

The predictive value of neurobiological measures for recidivism in delinquent male young adults

Josjan Zijlmans, PhD; Reshmi Marhe, PhD; Floor Bevaart, PhD;
Laura Van Duin, MSc; Marie-Jollette A. Luijckx, MSc; Ingmar Franken, PhD;
Henning Tiemeier, MD, PhD; Arne Popma, MD, PhD

Background: Neurobiological measures have been associated with delinquent behaviour, but little is known about the predictive power of these measures for criminal recidivism and whether they have incremental value over and above demographic and behavioural measures. This study examined whether selected measures of autonomic functioning, functional neuroimaging and electroencephalography predict overall and serious recidivism in a sample of 127 delinquent young adults. **Methods:** We assessed demographics; education and intelligence; previous delinquency and drug use; behavioural traits, including aggression and psychopathy; and neurobiological measures, including heart rate, heart rate variability, functional brain activity during an inhibition task and 2 electroencephalographic measures of error-processing. We tested longitudinal associations with recidivism using Cox proportional hazard models and predictive power using C-indexes. **Results:** Past offences, long-term cannabis use and reactive aggression were strongly associated with recidivism, as were resting heart rate and error-processing. In the predictive model, demographics, past delinquency, drug use and behavioural traits had moderate predictive power for overall and for serious recidivism (C-index over 30 months [fraction of pairs in the data, where the higher observed survival time was correctly predicted]: $C_{30} = 0.68$ and 0.75 , respectively). Neurobiological measures significantly improved predictive power ($C_{30} = 0.72$ for overall recidivism and $C_{30} = 0.80$ for serious recidivism). **Limitations:** Findings cannot be generalized to females, and follow-up was limited to 4 years. **Conclusion:** Demographic and behavioural characteristics longitudinally predicted recidivism in delinquent male young adults, and neurobiological measures improved the models. This led to good predictive function, particularly for serious recidivism. Importantly, the most feasible measures (autonomic functioning and electroencephalography) proved to be useful neurobiological predictors.

Introduction

Delinquent behaviour is a major psychiatric and public safety issue with severe consequences for individuals and for society.¹ Several longitudinal studies have advanced our etiological understanding of delinquent behaviour and helped to identify various relevant risk factors for these behaviours, including age, sex, criminal history, drug use and behavioural problems.² However, despite advances in knowledge about the involvement of neurobiological mechanisms in the etiology of delinquency, only a few studies have integrated neurobiological measures with longitudinal prediction.

One of these advances is our knowledge of the relationship between the autonomic nervous system and delinquency. Resting heart rate is arguably the best-replicated physiologic correlate of antisocial behaviour.^{3,4} Several decades ago, Wadsworth and colleagues⁵ demonstrated that lower childhood heart rate predicted delinquency. Similarly, Raine and colleagues⁶ related lower heart rate to anti-

social behaviour in the general population and high-risk samples. A seminal study in Swedish conscripts provided strong evidence that heart rate at age 18 is longitudinally related to violent and nonviolent delinquency.⁷ Heart rate variability, as measured by respiratory sinus arrhythmia, is a related neurobiological measure associated with antisocial behaviour, albeit less consistently (e.g., Beauchaine and colleagues⁸ and Dietrich and colleagues⁹). More recently, functional imaging research has been used in cross-sectional studies to compare brain activity between antisocial cases and controls. These studies have shown that prefrontal brain functioning seems to be impaired in antisocial behaviour, specifically in the orbitofrontal, dorsolateral frontal and anterior cingulate cortex (ACC; see Yang and colleagues¹⁰). Finally, a series of electrophysiological studies compared evoked potential in studies of delinquent populations and found aberrant brain responses to errors in delinquent individuals.^{11,12} However, it is unclear whether these neurobiological variables independently predict

Correspondence to: Henning Tiemeier, Harvard T. H. Chan School of Public Health, Department of Social and Behavioral Sciences, 677 Huntington Ave., Boston, MA 02115, United States; tiemeier@hsph.harvard.edu

Submitted May 28, 2020; Revised Sep. 7, 2020; Accepted Nov. 1, 2020

DOI: 10.1503/jpn.200103

© 2021 Joule Inc. or its licensors

delinquency. This knowledge is important for optimizing risk assessment and tailoring treatment for at-risk populations.

Although some studies have associated neurobiological measures with delinquent behaviour, few have employed such measures in formal prediction models to make adequate and reliable prognoses. In recent years, a few epidemiological studies have assessed the validity and utility of prediction models for delinquent behaviour using a formal approach. Aharoni and colleagues¹³ employed functional MRI (fMRI) in criminal offenders ($n = 96$) to investigate whether a neurobiological measure of inhibition (i.e., functional brain activity during a go/no-go task) added to the prediction of recidivism. They found that decreased activity in the ACC was predictive of rearrest. Thus far, the only study to include several neurobiological measures has been performed in a small subsample ($n = 45$) of the same cohort, in which 2 electroencephalography (EEG) measures of error-processing were also investigated. Error-related negativity (ERN; a measure of early error-processing) was not predictive of rearrest, but greater error positivity (Pe; a measure of late error-processing) did predict rearrest.¹⁴ When the EEG and fMRI measures were included in the same model, ACC activity no longer predicted rearrest, possibly because the various factors measured the same process. Because well-powered studies are lacking of the joint and individual contributions of several neurobiological parameters tested simultaneously, to date it is unknown whether different neurobiological parameters independently predict delinquency. The current study aimed to address this gap in knowledge.

We conducted this study in a large sample of young adults who were followed for up to 4 years to evaluate the incidence of delinquency. We assessed an extensive battery of demographic and self-report measures — including age, ethnicity, education, intelligence, internalizing and externalizing psychopathology, and aggressive and psychopathic behaviour — and evaluated their predictive power. This is important, because assessment of these variables is feasible in most instances and practitioners would certainly not include costly neurobiological measures in prediction models if they did not outperform interview or questionnaire measures. Our choice of neurobiological predictors was guided by 2 theoretical considerations. First, we included baseline heart rate and baseline respiratory sinus arrhythmia as physiologic measures of arousal. Following low arousal and fearlessness theory,¹⁵ low arousal is thought to be related to increased delinquency, a concept that is supported by empirical research (e.g., Latvala and colleagues⁷). Second, we included EEG measures of error-processing (ERN and Pe) and fMRI activity of the ACC during response inhibition. These measures have been studied as key indicators of impulsivity,^{13,14} persistent lack of restraint and consideration of consequences. Self-report questionnaires and observational tests of impulsivity are good predictors of delinquency, and neural measures of impulsivity capture variance independent of the former measures. The first aim of this study was to test the prospective association of demographic, behavioural and neurobiological measures with criminal recidivism. The second aim was to formally test the predictive power of different domains of variables. The main goal was to examine which neurobiological

parameters meaningfully added to the prediction of overall and serious recidivism, over and above demographic and behavioural measures.

Methods

Participants

Participants were 127 male young adults aged 18–27 years (mean \pm standard deviation [SD] = 21.92 \pm 2.40), recruited when they started a day-treatment program at De Nieuwe Kans (translated as “New Opportunities”) in Rotterdam, the Netherlands. De Nieuwe Kans provides a multimodal day-treatment program for “multi-problem” young adults who present with a plethora of issues, including a history of delinquency (81% had a criminal record), low or no income, poor job skills and drug use (53% have used cannabis regularly for at least 5 years). The program employs cognitive behavioural techniques and rehabilitation components, such as cognitive skills training, drug treatment and education (see Luijckx and colleagues¹⁶). This study was part of a larger cohort study that included 696 young adults followed for 14 months and interviewed on 4 separate occasions.¹⁶ Data presented here are from the baseline measurements and from judicial documentation from the Ministry of Security and Justice of the Netherlands. Because most participants had a criminal record (81%), our research was mainly a study of recidivism in particular, rather than of any delinquency. To ensure results were generalizable, we performed analyses on the total sample ($n = 127$) and supplemental analyses on only participants with a criminal record ($n = 103$).

All procedures were in accordance with the ethical standards of the institutional or national research committee, and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. The study was approved by the Medical Ethical Committee of the Vrije Universiteit (VU) Medical Center (registration number 2013.422–NL46906.029.13), and all participants provided written informed consent. Participants received a reimbursement of 30 euros for their participation in the fMRI, autonomic nervous system and EEG protocols.

Predictors

We organized baseline variables into 5 groups: demographics, education and intelligence, delinquency and drug use, behavioural traits, and neurobiological variables.

The demographics category included age and ethnicity (Western, Caribbean, Moroccan, Cape Verdean and other non-Western). The categorization of ethnicity was specific to the Netherlands; it was based on the definition from the Centraal Bureau voor de Statistiek (Statistics Netherlands) of the Dutch government.

We categorized education as primary only, junior secondary school and senior secondary school. We measured intelligence quotient using 4 subscales of the Wechsler Adult Intelligence Scale, third version (digit symbol coding, information, block design and arithmetic).¹⁷

We assessed history of delinquency as the number of past offences registered in the Research and Policy database from

the Judicial Documentation of the Ministry of Security and Justice in the Netherlands. We assessed years of regular (i.e., weekly) cannabis use and years of regular alcohol use with the Measurements in the Addictions for Triage and Evaluation Questionnaire.¹⁸ Because the data on alcohol use were heavily skewed, we presented them in categories (0, 1–5 and 6+ years).

Behavioural predictors consisted of reactive aggression and proactive aggression as assessed using the Reactive Proactive Aggression Questionnaire.^{19,20} It is a 23-item self-report measure of aggression scored on a 3-point Likert scale and yields reactive and proactive aggression scores. We assessed interpersonal, affective and behavioural psychopathic traits with the Youth Psychopathy Inventory–Short Version,²¹ an 18-item self-report measure scored on a 4-point Likert scale. We assessed internalizing and externalizing problems using the Adult Self-Report,²² which consists of 123 items scored on a 3-point Likert scale. Finally, we investigated the accuracies of both the Flanker-task (total accuracy) and the go/no-go task (accuracy on no-go trials) as predictors.

Autonomic nervous system

Neurobiological predictors including heart rate and respiratory sinus arrhythmia at rest were collected at the Erasmus Behavioural Laboratory of the Institute for Psychology at the Erasmus University Rotterdam, using the VU Ambulatory Monitoring System (VU-AMS).²³ We recorded both an electrocardiogram and an impedance cardiogram while participants watched a 5-minute excerpt from the video *Coral Sea Dreaming* (Small World Music Inc.).²⁴ Participants were seated in a comfortable chair in a sound-attenuated room with dimmed lights. We used a sampling rate of 1000 Hz and processed the data using VU-AMS Data, Analysis and Management Software (VU-DAMS).²⁵ We derived the R-peak time series via an automated detection algorithm within VU-DAMS. We manually checked for missing or incorrect R-wave peaks and abnormalities. Heart rate was assessed by automated counting of R peaks (beats per minute). We defined respiratory sinus arrhythmia as the longest period between heart beats during expiration minus the shortest period between heart beats during inspiration.

Electrophysiological measures of error-processing

We acquired ERN and Pe measures of error-processing during the same sessions as the measurements of heart rate and respiratory sinus arrhythmia. In the same room, participants performed an Eriksen–Flanker task previously used by Marhe and colleagues.²⁶ In short, participants responded to letter strings (HHHHH, SSSSS, HSHSH, SSHSS) by pressing a button with their left or right index finger, depending on the middle letter of the string. Each string was presented for 50 ms, the maximum response time was 650 ms, and a stimulus was shown once every 1450 ms. In total, 400 trials were presented to each participant. We defined ERN as the difference wave (difference in mean amplitude between correct and incorrect trials) in a post-response window of 25 ms to 100 ms on the FCz electrode. We defined Pe as the difference wave in a post-response window of 250 ms to 400 ms on the

Pz electrode. We used the pre-response period (–100 ms to 0 ms) as a baseline. We used a Biosemi ActiveTwo System amplifier to measure EEG. We placed silver chloride (Ag/AgCl) electrodes on the scalp using the international 10–20 system, with 2 electrodes on the left and right mastoids as a reference. We assessed vertical and horizontal electro-oculograms to control for ocular artifacts. Signals were digitized with a sampling rate of 512 Hz and 24-bit analogue-to-digital conversion, and digitally filtered offline using a low cut-off of 0.15 Hz and a high cut-off of 30 Hz (24 dB per octave slope) with BrainVision Analyzer 2. Additional artifact rejection ($\pm 100 \mu\text{V}$) was performed automatically. We used the EEG data if at least 6 error trials were available for analysis.²⁷ The mean (\pm SD) number of trials used after inclusion was 62.05 ± 43.00 .

fMRI measures of inhibition

We assessed ACC activity during inhibition using fMRI. Participants performed a go/no-go task previously used by Luijten and colleagues.²⁸ In short, participants responded to letters (go trials) presented at 1 Hz, but had to refrain from responding when the letter was the same as the previous one (no-go trials). In total, 817 go and 110 no-go trials were presented. We defined ACC activity during inhibition as the brain activity in an region of interest defined a priori (14 mm radius-sphere at $x, y, z = 3, 24, 33$)¹³ during the contrast of commission errors versus correct hits, which captures ACC recruitment by either error feedback (or by simple processing or awareness of an oddball stimulus) while holding constant motor behaviour. In a previous study, this contrast predicted rearrest in male offenders.¹³ We also investigated the additive predictive power of the contrast of correct rejections versus correct hits, as a possible indicator of brain recruitment by successful inhibition (or by simple oddball detection). We collected data on a 3 T GE Healthcare MRI scanner at the Erasmus Medical Center, Rotterdam, on a different day from the autonomic and EEG measurements. We acquired structural T_1 -weighted images with a fast-spoiled gradient pulse sequence in 180 sequential sagittal (superior/inferior) slices, with a thickness of 1.0 mm. The repetition time was 6.4 ms, the echo time was 2.8 ms, the flip angle was 12 degrees, the field of view was 240 mm and the matrix size 240×240 mm. We acquired blood-oxygen-level-dependent T_2^* -weighted images axially (right/left) with an echo planar imaging gradient echo pulse sequence in 42 slices of 3.5 mm with a slice spacing of 0.5 mm. The repetition time was 2000 ms, the echo time was 30 ms, the flip angle was 80° , the field of view was 220 mm and the matrix size 64×64 mm. We analyzed functional imaging data using SPM12 (www.fil.ion.ucl.ac.uk/spm/). As preprocessing steps for each participant, functional images were realigned and unwarped, and the structural scan was segmented and coregistered to the mean T_2^* -weighted image. Images were then normalized to the Montreal Neurological Institute template and smoothed with a Gaussian filter (8 mm full width at half maximum). We modelled 4 conditions and added 6 movement parameters as covariates of no interest. We extracted region-of-interest data using the MarsBaR toolbox for SPM (<http://marsbar.sourceforge.net/>).

We used fMRI data if at least 8 error trials were available for analysis.²⁹ No participants had fewer than 8 error trials, and the mean (\pm SD) number of used trials was 56.25 ± 56.00 . We tested the contrast of commission errors versus correct hits and the contrast of correct rejections versus correct hits.

Outcomes

We assessed information on 2 outcome measures during follow-up. The first outcome measure was time to arrest for any criminal offence. This included (as defined by the Research and Documentation Centre of the Ministry of Security and Justice of the Netherlands) relatively minor offences (e.g., resistance, drunk driving, possession of a weapon) and serious offences (those with a minimum 4-year sentence; e.g., assault, aggravated theft, rape).³⁰ The second outcome measure was time to arrest for serious offences. We obtained data from the Research and Policy database Judicial Documentation of the Ministry of Security and Justice of the Netherlands, and included arrests from the baseline measurement until January 2018. The median follow-up time was 31 months (range = 14–47 months).

Data analysis

We imputed to 40 complete data sets³¹ using SPSS 21 (IBM) for variables with a maximum missingness of 30%; no outcomes were imputed. First, we performed Cox proportional hazard models to assess potential predictors, adjusting for age for each of the 2 outcomes. We selected all variables that were significant at $p < 0.10$ in the model for any offence or serious offences and included them in the fully adjusted Cox models; we used a liberal value of $p < 0.10$ because use of a more traditional level of $p < 0.05$ often failed to identify important variables.³² We tested the proportional hazard assumption by testing the associations between Schoenfeld residuals³³ and time.

We standardized all effect sizes except for age. Second, using the risksetROC package³⁴ in R 3.5.1, we calculated (areas under) time-dependent receiver operating characteristic (ROC) curves to compare the predictive performance of the different variable groups over time and to assess whether adding variable groups increased predictive performance above simpler models that included fewer variables. We used time-dependent ROC curves to represent the relationship of time-dependent sensitivity and specificity, and areas under the ROC curves (AUCs) to measure the probability that the predictive value for a randomly selected offender exceeded the predictive value for a randomly selected nonoffender. Finally, we computed C-indexes for prediction at 6, 12, 18, 24 and 30 months, which was the concordance statistic over a time-period based on the weighted averages of the AUCs. The concordance statistic is best conceptualized as an AUC. It is commonly used to evaluate risk models in survival analysis, where data may be censored. Technically, it is a measure of goodness of fit for binary outcomes in a regression model. It is the fraction of pairs in the data, where the observation with the higher survival time also has the higher probability of survival as predicted by our model. It represents the predictive ability of determinant blocks: a C-index above 0.70 indicates a good model, and a C-index above 0.80 indicates a strong model.³²

Results

In the follow-up period, 61 (48.0%) of the 127 participants were arrested for any offence, of which 36 (28.3% of total) were arrested for a serious offence. In the age-adjusted models, the following variables predicted any offences and serious offences: past offences, cannabis use, reactive aggression, proactive aggression, heart rate and ERN (EEG and fMRI findings are shown in Figure 1). In contrast, ethnicity (Moroccan versus Western, and Cape Verdean versus

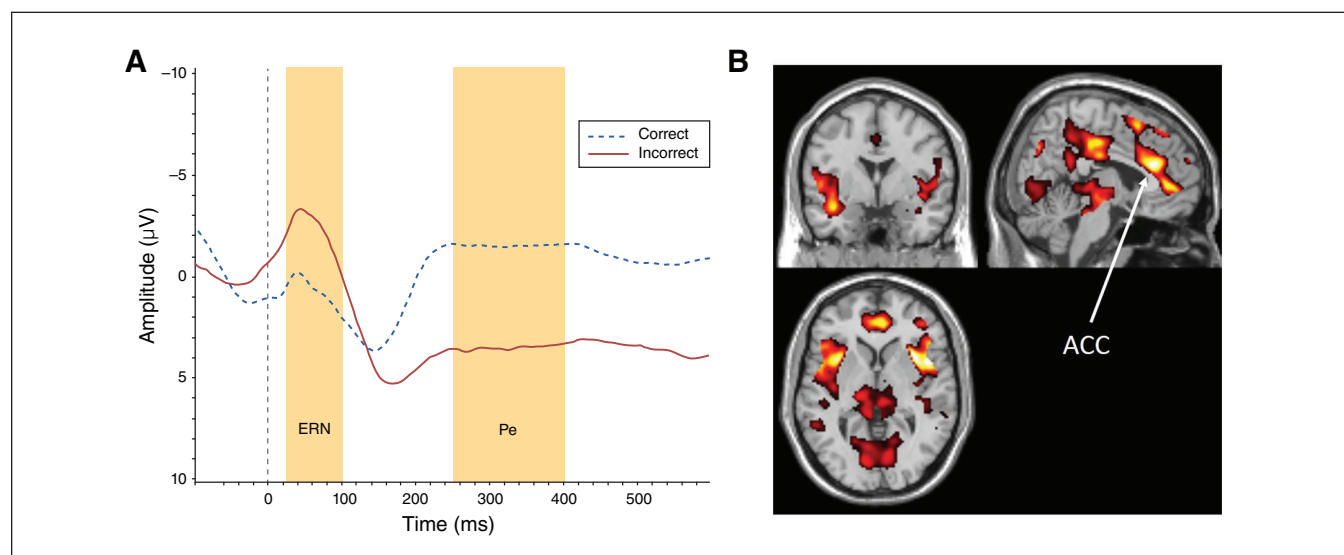


Fig. 1: (A) Electroencephalographic waveforms in response to correct and incorrect trials. (B) Whole-brain ($p_{FWE} < 0.05$) hemodynamic activity during the contrast of commission errors versus correct hits ($x, y, z = 49, 64, 38$). ACC = anterior cingulate cortex; ERN = error-related negativity; FWE = family-wise error corrected; Pe = error positivity.

Western) predicted only any offences, whereas behavioural psychopathic traits and the Pe predicted only serious offences. See Table 1 for full results with effect estimates. We included these 9 predictors in the mutually adjusted models, retaining age as a covariate.

In the final model, 3 predictors were significantly related to any offence: ethnicity Moroccan versus Western ($p < 0.05$, RR 3.07, 95% CI 1.04–9.10), past offences ($p < 0.001$, RR 1.45, 95% CI 1.17–1.79), and a smaller ERN ($p < 0.05$, RR 1.61, 95% CI 1.13–2.28; Table 2). Likewise, 3 predictors were significantly related to serious offences: past offences ($p < 0.001$, RR 1.71, 95% CI 1.33–2.23), behavioural psychopathic traits ($p < 0.05$,

RR 1.54, 95% CI 1.00–2.36) and low heart rate ($p < 0.05$, RR 0.62, 95% CI 0.38–0.99; Table 2).

In the supplemental analysis of the subgroup with a criminal record ($n = 103$), results were similar, although ethnicity failed to reach significance and ERN did reach significance ($p = 0.046$ in the subgroup analysis, compared with $p = 0.08$ in the main analysis); see Appendix 1, Table S1, available at jpn.ca/200103-a1.

Any offence

For each of the 4 variable groups, we calculated AUCs for different time points as a measure of predictive power over

Table 1: Descriptive statistics and univariate Cox regressions

Characteristic	Baseline*	General offences		Serious offences	
		RR (95% CI)	<i>p</i> value	RR (95% CI)	<i>p</i> value†
Demographics					
Age	21.98 ± 2.4	1.01 (0.91–1.13)	0.85	0.94 (0.81–1.08)	0.37
Ethnicity					
Western	22 (17.3)		ref		ref
Caribbean	55 (43.3)	1.98 (0.82–4.78)	0.13	1.70 (0.57–5.05)	0.34
Moroccan	21 (16.5)	2.70 (1.02–7.13)	0.045	2.07 (0.62–6.87)	0.24
Cape Verdean	8 (6.3)	2.78 (0.89–8.69)	0.078	1.30 (0.24–7.15)	0.76
Other non-Western	21 (16.5)	1.29 (0.44–3.72)	0.64	1.19 (0.32–4.44)	0.80
Education					
Senior secondary education	23 (18.1)		ref		ref
Junior secondary education	48 (37.8)	1.15 (0.54–2.45)	0.72	0.62 (0.23–1.67)	0.35
Primary education	56 (44.1)	1.45 (0.69–3.04)	0.33	1.09 (0.44–2.71)	0.85
Intelligence quotient	82.07 ± 10.2	0.82 (0.64–1.07)	0.14	0.86 (0.61–1.22)	0.40
Delinquency and drug use					
Past offences, <i>n</i>	5.06 ± 5.70	1.51 (1.27–1.79)	< 0.001	1.72 (1.40–2.10)	< 0.001
Cannabis use, yr	4.28 ± 3.83	1.27 (0.98–1.65)	0.069	1.43 (0.93–2.05)	0.051
Alcohol use, yr					
0	73 (57.5)		ref		ref
1–5	31 (24.4)	0.88 (0.45–1.71)	0.97	0.93 (0.40–2.15)	0.86
6+	23 (18.1)	0.81 (0.40–1.65)	0.57	0.90 (0.34–2.37)	0.84
Behaviour					
Reactive aggression	11.28 ± 4.65	1.53 (1.17–2.00)	0.002	1.57 (1.11–2.22)	0.011
Proactive aggression	5.28 ± 4.28	1.38 (1.10–1.72)	0.005	1.47 (1.12–1.93)	0.005
Psychopathy, interpersonal	11.27 ± 3.86	0.90 (0.70–1.15)	0.38	1.07 (0.79–1.46)	0.66
Psychopathy, affective	10.83 ± 3.51	1.04 (0.82–1.31)	0.78	1.13 (0.83–1.53)	0.44
Psychopathy, behavioural	12.21 ± 3.14	1.07 (0.83–1.38)	0.60	1.38 (1.00–1.89)	0.047
Internalizing problems	72.52 ± 24.19	0.99 (0.77–1.27)	0.92	0.91 (0.66–1.26)	0.57
Externalizing problems	69.22 ± 23.83	1.07 (0.82–1.40)	0.63	1.17 (0.82–1.69)	0.39
Accuracy, Flanker	0.82 ± 0.16	0.82 (0.63–1.06)	0.13	0.74 (0.54–1.03)	0.11
Accuracy, go/no-go	0.50 ± 0.16	1.05 (0.77–1.43)	0.77	1.12 (0.75–1.68)	0.58
Neurobiology					
Heart rate	65.78 ± 9.11	0.70 (0.52–0.93)	0.014	0.53 (0.35–0.79)	0.002
Respiratory sinus arrhythmia	95.70 ± 42.48	1.17 (0.94–1.46)	0.17	1.22 (0.91–1.64)	0.18
Error-related negativity	-5.17 ± 4.72	1.47 (1.10–1.96)	0.008	1.53 (1.05–2.25)	0.029
Positivity error	6.13 ± 5.07	0.83 (0.64–1.08)	0.17	0.70 (0.49–0.99)	0.043
ACC activity (commission errors v. correct hits)	2.87 ± 2.34	1.05 (0.81–1.36)	0.74	1.27 (0.90–1.80)	0.18
ACC activity (correct rejections v. correct hits)	2.56 ± 1.85	0.90 (0.78–1.05)	0.18	0.90 (0.73–1.10)	0.28

ACC = anterior cingulate cortex; CI = confidence interval; RR = risk ratio; SD = standard deviation.

*Values are presented as mean ± SD for continuous variables or *n* (%) for categorical variables.

†Considered significant at $p < 0.10$.

Table 2: Full Cox models for general and serious offences in the total sample (n = 127)

Characteristic	General offences		Serious offences	
	RR (95% CI)	p value	RR (95% CI)	p value*
Demographics				
Age	0.98 (0.86–1.12)	0.79	0.88 (0.74–1.05)	0.15
Ethnicity				
Western		ref		ref
Caribbean	2.25 (0.86–5.93)	0.10	1.71 (0.53–5.57)	0.37
Moroccan	3.07 (1.04–9.10)	0.043	2.07 (0.55–7.81)	0.28
Cape Verdean	2.65 (0.73–9.69)	0.14	0.75 (0.10–5.50)	0.78
Other non-Western	1.62 (0.52–5.08)	0.41	1.27 (0.29–5.54)	0.75
Delinquency and drug use				
Past offences, n	1.45 (1.17–1.79)	< 0.001	1.71 (1.33–2.23)	< 0.001
Cannabis use, yr	1.21 (0.90–1.63)	0.20	1.39 (0.92–2.09)	0.12
Behaviour				
Reactive aggression	1.26 (0.87–1.81)	0.22	0.95 (0.60–1.52)	0.84
Proactive aggression	1.01 (0.72–1.42)	0.94	1.07 (0.70–1.62)	0.76
Psychopathy, behavioural	1.08 (0.78–1.50)	0.63	1.54 (1.00–2.36)	0.049
Neurobiology				
Heart rate	0.88 (0.65–1.19)	0.39	0.62 (0.38–0.99)	0.023
Error-related negativity	1.61 (1.13–2.28)	0.008	1.52 (0.96–2.40)	0.08
Positivity error	0.93 (0.69–1.26)	0.65	0.75 (0.49–1.16)	0.27

CI = confidence interval; RR = risk ratio.
*Considered significant at $p < 0.05$.

time. Figure 2A shows that the AUCs for all groups of variables were fairly stable over time. The neurobiological measures, as well as delinquency and cannabis use, performed best (C_{30} for 30 months [C_{30}] = 0.64), the behaviour group performed similarly (C_{30} = 0.63), and age and ethnicity performed worst (C_{30} = 0.57; Table 3).

Figure 2B shows the predictive power of models step-wise, including more variable groups. The figure demonstrates that the neurobiological measures added to the predictive power above all other variables, and that the added predictive power decreased slightly over time. The model without neurobiological measures performed moderately well (C_{30} = 0.68), but adding neurobiological measures increased the predictive power (C_{30} = 0.72).

The results of the supplemental analysis on the subgroup with a criminal record were very similar, with identical C -indexes for the complete model (Appendix 1, Table S2).

Serious offences

Figure 2C shows the predictive power of the variable groups over time, which was again fairly stable for each variable group. Table 3 shows that neurobiological measures (C_{30} = 0.70) and delinquency and cannabis use (C_{30} = 0.68) outperformed demographic (C_{30} = 0.60) and behavioural measures (C_{30} = 0.64). Figure 2D demonstrates that the neurobiological predictors consistently added predictive power above the other variables. The model without neurobiological predictors performed well (C_{30} = 0.75), but adding neurobiological measures also increased its predictive performance (C_{30} = 0.80).

The results of the supplemental analysis on the subgroup with a criminal record were very similar, with identical C -indexes for the complete model (see Appendix 1, Table S2).

Discussion

In this study, we showed that in a sample of delinquent individuals, models including demographic variables, delinquency and drug use, behavioural traits and neurobiological variables did a good job of predicting overall and serious recidivism. We assessed the combined and independent predictive value of several multimodal neurobiological measures: resting heart rate and respiratory sinus arrhythmia, 2 EEG measures of error-processing and an fMRI measure of inhibition. Three of these neurobiological measures were associated longitudinally with criminal recidivism in age-adjusted analyses, and more importantly, these measures had incremental predictive value above traditional risk factors.

A wide array of variables univariately predicted recidivism, but understandably, most predictors failed to reach significance in the final, mutually adjusted models. Delinquency is a complex multifactorial phenomenon, and no single variable reliably predicts recidivism except number of past offences. Low resting heart rate predicted serious recidivism, in line with low arousal and fearlessness theory.¹⁵ Another previous study found a similar association that was stronger for violent than for nonviolent crimes.⁷ Self-reported behavioural psychopathic traits also independently contributed to the prediction of serious recidivism. Concerning overall recidivism, ethnicity and the ERN were the only

additional significant predictors; a smaller ERN (indicating diminished early error-processing) was related to increased recidivism. Although the ERN measure failed to reach significance in the model of serious recidivism, effect sizes in both models were very similar (RR 1.61 for overall recidi-

vism and RR 1.52 for serious recidivism). In a study by Steele and colleagues,¹⁴ ERN was not predictive of recidivism, but this study was limited by its small sample size ($n = 45$). In line with findings that a smaller ERN is related to antisocial behaviour in general,³⁵ our results suggest that

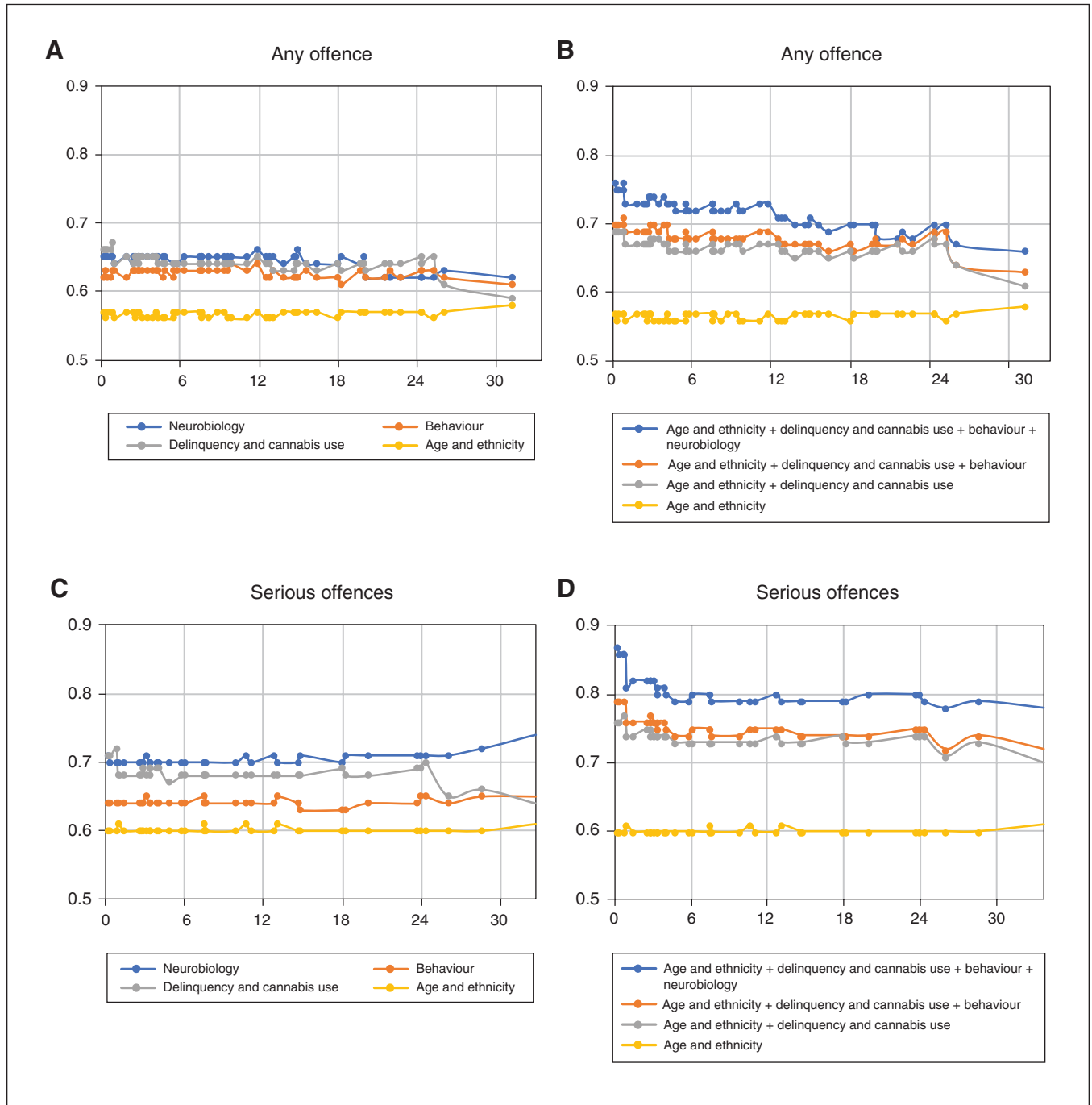


Fig. 2: Areas under time-dependent receiver-operating-characteristic curves for the prediction of (A) any offence by different variable groups; (B) any offence by models that include multiple variable groups; (C) serious offences by different variable groups; (D) serious offences by models that include multiple variable groups. The categorization of ethnicity was specific to the Netherlands; it was based on the definition from the Centraal Bureau voor de Statistiek (Statistics Netherlands) of the Dutch government (www.cbs.nl/en-gb/onze-diensten/methods/definitions/person-with-a-migration-background).

Table 3: C-indexes for different time intervals in the full sample (n = 127)

Parameter	Follow-up, months				
	6	12	18	24	30
General offences					
Group					
Age and ethnicity	0.57	0.57	0.57	0.57	0.57
Delinquency and cannabis use	0.65	0.65	0.64	0.64	0.64
Behaviour	0.63	0.63	0.63	0.63	0.63
Neurobiology	0.65	0.65	0.65	0.65	0.64
Model					
Age and ethnicity	0.57	0.57	0.57	0.57	0.57
Age and ethnicity + delinquency and cannabis use	0.68	0.67	0.67	0.67	0.67
Age and ethnicity + delinquency and cannabis use + behaviour	0.69	0.69	0.69	0.68	0.68
Age and ethnicity + delinquency and cannabis use + behaviour + neurobiology	0.74	0.73	0.73	0.72	0.72
Serious offences					
Group					
Age and ethnicity	0.60	0.60	0.60	0.60	0.60
Delinquency and cannabis use	0.69	0.69	0.69	0.69	0.68
Behaviour	0.64	0.64	0.64	0.64	0.64
Neurobiology	0.70	0.70	0.70	0.70	0.70
Model					
Age and ethnicity	0.60	0.60	0.60	0.60	0.60
Age and ethnicity + delinquency and cannabis use	0.74	0.74	0.74	0.74	0.74
Age and ethnicity + delinquency and cannabis use + behaviour	0.76	0.76	0.76	0.75	0.75
Age and ethnicity + delinquency and cannabis use + behaviour + neurobiology	0.82	0.81	0.81	0.81	0.80

the inability to adequately process adverse consequences may underlie the continuation of criminal behaviour despite its negative consequences.

In contrast with a previous study,¹³ we did not find an independent association between recidivism and either error-awareness-related or inhibition-related ACC recruitment. Because we employed a very similar task and found expected brain activity, it seems unlikely that we measured a different process. The difference may have been due to variation between investigated samples. Aharoni and colleagues¹³ studied incarcerated offenders with a large age range (20–52 years) who performed the task before being released from prison, whereas our sample consisted of young adults (18–27 years) who were not incarcerated. The relevance of inhibitory processes for delinquent behaviour may increase with age, or the setting of the experiment may influence results. Another possibility for the incongruent findings is the methodological differences between the present study and that of Aharoni and colleagues. We employed a different motion-correction algorithm, used a different pulse sequence and conducted scans on a machine from a different vendor. In general, replication of fMRI findings has proven to be difficult,³⁶ complicating the interpretation of discrepant results.

Although Pe did not reach significance in the mutually adjusted models, we did find that a smaller Pe univariately predicted serious recidivism, and the effect size remained similar in the full model. This finding was in line with research that found associations between a smaller Pe and antisocial behaviour in adolescents (16–20 years),³⁷ but at odds with findings by Steele and colleagues,¹⁴ who found that a larger

Pe was predictive of rearrest in a sample of 45 incarcerated adults. Again, the age differences between participant groups may have been a relevant factor in explaining these discrepancies. As well, the 2 studies had sizable methodological differences. Whereas we employed a difference wave analysis to ascertain EEG variables, Steele and colleagues employed principal component analysis to reduce data before prediction. These different analytical choices may have affected results, although Maurer and colleagues³⁷ employed the same principal component analysis and reported results more in line with the current findings.

For both overall and serious recidivism, neurobiological measures had greater predictive value than the other groups of variables, and importantly, they improved predictive performance when added to the models without neurobiological measures. Although our study warrants extension and replication, our finding related to low heart rate was in accordance with theory and previous research in both population-based and at-risk samples, bolstering the idea that heart rate can be a reliable indicator of future delinquency. Concerning EEG and MRI measures, it cannot definitively be concluded which measures of executive functioning have greater predictive potential based on the available studies, but using EEG rather than fMRI measures has considerable practical advantages: the former is relatively cheap, less time-consuming, technically relatively easy to acquire and less invasive. Measurements of the autonomic nervous system have the same benefits. Overall, our findings suggest that adding neurobiological measures to risk assessments may be a viable way of improving predictive accuracy for recidivism. However, at present

not enough data are available to justify implementing this type of risk assessment in judicial practice, and studies in more regulated environments such as prisons are warranted to assess the practical utility of such risk assessment. Future research would benefit from larger sample sizes and longer follow-up periods to assess whether prediction is stable across time, and whether neurobiological measures are truly viable as risk factors (for a discussion, see Poldrack and colleagues³⁸).

Limitations

This study had several limitations. First, we included only male participants, and as such our findings cannot be generalized to female populations. Second, although predictor choice was theoretically informed and our sample size was large for this type of study, the number of investigated predictors limited the power of the analysis. Third, our follow-up time was up to 4 years, leaving it an open question whether the predictive models would hold for longer time periods.

Conclusion

A group of 3 multimodal neurobiological measures added to the prediction of overall and serious recidivism beyond demographic and behavioural measures. As the most feasible measures, autonomic functioning and electroencephalography are also most predictive; in the future they may be viable for including in risk assessment.

Acknowledgements: This research project is funded by De Verre Bergen Foundation. De Verre Bergen Foundation is a venture philanthropy organization that aims to build a better Rotterdam through substantial investments in innovative, impactful social ventures. The financier was not involved in the design of the study or the drafting of the manuscript. Furthermore, the financier was not involved in the process of data collection, analysis or interpretation. Contact information: Nanne Boonstra, Parklaan 22, 3016 BB Rotterdam, the Netherlands; Tel: 0031 10 209 2000; email: nboonstra@sduvb.com.

Affiliations: From the Amsterdam University Medical Centers, Department of Child and Adolescent Psychiatry, Amsterdam, Netherlands (Zijlmans, Marhe, Bevaart, van Duin, Luijks, Popma); the Erasmus University Rotterdam, Department of Psychology, Education and Child Studies, Rotterdam, Netherlands (Marhe, Franken); the Harvard T. H. Chan School of Public Health, Department of Social and Behavioral Sciences, Boston, Massachusetts (Tiemeier); the Erasmus University Medical Center, Department of Child and Adolescent Psychiatry, Rotterdam, Netherlands (Tiemeier); and Leiden University, Department of Criminal Law and Criminology, Leiden, Netherlands (Popma).

Competing interests: None declared.

Contributors statement: All authors designed the study. J. Zijlmans, R. Marhe, L. Van Duin and M.-J. Luijks acquired the data, which J. Zijlmans, R. Marhe, H. Tiemeier and A. Popma analyzed. J. Zijlmans and R. Marhe wrote the article, which all authors reviewed. All authors approved the final version to be published and can certify that no other individuals not listed as authors have made substantial contributions to the paper.

Content licence: This is an Open Access article distributed in accordance with the terms of the Creative Commons Attribution (CC BY-NC-ND 4.0) licence, which permits use, distribution and reproduction in any medium, provided that the original publication is

properly cited, the use is noncommercial (i.e., research or educational use), and no modifications or adaptations are made. See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>

References

1. McCollister KE, French MT, Fang H. The cost of crime to society: new crime-specific estimates for policy and program evaluation. *Drug Alcohol Depend* 2011;108:98-109.
2. Cottle CC, Lee RJ, Heilbrun K. The prediction of criminal recidivism in juveniles: a meta-analysis. *Crim Justice Behav* 2001;28:367-94.
3. Lorber MF. Psychophysiology of aggression, psychopathy, and conduct problems: a meta-analysis. *Psychol Bull* 2004;130:531-52.
4. Ortiz J, Raine A. Heart rate level and antisocial behavior in children and adolescents: a meta-analysis. *J Am Acad Child Adolesc Psychiatry* 2004;43:154-62.
5. Wadsworth MEJ. Delinquency, pulse rates and early emotional deprivation. *Br J Criminol* 1976;16:245-56.
6. Raine A, Venables PH, Mednick SA. Low resting heart rate at age 3 years predisposes to aggression at age 11 years: evidence from the Mauritius Child Health Project. *J Am Acad Child Adolesc Psychiatry* 1997;36:1457-64.
7. Latvala A, Kuja-Halkola R, Almqvist C, et al. A longitudinal study of resting heart rate and violent criminality in more than 700 000 men. *JAMA Psychiatry* 2015;72:971-8.
8. Beauchaine TP, Hong J, Marsh P. Sex differences in autonomic correlates of conduct problems and aggression. *J Am Acad Child Adolesc Psychiatry* 2008;47:788-96.
9. Dietrich A, Riese H, Sondejker FEPL, et al. Externalizing and internalizing problems in relation to autonomic function: a population-based study in preadolescents. *J Am Acad Child Adolesc Psychiatry* 2007;46:378-86.
10. Yang Y, Raine A. Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: a meta-analysis. *Psychiatry Res* 2009;174:81-8.
11. Brazil I a, de Bruijn ER a, Bulten BH, et al. Early and late components of error monitoring in violent offenders with psychopathy. *Biol Psychiatry* 2009;65:137-43.
12. Steele VR, Maurer JM, Bemat EM, et al. Error-related processing in adult males with elevated psychopathic traits. *Pers Disord* 2015;20:1878-90.
13. Aharoni E, Vincent GM, Harenski CL, et al. Neuroprediction of future rearrest. *Proc Natl Acad Sci U S A* 2013;110:6223-8.
14. Steele VR, Claus ED, Aharoni E, et al. Multimodal imaging measures predict rearrest. *Front Hum Neurosci* 2015;9:425.
15. Raine A. *The psychopathology of crime: criminal behavior as a clinical disorder*. San Diego (CA): Academic Press; 1993.
16. Luijks MA, Bevaart F, Zijlmans J, et al. A multimodal day treatment program for multi-problem young adults: study protocol for a randomized controlled trial. *Trials* 2017;18:225.
17. Blyler CR, Gold JM, Iannone VN, et al. Short form of the WAIS-III for use with patients with schizophrenia. *Schizophr Res* 2000;46:209-15.
18. Schippers GM, Broekman TG, Buchholz A, et al. Measurements in the Addictions for Triage and Evaluation (MATE): an instrument based on the World Health Organization family of international classifications. *Addiction* 2010;105:862-71.
19. Cima M, Raine A, Meesters C, et al. Validation of the Dutch Reactive Proactive Questionnaire (RPQ): differential correlates of reactive and proactive aggression from childhood to adulthood. *Aggress Behav* 2013;39:99-113.
20. Raine A, Dodge K, Loeber R, et al. The Reactive-Proactive Aggression Questionnaire. *Aggress Behav* 2006;232:159-71.
21. Van Baardewijk Y, Andershed H, Stegge H, et al. Development and tests of short versions of the Youth Psychopathic Traits Inventory and the Youth Psychopathic Traits Inventory-Child Version. *Eur J Psychol Assess* 2010;26:122-8.
22. Achenbach TM, Rescorla LA. *Manual for the ASEBA adult forms and profiles*. Burlington (VT): University of Vermont, Research Center for Children, Youth and Families; 2003.
23. Klaver C, De Geus E, De Vries J. Ambulatory monitoring system. In: Maarsse FJ, Akkerman AE, Brand AN, et al, editors. *Computers in psychology: applications, methods, and instrumentation*. Amsterdam: Swets and Zeitlinger; 1994: 254-68.

24. Piferi RL, Kline KA, Younger J, et al. An alternative approach for achieving cardiovascular baseline: viewing an aquatic video. *Int J Psychophysiol* 2000;37:207-17.
25. *Data analysis and management software (DAMS) for the Vrije Universiteit ambulatory monitoring system (VU-AMS). Manual version 1.5.* Amsterdam: Vrije Universiteit; 2019. Available: www.vu-ams.nl/fileadmin/user_upload/manuals/VU-DAMS_manual_v1.5.pdf (accessed 2021 Mar. 17).
26. Marhe R, van de Wetering BJM, Franken IH a. Error-related brain activity predicts cocaine use after treatment at 3-month follow-up. *Biol Psychiatry* 2013;73:782-8.
27. Olvet DM, Hajcak G. The stability of error-related brain activity with increasing trials. *Psychophysiology* 2009;46:957-61.
28. Luijten M, Veltman DJ, Hester R, et al. The role of dopamine in inhibitory control in smokers and non-smokers: a pharmacological fMRI study. *Eur Neuropsychopharmacol* 2013;23:1247-56.
29. Steele VR, Anderson NE, Claus ED, et al. Neuroimaging measures of error-processing: extracting reliable signals from event-related potentials and functional magnetic resonance imaging. *Neuroimage* 2016;132:247-60.
30. Wartna B, Blom M, Tollenaar N. *The Dutch recidivism monitor.* Den Haag: Research and Documentation Centre of the Ministry of Security and Justice; 2011. Available: www.wodc.nl/binaries/the-dutch-recidivism-monitor-2011-20110803_tcm28-78143.pdf (accessed 2021 Mar. 17).
31. White IR, Royston P, Wood AM. Multiple imputation using chained equations: issues and guidance for practice. *Stat Med* 2011;30:377-99.
32. Hosmer D, Lemeshow S. *Applied logistic regression.* 2nd ed. New York: John Wiley & Sons; 2000.
33. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika* 1982;69:239-41.
34. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves 1. *Biometrics* 2005;61:92-105.
35. Hall JR, Bernat EM, Patrick CJ. Externalizing psychopathology and the error-related negativity. *Psychol Sci* 2007;18:326-33.
36. Bennett CM, Miller MB. How reliable are the results from functional magnetic resonance imaging? *Ann N Y Acad Sci* 2010;1191:133-55.
37. Maurer JM, Steele VR, Cope LM, et al. Dysfunctional error-related processing in incarcerated youth with elevated psychopathic traits. *Dev Cogn Neurosci* 2015;57:742-68.
38. Poldrack RA, Monahan J, Imrey PB, et al. Predicting violent behavior: what can neuroscience add? *Trends Cogn Sci* 2018;22:111-23.