.67)	53.78***
	N (%)				Pearson's χ^2
Any Recidivism	1061(70.73)	123 (7.12)	265 (83.86)	86 (82.69)	1668.6***
Technical violations	763 (50.87)	79 (4.57)	197 (62.34)	69 (66.35)	1058.2***
Nonviolent criminal recidivism	284 (18.93)	23 (1.33)	78 (24.68)	22 (21.15)	328.65***
Violent recidivism	182 (12.13)	21 (1.22)	53 (16.77)	12 (11.54)	189.86***

$N = 3648$ participants, assigned to groups based on posterior probabilities

RoC*RoI = Risk of Reconviction* Risk of Reimprisonment (Bakker et al. 1999)

DRAOR = Dynamic Risk Assessment for Offender Re-entry (Serin 2007); Stable = DRAOR Stable subscale (0–12 points possible); Acute = DRAOR Acute subscale (0–14 points possible); Protect = DRAOR Protect subscale (0–12 points possible)

Baseline refers to the assessment closest to time of return from incarceration

***p-value < 0.001

Table 3 Predictive discrimination and calibration of DRAOR Acute using selected joint latent class model and equivalent model without latent class structure

Model	Prediction window					
	4 weeks through 12 weeks		12 weeks through 20 weeks		24 weeks through 32 weeks	
	AUC*100	Brier score*100	AUC *100	Brier score*100	AUC *100	Brier score*100
Four class JLCM	72.47	11.62	67.50	14.18	74.36	10.04
Equivalent SREM	73.92	10.85	74.77	11.11	76.47	9.77

DRAOR Acute = Dynamic Risk Assessment for Offender Re-entry (Serin 2007), Acute subscale

JLCM = joint latent class model

SREM = shared random effects model

AUC indicates area under receiver operating characteristic (ROC) curve

Survival predictions from JLCM used Weibull baseline hazard stratified on probable group membership with Risk of Reconviction* Risk of Reimprisonment (RoC*RoI; Bakker et al. 1999) as covariate. Survival predictions from SREM used Weibull baseline hazard with RoC*RoI as covariate

calibration or in discriminating between recidivists and non-recidivists. In all three prediction windows, the joint model with no latent class structure outperformed the selected joint latent class model, in both discrimination and calibration, with differences being greatest for the second prediction window (12 through 20 weeks, see Table 3). The predictive ability of selected DRAOR Stable and Protect models are presented in Appendix 9.

Discussion

Previous analyses of dynamic risk across community re-entry in this and other similar samples have indicated that risk levels generally decrease over time (Davies et al. 2021; Lloyd et al. 2020; Stone et al. 2021). The present results confirm this was the case for most participants but decreases in risk scores were not uniform and were not uniformly associated with recidivism. Although most individuals evidenced more gradual reduction in risk, JLCM revealed a small group of individuals who demonstrated relatively more extreme variability characterized by rapid short-term improvement. However, this group also recidivated at a high rate shortly after reentry. Other analytic approaches would be unlikely to capture this heterogeneity. Additionally, there was a small but reliably estimated trajectory with relatively rapidly increasing risk scores. Likely membership in this group was also associated with high rates of recidivism.

Should assessors and justice agencies consider trajectories of acute dynamic risk factors when evaluating risk for recidivism? Our results do not support trajectory methods as means to discriminate between recidivists and non-recidivists in practice. Indeed, stratifying hazard estimates on probable latent group membership resulted in poorer predictive discrimination. Failure to improve discrimination is perhaps unsurprising, as Lloyd et al. (2020) found scores representing demographics and criminal history (i.e., RoC*RoI) drive predictive discrimination in this sample (although DRAOR Acute scores contributed incrementally to prediction). Supervision officers should not use dynamic DRAOR scores as the sole determinant when clients are at risk of criminal behavior or breaching supervision

conditions. Rather, DRAOR should be used as a case management tool to identify specific areas of risk among clients known to be high risk based on a static risk assessment, then use targeted interventions and referrals to ameliorate these risks.

Why is it important that trajectories of risk varied in this sample, if those differences did not predict recidivism? On one hand, these findings largely substantiate what is already known about risk for recidivism: dynamic risk generally decreases over time, and older participants with lower static risk are less likely to recidivate (see specifically the *low decreasing* group). However, the results also provide unique evidence supporting theories emphasizing multiple pathways to recidivism. Specifically, whereas there appeared to be only one pathway towards desistance in this sample, the pathways to recidivism varied. As suggested by Zamble and Quinsey (2001), different series of events, internal or external, may lead to criminal behavior. This requires supervision officers to be vigilant about considering varied client-informed, subjective motivations and goals that may arise from multiple situational triggers while avoiding assumptions about singular pathways and pitfalls. This is also consistent with a framework that hypothesized eight unique pathways to sexual criminal behavior based on different patterns of affect leading to the sexual behavior (Hudson et al. 1999). As such, there is no uniform, single risk factor or pattern of re-entry experiences that will signal recidivism for all individuals. In our sample, underlying differences in a collection of risk factors gave rise to three trajectory groups with high rates of recidivism.

Unfortunately, addressing and reducing acute risk factors may be insufficient to reduce recidivism in the absence of other simultaneous changes. From a coping-criminality perspective, for example, a decrease in acute risk factors (e.g., stability through housing and employment) may mean little for long-term desistance if not accompanied by improved coping or problem solving when encountering inevitable stressors. Many desistance theories emphasize that change needs to occur at a deeper intrapersonal level beyond situational adjustments. For example, interactionist theories of desistance posit that engagement in prosocial institutions facilitates desistance when accompanied with concomitant changes in identity and agency (Weaver 2019; see also LeBel et al. 2008). Further, interactionist theories recognize that the transition from criminal behavior to desistance is not linear, but that episodes of crime occur despite gradual progress towards desistance (Bottoms et al. 2004). Maruna (2001) identified "condemnation scripts" in his interviews of individuals whose desistance efforts did not lead to desistance: attempts to "make good" eventually gave way to resignation and beliefs that recidivism was inevitable. Perhaps for participants in the *moderate decreasing* or *rapid decreasing* groups, effortful reductions in acute risk (e.g., avoiding substance use, seeking employment, avoiding romantic conflict) eventually proved unsustainable, or behavioral efforts were insufficient in the absence of underlying changes in desistance belief structures (see Lloyd & Serin 2012).

Although a re-entry stress-coping framework (Zamble & Porporino 1988) suggests recidivism can occur when legitimate efforts toward decreasing one's risk factors start to become outstripped by stressful events and available coping resources, we did not expect similar survival curves for the *rapid decreasing* and *increasing* groups. This result indicates an important contribution of the joint latent class modelling approach. Namely, JLCM allows observation of patterned differences in how long individuals remain and are assessed in the community. Given existing approaches that either do not control for recidivism when examining patterns of change in assessed risk or assume a uniform relationship between risk and recidivism across participants, it is unsurprising that the pattern observed in the *rapid decreasing* group has not been previously observed. The group is unusual, not only in their high rates of recidivism despite apparent improvement but in

the rapidity with which they recidivated (within the first 2 months of re-entry on average). Data from these participants would typically be removed from analyses using a two time-point design or even a traditional group-based trajectory panel design, as these designs would generally span longer timeframes and do not allow retention of data from censored cases. Conversely, using a hazard model that assumes a uniform relationship between risk and recidivism would absorb the variability contributed by this small group with relatively few assessments into the larger group. JLCM is uniquely able to parcel out heterogenous information about risk and recidivism in the weeks immediately following release from incarceration. This information may have important implications for how to best contextualize instability and support individuals with multiple situational, behavioral, and affective risk factors early in their re-entry process.

Why would a group that appears to be improving have similar outcomes to a group that appears to be deteriorating? Results are surprising if one examines change in acute risk in isolation but are less surprising when one considers initial acute risk and static risk. Participants with *rapid decreasing* acute risk had, on average, identical static risk and potentially twice as many or twice the intensity of acute risk factors at the time of re-entry relative to those with *increasing* acute risk. Participants in the *rapid decreasing* group returned from incarceration to lives characterized by initial instability, indicated by multiple situational, behavioral, and emotional risk factors. The brief time they spent in the community suggests that, despite some improvement, these participants never quite achieved stability. Instead, their rapidly decreasing risk scores indicate a different form of lifestyle instability as they experienced many simultaneous changes in multiple important life areas.

Still, it is possible that supervision officers did not accurately assess risk for a small proportion of clients characterized by a unique presentation. Supervision officers may have overestimated improvements in the *rapid decreasing* group, or these clients' risk may have been most closely tied to risk factors unassessed on DRAOR. This latent trajectory is also consistent with supervision officers using more accurate file information when initially assessing DRAOR, but needing to rely on inaccurate information from interview and collateral information for subsequent assessments, such that clients who are better able or more motivated to obscure their community functioning may appear to improve, despite remaining high risk. This may indicate a weakness in our measure of acute risk factors, as supervision officers may not adequately account for response bias when scoring acute items. Alternatively, the pattern observed in the *rapid decreasing* group might reflect instability (or even regression to the mean) in a small, high-risk group. It is possible that the high fluctuation in weekly ratings (as evidenced by mean net change) contributed to increased recidivism rates in both *rapid decreasing* and *increasing* groups. Even as individuals may be trying their best to "try on" a new identity as a desisted person, fluctuation may represent affective dysregulation, faltering attempts at abstinence from substance use, or unstable relationships (including prosocial relationships). This is consistent with findings that higher variability in acute risk is associated with higher rates of recidivism (Davies 2019; Stone et al. 2021). It also accords with several criminological and desistance theories that suggest circumstances conducive to stability (e.g., consistent engagement in employment, meaningful investment in relationships with prosocial others, ongoing efforts to maintain sobriety, etc.) are also less conducive to criminal behavior (e.g., strain theory, control theory; Bonta & Andrews 2017), or that instability and poor coping together lead to criminal behavior (e.g., coping-criminality hypothesis; Zamble & Porporino 1988).

Both explanations, and the patterns found across groups, have implications for supervision practice. First, the results support that acute dynamic risk factors should be considered incrementally to and in the context of static risk when predicting recidivism (Brown et al.

2009; Serin et al. 2016). Supervision officers should be cautious discounting an extensive history of serious criminal behavior on the basis of potentially superficial current improvements, especially in the first months of re-entry. Second, results indicate consistent, incrementally achieved change over time is associated with improved outcomes. Effective supervision should include facilitating overall stability characterized by gradual improvements, as this characterized the single trajectory most closely associated with desistance. For example, coaching clients how to keep a job and problem-solve through stressful temptations to give up (i.e., maintaining existing strengths) may be as important as initially helping a client to find a job. In practice, rather than focusing on patterns of change as a means to identify individuals at high risk, it may be more useful to use steadily decreasing change over sufficient time to identify desisting individuals for whom it would be appropriate to lower supervision intensity.

Limitations

The primary limitations of our study are common to applied research. We were limited to observing only the outcomes recorded by law enforcement and criminal justice agencies and using only the risk factor information provided by corrections staff, potentially colored by staffs' personal bias or participants' response bias. These data are important, as they provide the basis for case management decisions; however, these data cannot tell the entire story of desistance or recidivism, and assessors should always understand that there are individual variations masked by aggregate data.

We relied on officially recorded recidivism as an outcome. This limitation is twofold. First, law enforcement does not detect or respond to all criminal behavior equally. Multiple external factors and structural inequality influence patterns of detection and prosecution. Second, a brief timeframe and a binary recidivism outcome cannot adequately describe the process of desistance. Even those who are reincarcerated with new charges may already be making substantial changes towards desistance through reduced frequency or severity. It is also important that most of the participants who recidivated did so with convictions for technical violations. Although these violations merited prosecution and reincarceration in this jurisdiction, results may not generalize to jurisdictions where technical violations do not usually result in parole cancellation.

Another limitation of our data is that supervision officers only assessed many participants for 6 months consistent with New Zealand's policies for mandatory supervision. JLCM is well-suited to right censoring, and tests indicated that results did not differ when we restricted our analyses only to six months of data. Yet, our decision to truncate data once assessments became inconsistent led us to censor all cases at 52 weeks and this limited the scope of this study. However, prior studies in New Zealand indicated approximately 60% of recidivism events observed over several years occurred within the first year in the community (Nadesu 2007). Although longer timeframes may be additionally informative, high first-year recidivism rates and low assessment rates beyond 52 weeks suggest our results meaningfully describe core risk-recidivism processes.

Our decision to focus on only acute risk factors may be an additional limitation. In particular, as the JLCM design intends covariates to be time-independent, we chose not to include other time-dependent risk dimensions as covariates. DRAOR Stable scores did not predict recidivism incrementally to RoC*RoI and DRAOR Acute in this sample (Lloyd et al. 2020), but it is still possible that variation in group trajectories is related to stable dynamic risk scores. A multi-factorial group trajectory model is beyond the scope of the

current study but may be of interest to future researchers. Further, our use of a summary score obscures changes at the item-level, introducing another limitation. Our analyses cannot reveal the extent to which change in risk scores indicated changes in the variety of risk factors or their intensity.

Additional limitations concern model selection. As we constructed our models, we found that relatively minor changes to model specification often resulted in changes in the balance of model selection criteria, such that a different optimal number of groups was indicated. This is consistently cited as a critical problem with group-based trajectory models (D'Unger et al. 1998; Erosheva et al. 2014; Roeder et al. 1999; Sampson et al. 2004). Indeed, as selection criteria rarely converged on a single solution, model selection seemed disconcertingly arbitrary. It is probable that there is no single real or correct number of latent groups to be identified within our sample. This is not an indictment of the statistical concept of latent classes, but rather demonstrates a need for transparency and clarity when presenting results of group-based trajectory models. *The groups identified are defined by the trajectories and survival times of the individuals observed, and not* vice versa. This echoes Nagin and Tremblay's (2005) discussion of misconceptions of how group-based trajectory modelling functions, particularly the misconceptions that individuals belong to trajectory groups and that the number of groups is immutable. Testing trajectory estimations across a split sample confirmed that the trajectories were consistently identifiable in our data. Plotting during the selection process (Appendix 9) suggested that trajectories identified in a model with four latent classes were still present when a fifth class was added. However, baseline risk functions based on these latent classes did not improve predictive discrimination or model calibration, compared to risk functions based on a single distribution of random effects. Researchers must understand the latent groups as the identification of patterns of unobserved heterogeneity within our data, a starting point for further examination. They should exercise caution in their communication of this, especially with stakeholders less versed in statistical concepts, being careful not to imply that the directional pattern of scores at any one point in time in practice are meaningful for labelling individuals, predicting an individual's future risk trajectory, or predicting an individual's subsequent recidivism.

Future Directions

Our study had the advantages of a large sample with frequent reassessment recorded in close proximity to recidivism outcomes. It will be informative to recreate this study in similar corrections samples. The identification of similar trajectory groups in different samples would suggest that these patterns of heterogeneity are common to the re-entry process, rather than unique to this specific cohort.

If trajectories are replicated in new samples, it will then be helpful to identify potential sources of heterogeneity. Within DRAOR subscales, it may be useful to examine if specific risk items are driving differences in trajectories. For instance, in the group identified as increasing in acute dynamic risk, it would be helpful to examine if these increases are due to deterioration in situational factors (e.g., employment, relationship conflict), affective factors (e.g., negative mood, anger) or both. Additionally, it would be informative to examine if within-group changes in overall acute scores are due to changes in the number of risk factors observed, their apparent intensity, or both. Researchers could also test whether certain stable risk factors (e.g., impulsivity, attitudes towards authority) moderate the relationship between acute risk and recidivism. Further heterogeneity might be explained by

examining file information beyond RoC*RoI summary scores (e.g., specific dimensions of criminal history) or by including information on the supervision officers conducting assessments (i.e., nesting observations conducted by specific supervision officers). Our data set did not include indicators that identified different assessors.

The question of stability in risk should be examined more closely. Mean net change provides a general measure of how much scores varied from week-to-week but does not indicate direction of change or consider how fluctuation varies over time. Methods designed specifically to assess stability, such as P(Δ) (Barnes et al. 2017) or damped oscillator models (Chow et al. 2005) may be informative.

Future directions also include expanded application of dynamic predictions, and exploration of dynamic predictions from joint models using a shared random-effects approach. Potential applications include comparing predictive accuracy of different risk models or comparing models' discriminative ability across different prediction windows. Eventually, individualized dynamic predictions may be a useful tool for identifying individuals requiring reductions in the amount of control imposed by their community supervision.

Appendix 1: Equation for Relative Entropy

The equation and syntax for calculating relative entropy were provided by Brilleman (2019 p.42; personal communications). Relative entropy is calculated as

$$1 + \frac{\sum_{i=1}^{N} \sum_{g=1}^{G} \widehat{\pi}_{ig} \log(\widehat{\pi}_{ig})}{N \log G},$$

Where $\widehat{\pi}_{ig}$ is the estimated posterior probability of individual i ($i = 1 \ldots, N$) being in latent class g ($g = 1 \ldots, G$).

Appendix 2: Estimation of Equivalent Joint Models Without Latent Class Structure

Shared random-effects modelling (SREM; see Rizopoulos 2012) is an alternative approach to joint modelling that assumes a homogenous distribution of hazard. In this approach, random effects from a longitudinal model are included as covariates of a survival model.

In order to account for error in measurement and incomplete knowledge of the full trajectory of the longitudinal variable (dynamic risk in our project), SREM joint models use observed values to estimate the true values of the longitudinal variable. These are used to reconstruct a complete longitudinal history for each individual to inform the hazard model.

We tested predictive ability of SREM models with identical covariates, link functions, and baseline hazard distributions as our JLCM models.

Appendix 3: Assumption Testing: Variability of Slopes and Intercepts

A fundamental criticism of group-based trajectory analysis is that statistical programs are likely to identify multiple trajectory groups, even when the data reflects a single underlying mean trajectory (Erosheva et al. 2014; Bauer and Curran 2003). According to the framework suggested by Qureshi and Fang (2011) one must demonstrate that individuals differ both in terms of starting points (intercepts) and growth over time (slopes) before determining that a group-based trajectory method is appropriate. In the context of this project, we sought to establish that different individuals entered the community with different levels of risk and that there were between-individual differences in how risk changed over time. We sought to demonstrate this using multilevel modelling, as it is readily able to accommodate the large number of reassessments and unbalanced measurement schedules.

Using the `lme4` package (Bates et al. 2015) for R (R Core Team 2015), we first compared model fit for linear and higher order models of the relationship between DRAOR subscale scores and time. Weeks were selected as our meaningful units of time. For each DRAOR subscale, mixed-models were run using linear, quadratic, and cubic specifications of time. For each subscale, higher order specifications of time resulted in optimal fit (see Table 4).

After selecting an appropriate representation of time, we then examined the variance component of intercept-only models to establish whether significant variation around intercepts was present. To establish whether individual differences existed in change in risk over time, we compared models with fixed slopes to models where slopes were allowed to vary across individuals. For each of the DRAOR subscales, allowing slopes to vary between individuals resulted in significant improvements of model fit (see Table 5).These results indicate that there is significant variation in both initial risk levels and change in risk over time. Therefore it was appropriate to move forward with testing whether distinct classes of risk trajectories were identifiable based on this unobserved heterogeneity.

Table 4 Model fit comparisons for linear and higher order specifications of time and DRAOR scores

	Model 1: Random intercept, random linear slope	Model 2: Random intercept, random linear slope, fixed quadratic term	Model 3: Random intercept, random linear slope, fixed quadratic term, fixed cubic term
AIC/BIC	261,569/261625	259,263/259329	258,941/259017
χ^2 Δ from Model 1	–	2307.55***	2631.43***
χ^2 Δ from Model 2	–	–	323.87***
χ^2 Δ from Model 3	–	–	–
AIC/BIC	290,998/291055	288,977/289043	288,544/288620
χ^2 Δ from Model 1	–	2023.42***	2458.13***
χ^2 Δ from Model 2	–	–	434.71***
χ^2 Δ from Model 3	–	–	–
AIC/BIC	259,819/259876	257,118/257184	256,682/256757
χ^2 Δ from Model 1	–	2703.38***	3141.14***
χ^2 Δ from Model 2	–	–	437.77***
χ^2 Δ from Model 3	–	–	–

N = 92,104 observations of 3421 individuals. DRAOR = Dynamic Risk Assessment for Offender Re-Entry (Serin 2007). AIC = Akaike Information Criteria. BIC = Bayesian Information Criteria

p* < .05, p** < .01, p*** < .001

Table 5 Multilevel models demonstrating variability around slopes and intercepts

Variable	Model 1: random intercept		Model 2: random intercept and fixed slope		Model 3: random intercept and random slope	
	B (SE)	*t*-value	B (SE)	*t*-value	B (SE)	*t*-value
DRAOR stable (N = 75,917 observations from 3416 individuals)						
Intercept	5.68 (0.05)	126	5.55 (0.04)	127.03	5.56 (0.04)	123.80
Linear slope (Time in weeks)	–	–	−108.20 (1.30)	−82.97	−105.99 (4.97)	−21.32
Quadratic slope (Time in weeks, squared)	–	–	26.89 (1.22)	22.11	29.86 (3.15)	9.47
Model fit						
AIC/BIC	261,465/261492		254,519/254566		213,267/213360	
ICC/ χ^2 Δ from prior model	ICC = 0.82		χ^2 Δ = 6949.2***		χ^2 Δ = 41,262.3***	
DRAOR acute (N = 75,917 observations from 3416 individuals)						
Intercept	4.97 (0.04)	136.5	4.82 (0.03)	138.37	4.81 (0.04)	136.67
Linear slope (Time in weeks)	–	–	−130.81 (1.42)	−92.42	−127.24 (1.42)	−89.68
Quadratic slope (Time in weeks, squared)	–	–	37.53 (1.32)	28.41	44.98 (3.46)	12.98
Model fit						
AIC/BIC	273,815/273843		265,119/265165		256,115 256,207	
ICC/ χ^2 Δ from prior model	ICC = 0.71		χ^2 Δ = 8700.7***		χ^2 Δ = 9013.9***	
DRAOR protect (N = 75,917 observations from 3416 individuals)						
Intercept	6.81 (0.04)	165.7	6.95 (0.04)	176.37	6.95 (0.04)	171.09
Linear slope (Time in weeks)	–	–	117.35 (1.26)	92.84	120.36 (4.67)	25.78
Quadratic slope (Time in weeks, squared)	–	–	−32.32 (1.18)	−27.42	−34.06 (3.04)	−11.22
Model fit						
AIC/BIC	258,044/258072		249,329/249375		213,681/213774	
ICC/ χ^2 Δ from prior model	ICC = 0.80		χ^2 Δ = 8719.3***		χ^2 Δ = 35,657.8***	

N = 75,917. DRAOR = Dynamic Risk Assessment for Offender Re-Entry (Serin 2007). AIC = Akaike Information Criteria. BIC = Bayesian Information Criteria. Quadratic specification was used instead of cubic, as singularity became an issue with the cubic specification of the DRAOR Stable subscale

p* < .05, p** < .01, p*** < .001

Appendix 4: Selection of Initial Values for Joint Latent Class Analyses

Latent class joint modelling in `lcmm` uses a maximum likelihood framework. Specifying appropriate initial values is critical, as it reduces computation time (see appendix of Brilleman et al. 2019) and increases the likelihood that convergence reflects a *global* maxima rather than a local maxima (Proust-Lima et al. 2017). Our approach for specifying initial values was to first estimate a model with the same structure as our intended joint model but with a single class. The parameter estimates from this model were used along with `lcmm`'s `gridsearch` function to randomly generate three sets of initial values (argument 'rep = 3'), each ran for 50 maximum iterations (argument 'maxiter = 50'). In cases where this was not sufficient to facilitate model convergence, the number of reps was increased to 5. If models still failed to converge, we increased reps to 7 and maximum iterations to 150.

Appendix 5: Model Selection Outcomes for DRAOR Stable and DRAOR Protect

Model selection criteria for DRAOR Stable and Protect trajectories are presented in Table 6. For DRAOR Stable subscales, BIC values continued to drop as more groups were added. We eliminated the six-group solution as group membership fell below 1% for one of the groups. To select between four- and five- group solutions we considered trajectory plots and chose to use the more parsimonious four-group solution. Trajectory estimates from the test sample did not match the calibration sample (increasing group not replicated). We therefore tested the five-group solution across the split sample. Trajectories from the calibration sample were consistent with the test sample (see Appendix F) and the five-group solution was upheld.

Models for DRAOR Protect had progressively better fit as indicated by BIC as more groups were added. Likewise, relative entropy increased as more classes were added. The six-group solution failed to converge. We therefore selected a five-group solution.

Appendix 6: Split Sample Model Comparisons

For each of the DRAOR subscales, the models selected from the 60% calibration sample (calibration sample) were applied to the remaining 40% of the sample (test sample) for visual comparison. We accepted the model from the calibration sample if the respective

Table 6 DRAOR stable and DRAOR protect model selection criteria derived from 60% calibration sample

Model	BIC	Max LL	CI test statistic	Relative entropy	Mean posterior probability within class (percentage)					
					1	2	3	4	5	6
DRAOR stable										
2 classes	154,675.32	−77,257.59	0.196	0.48	0.86 (51.17)	0.82 (48.83)				
3 classes	154,324.20	−77,059.15	0.79	0.66	0.86 (54.34)	0.82 (43.22)	0.89 (2.44)			
4 classes	154,236.00	−76,992.18	0.997	0.72	0.85 (50.73)	0.82 (45.41)	0.88 (2.83)	0.86 (1.02)		
5 classes	154,082.54	−76,892.57	2.11	0.71	0.85 (52.2)	0.78 (37.76)	0.78 (7.37)	0.86 (1.51)	0.84 (1.17)	
6 classes	154,008.22	−76,832.53	2.07	0.73	0.84 (50.20)	0.76 (37.17)	0.76 (9.22)	0.87 (1.61)	0.84 (1.32)	0.93 (0.49)
DRAOR protect										
2 classes	152,242.05	−76,040.96	1.92	0.51	0.86 (51.32)	0.84 (48.68)				
3 classes	151,992.62	−75,893.37	2.20	0.66	0.86 (51.07)	0.82 (45.76)	0.83 (3.17)			
4 classes	151,910.65	−75,829.5	0.23	0.69	0.87 (56.05)	0.78 (36.44)	0.77 (6.20)	0.84 (1.32)		
5 classes	151,779.84	−75,741.22	0.12	0.71	0.85 (52.49)	0.77 (37.51)	0.80 (7.27)	0.86 (1.46)	0.86 (1.27)	

Calibration sample comprised of 51,633 observations of N = 2050 randomly selected participants. There were 787 recidivism events recorded in this sample

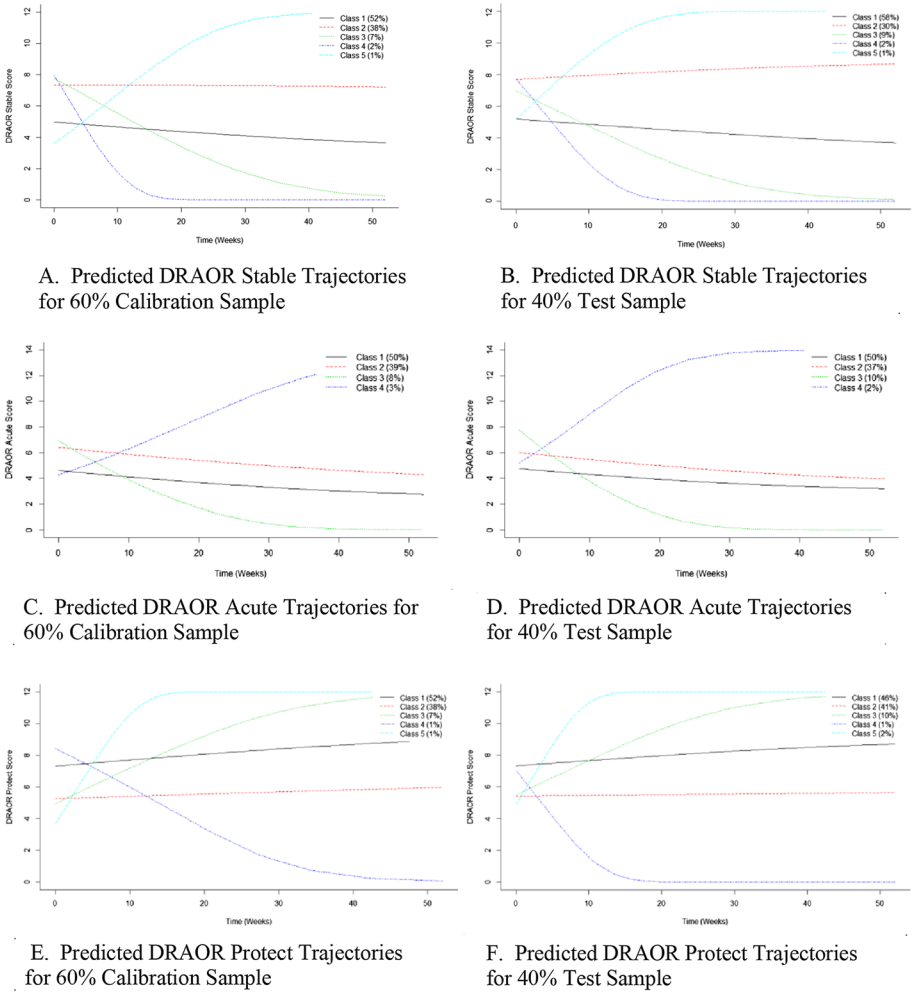
BIC indicates Bayesian information criterion

*Indicates significant conditional independence test statistic

intercepts, trajectory shapes, and proportion of participants identified were replicated in the test sample. Comparisons for the selected models are presented in Fig. 3.

For the four-group solution of DRAOR Stable, the intercepts, shapes, and proportion of participants allotted for three of the groups identified were similar, but the fourth, increasing group from the calibration sample was not replicated. We therefore rejected the four-group solution and revisited the five-group solution. For this solution, trajectories identified in the test sample were consistent with those identified in the calibration sample. The proportion of participants designated to each group was also similar across samples. We therefore upheld the five-group solution.

For DRAOR Acute, the proportion of participants designated into the first two groups was consistent across calibration and test samples, and the shapes and intercepts of the predicted trajectories were very similar. For this reason, we upheld the four-group solution.



A. Predicted DRAOR Stable Trajectories for 60% Calibration Sample

B. Predicted DRAOR Stable Trajectories for 40% Test Sample

C. Predicted DRAOR Acute Trajectories for 60% Calibration Sample

D. Predicted DRAOR Acute Trajectories for 40% Test Sample

E. Predicted DRAOR Protect Trajectories for 60% Calibration Sample

F. Predicted DRAOR Protect Trajectories for 40% Test Sample

Fig. 3 Cross sample comparisons of identified trajectories. Calibration sample comprised of 51,633 observations of $N = 2050$ randomly selected participants, with 787 recidivism events. Calibration sample comprised of 40,312 observations of $N = 1598$ remaining participants, with 646 recidivism events

For DRAOR Protect, the five-group solution yielded similar trajectories across calibration and test samples. The intercepts and trajectory shapes estimated were similar, and similar proportions of participants were identified as belonging in each trajectory group. We upheld the five-group solution accordingly.

Appendix 7: Descriptive Statistics of Selected DRAOR Stable Model

Among the five trajectory groups identified for DRAOR Stable scores some differed in both intercepts and trajectory shapes, while others differed primarily in intercepts (see Fig. 4). Each group was associated with different survival and hazard curves. Two groups had roughly parallel trajectories: one characterized by high initial DRAOR Stable scores that stayed high (*high stable* group, about 17% of participants; see Fig. 5 for sample trajectories within groups and Table 7 for descriptive statistics) and one characterized by low initial DRAOR Stable scores that decreased slightly (*low decreasing* group, about 8% of participants). A third group was characterized by moderate initial DRAOR Stable scores that decreased consistently across follow up (*moderate decreasing* group, about 70% of participants). A fourth group had initially high DRAOR Stable scores which decreased rapidly (*rapid decreasing* group, about 4% of participants). A fifth group had low initial DRAOR scores that increased over time (*increasing* group, about 1% of participants).

The *high stable*, *rapid decreasing*, and *increasing* groups had relatively high rates of recidivism and were assessed in the community for similar amounts of time. Participants designated to the *increasing* group had the highest rate of any recidivism and higher rates of each type of recidivism relative to other groups, despite having lower initial DRAOR scores. This group was the only to show substantial deterioration, and also had the highest mean net change, indicating greater fluctuation in scores from week to week. *High stable*

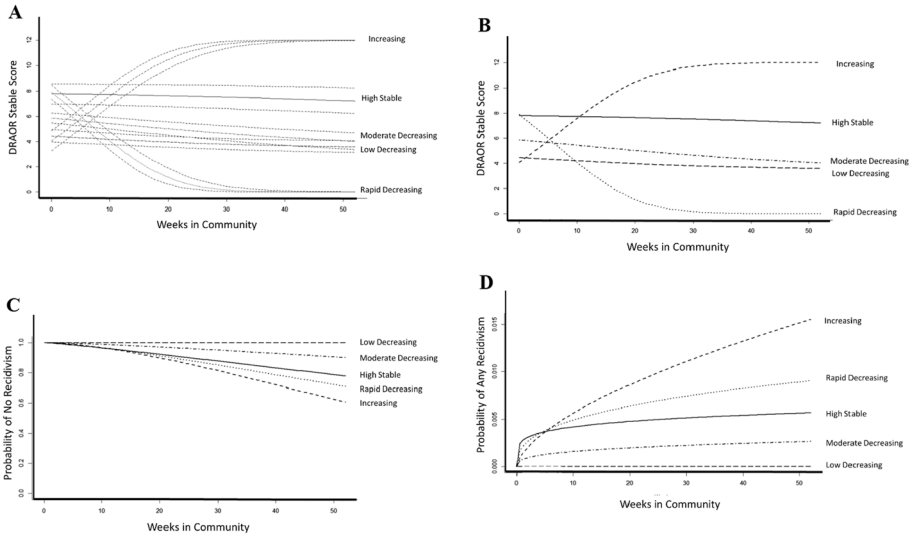


Fig. 4 Mean group predicted trajectories of DRAOR Stable plotted with (**A**) and without (**B**) Monte Carlo confidence intervals, mean group survival curves (**C**) and mean group hazard curves (**D**)

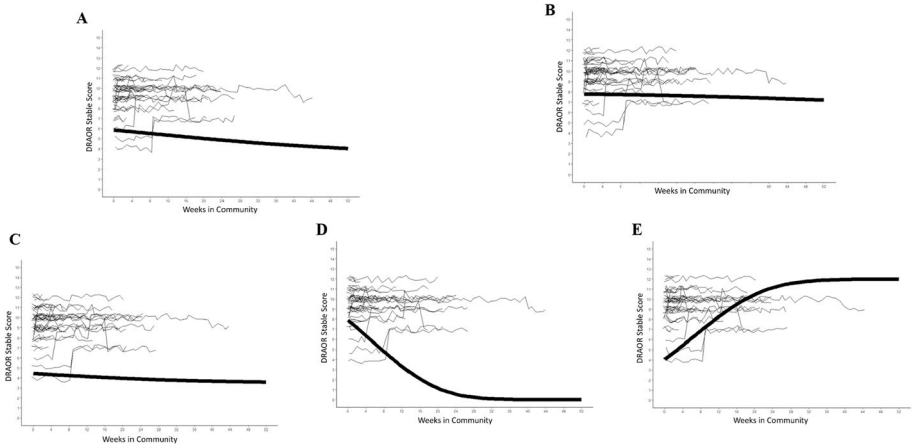


Fig. 5 Sample DRAOR Stable Trajectories with Mean Predicted Group Trajectories Superimposed. Heavily weighted line represents predicted mean trajectories of DRAOR Stable for **A** Moderate Decreasing, **B** High Stable, **C** Low Decreasing, **D** Rapid Decreasing, and **E** Increasing. Unweighted lines represent 50 randomly selected individual sample trajectories within each group (except increasing which only had 47 participants designated to it), jittered to reduce overlap. Jittering increases readability, facilitating illustration of within-group noise, but creates appearance of oscillation, where scores may be constant across measurement occasions

and *rapid decreasing* groups also had elevated mean net change in DRAOR Acute scores relative to *moderate decreasing* and *low decreasing* groups.

The *moderate decreasing* group and *low decreasing* group each had lower rates of recidivism than the other three groups. The *low decreasing* group had substantially less recidivism and was assessed in the community for longer than the *moderate decreasing* group despite the *moderate decreasing* group being older, with lower static risk. Conversely, the *low decreasing* group had lower initial risk as indicated by DRAOR scores, and had less fluctuation in DRAOR scores as indicated by mean net change.

Appendix 8: Descriptive Statistics of Selected DRAOR Protect Model

Of the five trajectory groups of DRAOR Protect scores identified, three reflected scores that increased over time (see Fig. 6). The first trajectory group was characterized by high initial Protect scores that increased somewhat over time (*high increasing*; about 57% of participants; see Fig. 7 for sample trajectories within groups). The participants in this group were older, spent more time in the community, and had low rates of any recidivism relative to other groups and sample averages. They had low risk at the time of re-entry and improved moderately across subscales, with low mean net change (descriptive statistics in Table 8).

The remaining two increasing groups were characterised by rapid increases in DRAOR Protect scores. One started with slightly higher initial Protect scores (*moderate-rapid increasing* group, about 7% of participants) and one started with low initial Protect scores (*low-rapid increasing* group, about 1.5% of participants. Participants in the *low-rapid increasing* group fared worse than those in the *moderate-rapid increasing* group, with less time spent in the community and higher rates of recidivism, although they demonstrated

Table 7 Descriptive qualities of five DRAOR stable trajectories

	Moderate decreasing N (%)	High stable	Low decreasing	Rapid decreasing	Increasing	
Number of participants	2544 (69.74)	617 (16.91)	297 (8.14)	143 (3.92)	47(1.29)	
	Mean (SD)					Kruskal–Wallis $\chi 2$
Age	35.58 (11.81)	33.79 (11.05)	32.80 (10.51)	31.55 (10.11)	28.00 (10.03)	42.91***
Weeks follow up	26.76 (16.46)	13.06 (11.78)	43.96 (11.82)	14.76 (8.92)	13.94 (7.98)	761.38***
RoC*RoI	0.48 (0.25)	0.57 (0.23)	0.61 (0.20)	0.56 (0.22)	0.60 (0.19)	128.69***
Baseline stable	5.94 (2.15)	9.05 (1.82)	4.14 (2.36)	8.19 (2.42)	4.57 (2.04)	1046.2***
Baseline acute	5.63 (2.23)	7.58 (2.39)	4.74 (2.19)	6.83 (2.80)	5.70 (2.11)	377.83***
Baseline protect	6.30 (2.20)	4.32 (2.28)	7.55 (2.32)	5.16 (2.41)	6.68 (2.49)	452.56***
Change stable	− 1.04 (2.25)	0.20 (1.35)	−0.88 (2.40)	−4.54 (2.49)	4.23 (1.59)	585.94***
Change acute	−1.40 (2.46)	−0.70 (2.34)	−1.64 (2.45)	−2.81 (2.67)	1.19 (2.50)	142.36***
Change protect	1.18 (2.26)	0.13 (1.81)	1.37 (2.41)	2.71 (2.67)	−1.91 (3.02)	275.89***
Mean net change stable	0.21 (0.46)	0.29 (0.72)	0.12 (0.13)	0.47 (0.45)	0.69 (0.65)	338.15***
Mean net change acute	0.36 (0.48)	0.50 (0.64)	0.25 (0.18)	0.46 (0.41)	0.86 (0.65)	95.05***
Mean net change protect	0.21 (0.48)	0.32 (0.69)	0.13 (0.16)	0.37 (0.42)	0.58 (0.76)	122.15***
	N (%)					Pearson's χ^2
Any recidivism	891 (35.02)	468 (75.85)	41 (13.80)	97 (67.83)	38 (80.85)	506.01***
Technical violations	626 (24.61)	348 (56.40)	29 (9.76)	74 (51.75)	31 (65.96)	356.35***
Nonviolent criminal recidivism	237 (9.32)	136 (22.04)	7 (2.36)	21 (14.69)	6 (12.77)	107.58***
Violent recidivism	153 (6.01)	83 (13.45)	5 (1.68)	18 (12.59)	9 (19.14)	69.81***

$N = 3648$ participants, assigned to groups based on posterior probabilities

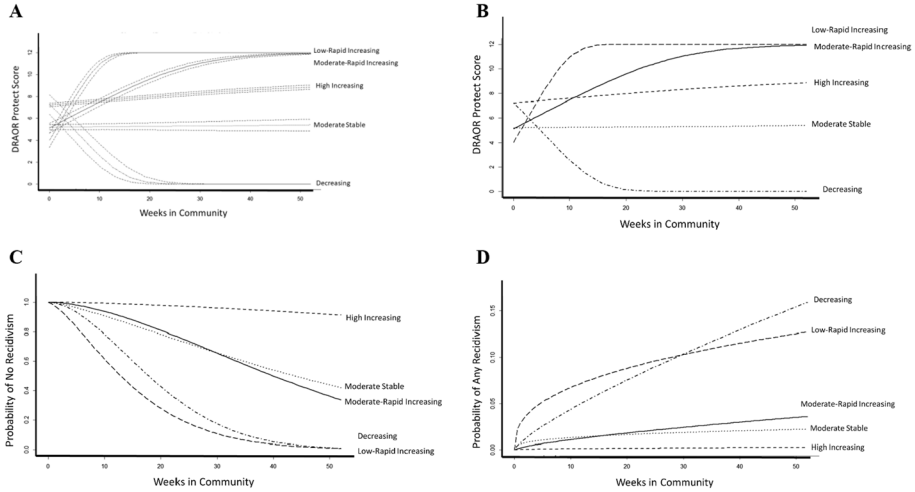


Fig. 6 Mean group predicted trajectories of DRAOR Protect plotted with (**A**) and without (**B**) Monte Carlo confidence intervals, mean group survival curves (**C**) and mean group hazard curves (**D**)

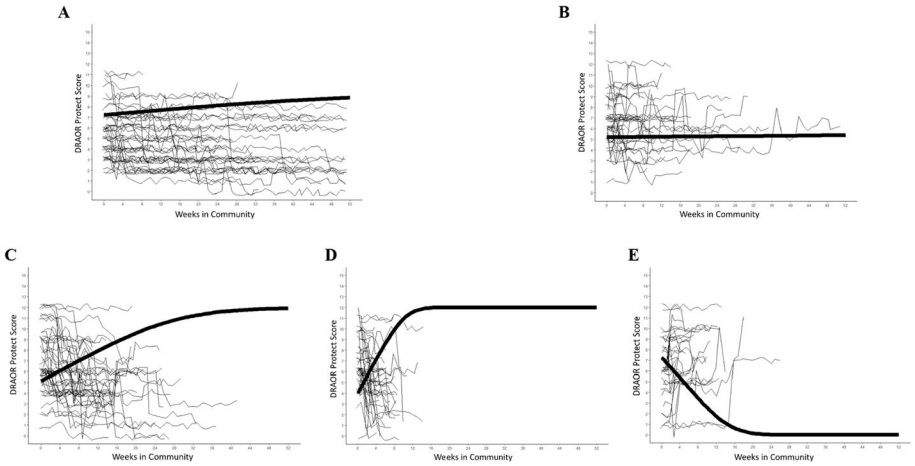


Fig. 7 Sample DRAOR Protect Trajectories with Mean Predicted Group Trajectories Superimposed. Heavily weighted line represents predicted mean trajectories of DRAOR Protect for **A** High Increasing, **B** Moderate Stable, **C** Moderate-Rapid Increasing, **D** Low-Rapid Increasing, and **E** Decreasing. Unweighted lines represent 50 randomly selected individual sample trajectories within each group (except decreasing which only had 33 participants designated to it), jittered to reduce overlap. Jittering increases readability, facilitating illustration of within-group noise, but creates appearance of oscillation, where scores may be constant across measurement occasions

greater improvement in DRAOR Acute and DRAOR Protect scores and had lower RoC*RoI scores. Participants in the *low-rapid increasing* group had high mean net change across subscales, indicating more fluctuation in scores across weeks.

Another trajectory group was characterized by moderate initial Protect scores that changed little (*moderate stable* group, about 34% of participants). These participants had

Table 8 Descriptive Qualities of Five DRAOR Protect Trajectories

	High increasing	Moderate stable	Moderate-rapid increasing	Low-rapid increasing	Decreasing	
	N (%)					
Number of participants	2069 (56.72)	1240 (33.99)	253 (6.94)	53 (1.45)	33 (0.90)	
	Mean (SD)					Kruskal–Wallis $\chi 2$
Age	37.43 (12.31)	31.06 (9.25)	32.20 (9.67)	34.13 (12.78)	33.70 (8.88)	233.48***
Weeks follow up	34.06 (16.00)	13.05 (9.95)	18.17 (8.00)	6.98 (3.20)	9.79 (5.01)	1330.10***
RoC*RoI	0.40 (0.23)	0.67 (0.16)	0.60 (0.19)	0.55 (0.22)	0.59 (0.19)	986.09***
Baseline stable	5.61 (2.40)	7.51 (2.27)	7.13 (2.48)	6.94 (2.50)	6.24 (2.97)	476.17***
Baseline acute	5.31 (2.20)	6.82 (2.39)	6.52 (2.67)	7.13 (2.53)	6.33 (2.15)	317.16***
Baseline protect	6.88 (2.13)	4.89 (2.23)	4.83 (2.26)	4.08 (2.49)	7.73 (2.35)	648.80***
Change stable	−1.27 (2.34)	0.05 (1.83)	−2.54 (2.97)	−1.62 (2.86)	1.79 (2.56)	430.84***
Change acute	−1.69 (2.39)	−0.50 (2.36)	−2.35 (2.80)	−2.38 (2.44)	0.30 (2.39)	262.53***
Change protect	1.40 (2.15)	−0.09 (1.58)	3.68 (2.25)	4.04 (2.72)	−4.73 (2.17)	916.25***
Mean net change stable	0.14 (0.25)	0.36 (0.73)	0.35 (0.54)	0.60 (0.79)	0.49 (0.53)	159.84***
Mean net change acute	0.27 (0.30)	0.53 (0.67)	0.53 (0.61)	0.65 (0.74)	0.64 (0.35)	211.25***
Mean net change protect	0.14 (0.28)	0.31 (0.72)	0.42 (0.41)	0.94 (0.79)	0.85 (0.51)	456.01***
	N (%)					Pearson's χ^2
Any recidivism	320 (15.47)	977 (78.79)	167 (66.01)	44 (83.02)	27 (81.81)	1404.2***
Technical violations	213 (10.29)	716 (57.74)	127 (50.20)	32 (60.38)	20 (60.61)	917.46***
Nonviolent criminal recidivism	68 (3.29)	271 (21.85)	46 (18.18)	14 (26.42)	8 (24.24)	303.21***
Violent recidivism	54 (2.61)	173 (13.95)	24 (9.49)	10 (18.87)	7 (21.21)	169.03***

$N = 3648$ participants, assigned to groups based on posterior probabilities

Table 9 Predictive discrimination and calibration of selected joint latent class model and equivalent model without latent class structure for DRAOR Stable and Protect

Model	Prediction window					
	4 weeks through 12 weeks		12 weeks through 20 weeks		24 weeks through 32 weeks	
	AUC*100	Brier score*100	AUC *100	Brier score*100	AUC *100	Brier score*100
DRAOR stable						
Five class JLCM	73.46	11.24	70.16	12.49	74.50	9.79
Equivalent SREM	73.44	10.87	73.63	11.28	75.28	9.94
Five class JLCM	70.85	13.29	68.88	13.76	71.60	10.22
Equivalent SREM	73.56	10.83	75.02	11.18	75.37	9.84

Survival predictions from JLCM used Weibull baseline hazard stratified on probable group membership with RoC*RoI as covariate. Survival predictions from SREM used Weibull baseline hazard with RoC*RoI as covariate

slightly higher rates of recidivism than the *moderate-rapid increasing* group, with similar levels of mean net change. Of all groups identified, participants in this group had the highest static risk scores.

Finally, one group was characterised by decreasing Protect scores (*decreasing* group, about 1% of participants). This group had similar outcomes to the *low-rapid increasing* group, with relatively high recidivism rates and little time spent in the community. This was the only group to demonstrate deterioration over time, and they had a high degree of mean net change.

Appendix 9: Dynamic Predictions and Comparison of Predictive Ability for DRAOR Stable and DRAOR Protect

As with DRAOR Acute models, there was little difference between DRAOR Stable models with and without latent classes during the early prediction window (Table 9), whereas models without latent class structures outperformed those with latent classes during the second prediction window (12 through 20 weeks), and, to a lesser extent, during the third prediction window (24 through 30 weeks). For DRAOR Protect, the models without latent class structures outperformed the latent class models across all prediction windows.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10940-022-09566-5.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Data Availability The data used in this study are owned by the New Zealand Department of Corrections and were used under license agreement with the fifth author of the current study, and so are not publicly available. However, data are available from the authors upon reasonable request and with explicit written permission of New Zealand Department of Corrections. The data used for these analyses have been analyzed

previously in different applications, some of which have been published or submitted for publication. These include (Coulter et al. in press; Lloyd et al. 2020a,b; Stone et al. 2021, 2022). Please also note that the lead author has provided sample code and a manual with further instructions and considerations for using JLCM, available at https://github.com/agstone90/JLCM.

Declarations

Conflict of interest The views expressed are those of the authors and not necessarily those of the New Zealand Department of Corrections. Ralph C. Serin is the author and holds the copyright of the Dynamic Risk Assessment for Offender Re-entry (DRAOR) tool. Caleb D. Lloyd and Ralph C. Serin are co-authors of the 2017 version of the DRAOR scoring manual and co-developers of the DRAOR training program and training certification.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Andrews DA, Bonta J (2010) The psychology of criminal conduct, Anderson, Cincinnati

Australian Bureau of Statistics (2018) Corrective services, Australia, September Quarter 2018. www.abs.gov.au

Babchishin KM, Hanson RK (2020) Monitoring changes in risk of reoffending: a prospective study of 632 men on community supervision. J Consult Clin Psychol 88:886–898. https://doi.org/10.1037/ccp0000601

Baglivio MT, Wolff KT, Piquero AR, Howell JC, Greenwald MA (2017) Risk assessment trajectories of youth during juvenile justice residential placement. Crim Justice Behav 44:360–394. https://doi.org/10.1177/0093854816668918

Bakker L, Riley D, O'Malley J (1999) Risk of reconviction: statistical models predicting four types of re-offending. Wellington, NZ: Department of Corrections Psychological Service. https://www.corrections.govt.nz/__data/assets/pdf_file/0020/10667/roc.pdf

Barnes JC, El Sayed SA, TenEyck M, Nedelec JL, Connolly EJ, Schwartz JA, Anderson NE (2017) Estimating relative stability in developmental research: a critique of modern approaches and a novel method. J Quant Criminol 33:319–346. https://doi.org/10.1007/s10940-016-9298-5

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4. J Stat Softw 67:1–48. https://doi.org/10.18637/jss.v067.i01

Bauer DJ, Curran PJ (2003) Distributional assumptions of growth mixture models: implications for overextraction of latent trajectory classes. Psychol Methods 8(3):338. https://doi.org/10.1037/1082-989X.8.3.338

Berg MT, Huebner BM (2011) Reentry and the ties that bind: an examination of social ties, employment, and recidivism. Justice Q 28:382–410. https://doi.org/10.1080/07418825.2010.498383

Billen E, Garofalo C, Vermunt JK, Bogaerts S (2019) Trajectories of self-control in a forensic psychiatric sample: Stability and association with psychopathology, criminal history, and recidivism. Crim Justice Behav 46:1255–1275. https://doi.org/10.1177/0093854819856051

Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues JF, Jacqmin-Gadda H (2015) Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. Biometrics 71:102–113. https://doi.org/10.1111/biom.12232

Bonta J, Andrews DA (2017) The psychology of criminal conduct. Routledge, New York

Bonta J, Rugge T, Scott TL, Bourgon G, Yessine AK (2008) Exploring the black box of community supervision. J Offender Rehabil 47:248–270. https://doi.org/10.1080/10509670802134085

Bonta J, Bourgon G, Rugge T, Scott TL, Yessine AK, Gutierrez L, Li J (2011) An experimental demonstration of training probation officers in evidence-based community supervision. Crim Justice Behav 38:1127–1148. https://doi.org/10.1177/0093854811420678

Bottoms A, Shapland J, Costello A, Holmes D, Muir G (2004) Towards desistance: theoretical underpinnings for an empirical study. Howard J Crim Justice 43:368–389. https://doi.org/10.1111/j.1468-2311.2004.00336.x

Brilleman SL, Moreno-Betancur M, Polkinghorne KR, McDonald SP, Crowther MJ, Thomson J, Wolfe R (2019) Changes in body mass index and rates of death and transplant in hemodialysis patients: a latent class joint modeling approach. Epidemiology 30:38–47. https://doi.org/10.1097/EDE.0000000000000931

Brown SL, St Amand MD, Zamble E (2009) The dynamic prediction of criminal recidivism: a three-wave prospective study. Law Hum Behav 33:25–45. https://doi.org/10.1007/s10979-008-9139-7

Chadwick N, Dewolf A, Serin R (2015) Effectively training community supervision officers: a meta-analytic review of the impact on offender outcome. Crim Justice Behav 42:977–989. https://doi.org/10.1177/0093854815595661

Chow SM, Ram N, Boker SM, Fujita F, Clore G (2005) Emotion as a thermostat: representing emotion regulation using a damped oscillator model. Emotion 5:208–225. https://doi.org/10.1037/1528-3542.5.2.208

Clarke MC (2015) The examination of changes in dynamic risk factors and recidivism over time in youth offenders (Unpublished doctoral dissertation). University of Toronto, Canada

Coulter DJ, Lloyd CD, Serin RC (in press) Psychometric properties of a risk tool across Indigenous Māori and European samples in Aotearoa New Zealand: measurement invariance, discrimination, and calibration for predicting criminal recidivism. Assessment

Cullen FT (2013) Rehabilitation: Beyond nothing works. In: Tonry V (ed) Crime and justice: a review of research. University of Chicago Press, Chicago, pp 299–376

Cullen FT, Jonson CL, Mears DP (2017) Reinventing community corrections. Crime Justice 46:27–93. https://doi.org/10.1086/688457

D'unger AV, Land KC, McCall PL, Nagin DS (1998) How many latent classes of delinquent/criminal careers? Results from mixed Poisson regression analyses. Am J Sociol 103:1593–1630. https://doi.org/10.1086/231402

Davies ST (2019) An investigation of how change in dynamic risk and protective factors affects the prediction of imminent criminal recidivism (Unpublished doctoral dissertation). Victoria University of Wellington, New Zealand

Davies ST, Lloyd CD, Polaschek DLL (2021) Does reassessment enhance the prediction of imminent criminal recidivism? Replicating Lloyd et al. (2020) with high-risk parolees. Assessment. https://doi.org/10.1177/1073191121993216

Douglas KS, Skeem JL (2005) Violence risk assessment: getting specific about being dynamic. Psychol Public Policy Law 11:347–383. https://doi.org/10.1037/1076-8971.11.3.347

Dowden C, Andrews DA (2004) The importance of staff practice in delivering effective correctional treatment: a meta-analytic review of core correctional practice. Int J Offender Ther Comp Criminol 48:203–214. https://doi.org/10.1177/0306624X03257765

Erosheva EA, Matsueda RL, Telesca D (2014) Breaking bad: two decades of life-course data analysis in criminology, developmental psychology, and beyond. Annu Rev Stat Appl 1:301–332. https://doi.org/10.1146/annurev-statistics-022513-115701

Gendreau P, Ross B (1979) Effective correctional treatment: bibliotherapy for cynics. Crime Delinq 25:463–489. https://doi.org/10.1177/001112877902500405

Georgiou G (2014) Does increased post-release supervision of criminal offenders reduce recidivism? Evidence from a statewide quasi-experiment. Int Rev Law Econ 37:221–243. https://doi.org/10.1016/j.irle.2013.12.001

Hanson RK, Harris AJ, Scott TL, Helmus L (2007) Assessing the risk of sexual offenders on community supervision: the dynamic supervision project, Vol 5, No 6. Public Safety Canada, Ottawa, Canada

Hanson RK, Harris AJ (2000) Where should we intervene? Dynamic predictors of sexual offense recidivism. Crim Justice Behav 27:6–35. https://doi.org/10.1177/0093854800027001002

Hanson RK, Newstrom N, Brouillette-Alarie S, Thornton D, Robinson BE, Miner MH (2021) Does reassessment improve prediction? a prospective study of the Sexual Offender Treatment Intervention and Progress scale (SOTIPS). Int J Offender Ther Comp Criminol 65(16):1775–1803. https://doi.org/10.1177/0306624X20978204

Horney J, Osgood DW, Marshall IH (1995) Criminal careers in the short-term: intraindividual variability in crime and its relation to local life circumstances. Am Sociol Rev 60:655–673. https://doi.org/10.2307/2096316

Howard PD, Dixon L (2013) Identifying change in the likelihood of violent recidivism: causal dynamic risk factors in the OASys violence predictor. Law Hum Behav 37:163–174. https://doi.org/10.1037/lhb0000012

Hudson SM, Ward T, McCormack JC (1999) Offense pathways in sexual offenders. J Interpers Violence 14:779–798. https://doi.org/10.1177/088626099014008001

Hyatt JM, Barnes GC (2017) An experimental evaluation of the impact of intensive supervision on the recidivism of high-risk probationers. Crime Delinq 63:3–38. https://doi.org/10.1177/0011287145 55757

Jennings WG, Reingle JM (2012) On the number and shape of developmental/life-course violence, aggression, and delinquency trajectories: a state-of-the-art review. J Crim Justice 40:472–489. https://doi.org/10.1016/j.jcrimjus.2012.07.001

Kaeble D, Alper M (2020) Probation and parole in the United States, 2017–2018. https://bjs.ojp.gov/content/pub/pdf/ppus1718.pdf

Kazemian L, Farrington DP, Le Blanc M (2009) Can we make accurate long-term predictions about patterns of de-escalation in offending behavior? J Youth Adolesc 38:384–400. https://doi.org/10.1007/s10964-008-9338-z

Kennealy PJ, Skeem JL, Manchak SM, Eno Louden J (2012) Firm, fair, and caring officer-offender relationships protect against supervision failure. Law Hum Behav 36(6):496. https://doi.org/10.1037/h0093935

Lane J, Turner S, Fain T, Sehgal A (2005) Evaluating an experimental intensive juvenile probation program: supervision and official outcomes. Crime Delinq 51:26–52. https://doi.org/10.1177/0011128704 264943

LeBel TP, Burnett R, Maruna S, Bushway S (2008) The 'chicken and egg' of subjective and social factors in desistance from crime. Eur J Criminol 5:131–159. https://doi.org/10.1177/1477370807087640

Lipsey MW (1995) What do we learn from 400 research studies on the effectiveness of treatment with juvenile delinquents? In: McGuire J (ed) What works: reducing reoffending: guidelines from research and practice. Wiley, Chichester, UK, pp 63–78

Lipsey MW (1989) The efficacy of intervention for juvenile delinquency: Results from 400 studies. Paper presented at the annual meeting of the American Society of Criminology. Reno, NV

Lloyd CD (2015) Can a dynamic risk instrument make short-‐term predictions in "real time"? Developing a framework for testing proximal assessment of offender recidivism risk during re-entry (Doctoral dissertation, Carleton University, Ottawa, Canada). https://curve.carleton.ca/system/files/etd

Lloyd CD, Serin RC (2012) Agency and outcome expectancies for crime desistance: measuring offenders' personal beliefs about change. Psychol Crime Law 18:543–565. https://doi.org/10.1080/1068316X.2010.511221

Lloyd CD, Hanson RK, Richards DK, Serin RC (2020a) Reassessment improves prediction of criminal recidivism: a prospective study of 3,421 individuals in New Zealand. Psychol Assess 32:568–581. https://doi.org/10.1037/pas0000813

Lloyd CD, Perley-Robertson B, Serin RC (2020b) Age and strengths in a community corrections sample. Int J Forensic Ment Health 19(3):215–226. https://doi.org/10.1080/14999013.2019.1684404

Mann RE, Hanson RK, Thornton D (2010) Assessing risk for sexual recidivism: some proposals on the nature of psychologically meaningful risk factors. Sex Abuse 22:191–217. https://doi.org/10.1177/1079063210366039

Martinson R (1974) What works?-Questions and answers about prison reform. Public Interest 35:22–54

Martinson R, Wilks J (1978) Save parole supervision. Fed Probat 42:23–27

Maruna S (2001) Making good, American psychological association, Washington D.C.

Ministry of Justice (2018) Offender management statistics bulletin, England and Wales. Retrieved from-https://www.gov.uk/government/collections/offender-management-statistics-quarterly

Nadesu A (2007) Reconviction patterns of released prisoners: A 36-months follow-up analysis. New Zealand Department of Corrections. Retrieved from https://www.corrections.govt.nz/__data/assets/pdf_file/0015/10680/reimprisonment-report.pdf

Nagin DS (2005) Group-based modeling of development. Harvard Univ. Press, Cambridge, MA

Nagin DS, Tremblay RE (2005) Developmental trajectory groups: Fact or a useful statistical fiction? Criminology 43:873–904. https://doi.org/10.1111/j.1745-9125.2005.00026.x

New Zealand Department of Corrections (2020) Corrections volumes, 2019–2020. https://www.corrections.govt.nz/__data/assets/pdf_file/0016/41191/Corrections_Volumes_Report_2019-2020.pdf

Papalia N, Spivak B, Daffern M, Ogloff JR (2020) Are psychological treatments for adults with histories of violent offending associated with change in dynamic risk factors? A meta-analysis of intermediate treatment outcomes. Crim Justice Behav 47:1585–1608. https://doi.org/10.1177/0093854820956377

Petersilia J (2008) Influencing public policy: an embedded criminologist reflects on california prison reform. J Exp Criminol 4:335–356. https://doi.org/10.1007/s11292-008-9060-6

Petersilia J, Cullen FT (2015) Liberal but not stupid: Meeting the promise of downsizing prisons. Stanf J Crim Law Policy 2:1–43. https://doi.org/10.2139/ssrn.2458024

Petersilia J, Turner S (1993) Intensive probation and parole. Crime Justice 17:281–335. https://doi.org/10.1086/449215

Pettersen C, Nunes KL, Woods M, Maimone S, Hermann CA, Looman J, Spape J (2015) Does change in hostility predict sexual recidivism? Int J Offender Ther Comp Criminol 59(6):565–587. https://doi.org/10.1177/0306624X14547033

Piquero AR (2008) Taking stock of developmental trajectories of criminal activity over the life course. In: Liberman AM (ed) The long view of crime: a synthesis of longitudinal research. Springer, New York, pp 23–78

Proust-Lima C, Séne M, Taylor JM, Jacqmin-Gadda H (2014) Joint latent class models for longitudinal and time-to-event data: a review. Stat Methods Med Res 23:74–90. https://doi.org/10.1177/0962280212445839

Proust-Lima C, Philipps V, Liquet B (2017) Estimation of extended mixed models using latent classes and latent processes: the R package lcmm. J Stat Softw 78:1–56. https://doi.org/10.18637/jss.v078.i02

Public Safety Canada (2020) Corrections and conditional release statistical overview, 2019 annual report. https://www.publicsafety.gc.ca/cnt/rsrcs/pblctns/ccrso-2019/ccrso-2019-en.pdf

Quinsey VL, Harris GT, Rice ME, Cormier CA (2006a) Violent offenders: appraising and managing risk, 2nd edn. American Psychological Association, Washington, DC

Quinsey VL, Jones GB, Book AS, Barr KN (2006b) The dynamic prediction of antisocial behavior among forensic psychiatric patients: a prospective field study. J Interpers Violence 21:1539–1565. https://doi.org/10.1177/0886260506294238

Qureshi I, Fang Y (2011) Socialization in open source software projects: a growth mixture modeling approach. Organ Res Methods 14(1):208–238. https://doi.org/10.1177/1094428110375002

R Core Team (2015). R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. http://www.R-project.org/.

Rizopoulos D (2010) JM: an R package for the joint modelling of longitudinal and time-to-event data. J Stat Softw 35:1–33

Rizopoulos D (2012) Joint models for longitudinal and time-to-event data: with applications in R. CRC Press, Boca-Raton

Roeder K, Lynch KG, Nagin DS (1999) Modeling uncertainty in latent class membership: a case study in criminology. J Am Stat Assoc 94:766–776. https://doi.org/10.1080/01621459.1999.10474179

Sampson RJ, Laub JH, Eggleston EP (2004) On the robustness and validity of groups. J Quant Criminol 20(1):37–42. https://doi.org/10.1023/B:JOQC.0000016698.36239.91

Sarre R (2001) Beyond 'What Works?' A 25-year jubilee retrospective of Robert Martinsons famous article. Aust N Z J Criminol 34:38–46. https://doi.org/10.1177/000486580103400103

Schwarz G (1978) Estimating the dimension of a model. Ann Stat 6:461–464

Skeem JL, Schubert C, Odgers C, Mulvey EP, Gardner W, Lidz C (2006) Psychiatric symptoms and community violence among high-risk patients: a test of the relationship at the weekly level. J Consult Clin Psychol 74(5):967. https://doi.org/10.1037/0022-006X.74.5.967

Serin RC, Lloyd CD, Helmus L, Derkzen DM, Luong D (2013) Does intra-individual change predict offender recidivism? Searching for the Holy Grail in assessing offender change. Aggress Violent Behav 18:32–53. https://doi.org/10.1016/j.avb.2012.09.002

Serin RC, Gobeil R, Lloyd CD, Chadwick N, Wardrop K, Hanby L (2016) Using dynamic risk to enhance conditional release decisions in prisoners to improve their outcomes. Behav Sci Law 34:321–336. https://doi.org/10.1002/bsl

Serin RC (2007) The dynamic risk assessment for offender re-entry (DRAOR). Unpublished user manual

Smagula SF, Jia Y, Chang CCH, Cohen A, Ganguli M (2019) Trajectories of daytime sleepiness and their associations with dementia incidence. J Sleep Res 29:1–6. https://doi.org/10.1111/jsr.12952

Stone AG, Lloyd CD, Serin RC (2021) Dynamic risk factors reassessed regularly after release from incarceration predict imminent violent recidivism. Law Hum Behav. https://doi.org/10.1037/lhb0000463

Stone AG, Spivak BL, Lloyd CD, Papalia NL, Serin R (2022) Clients' current presentation yields best prediction of criminal recidivism: Jointly modelling repeated assessments of risk and associated recidivism outcomes in a community sample of paroled New Zealanders. J Consult Clin Psychol. https://doi.org/10.1037/ccp0000766

Szasz TS (1961) The myth of mental illness: foundations of a theory of personal conduct. Harper Perennial, New York

Taxman FS, Caudy MS (2015) Risk tells us who, but not what or how: empirical assessment of the complexity of criminogenic needs to inform correctional programming. Criminol Public Policy 14:71–103. https://doi.org/10.1111/1745-9133.12116

Viljoen JL, Cochrane DM, Jonnson MR (2018) Do risk assessment tools help manage and reduce risk of violence and reoffending? A systematic review. Law Hum Behav 42:181–214. https://doi.org/10.1037/lhb0000280

Weaver B (2019) Understanding desistance: a critical review of theories of desistance. Psychol Crime Law 25:641–658. https://doi.org/10.1080/1068316X.2018.1560444

Yesberg JA, Polascheck DI (2014) Assessing dynamic risk and protective factors in the community: examining the validity of the dynamic risk assessment for offender re-entry. Psychol Crime Law 21:80–99. https://doi.org/10.1080/1068316x.2014.935775

Yesberg JA, Polaschek DLL (2019) How does offender rehabilitation actually work? Exploring mechanisms of change in high-risk treated parolees. Int J Offender Ther Comp Criminol 63:2672–2692. https://doi.org/10.1177/0306624X19856221

Yesberg JA, Scanlan JM, Hanby LJ, Serin RC, Polaschek DLL (2015) Predicting women's recidivism: Validating a dynamic community-based 'gender-neutral' tool. Probation J 62:33–48. https://doi.org/10.1177/0264550514562851

Zamble E, Porporino FJ (1988) Problems and coping on the outside. In: Zamble E, Porporino FJ (eds) Coping behavior and adaptation in prison inmates. Springer, New York, pp 51–75

Zamble E, Quinsey VL (2001) The criminal recidivism process. Cambridge University Press, Cambridge

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.