Predicting forensic admission among the mentally ill in a multinational setting: A Bayesian modelling approach

Erkki J.O. Soini\textsuperscript{a,b,c,*}, Tarja Rissanen\textsuperscript{a,d}, Jari Tiihonen\textsuperscript{d,e}, Sheilagh Hodgins\textsuperscript{f}, Markku Eronen\textsuperscript{g}, Olli-Pekka Ryynänen\textsuperscript{h,i}

\textsuperscript{a}Department of Health Policy and Management, University of Kuopio, Kuopio, Finland
\textsuperscript{b}Department of Social Pharmacy (Centre for Pharmaceutical Policy and Economics), University of Kuopio, Kuopio, Finland
\textsuperscript{c}ESiOR Oy, P.O. Box 1188, FI-70211 Kuopio, Finland
\textsuperscript{d}Niuvanniemi Hospital, Kuopio, Finland
\textsuperscript{e}Department of Forensic Psychiatry, University of Kuopio, Finland
\textsuperscript{f}Department of Forensic Mental Health Science, Institute of Psychiatry, London, UK
\textsuperscript{g}Vanha Vaasa Hospital, Vaasa, Finland
\textsuperscript{h}General Practice Unit, Kuopio University Hospital, Kuopio, Finland
\textsuperscript{i}Department of Public Health and Clinical Nutrition, University of Kuopio, Kuopio, Finland

\textbf{A B S T R A C T}

Our objective was to explore protective and risk factors for a forensic admission among the mentally ill. The most influential risk factors included violent crime prior to the crime that leads to the index hospitalization, conviction of the biological father, and no use of psychotropic medications before the age of 18. The main contributions of this study in comparison to previous studies in this domain included the use of multinational setting, greedy Bayesian algorithm, generalized country-independent factors, a merger model with high predictive performance, various measures of predictive performance including relative predictive value, and independent learning and test sets.

\textbf{Keywords:}
Forensic psychiatry
Criminality
Violence
Model merging
Prediction
Knowledge validation

1. Introduction

Persons with severe mental illness, and most particularly those with schizophrenia and schizo-affective disorder, are at increased risk, as compared to the general population, to commit violent crimes \cite{1}. This is a robust finding. It has been reported by several independent research groups working in industrialized \cite{2-6} and underdeveloped countries \cite{7} with distinct cultures, health, social service and criminal justice systems, who have examined different cohorts and samples using various experimental designs including prospective, longitudinal investigations on birth cohorts \cite{2,3,5} and population cohorts \cite{6}, follow-up studies comparing patients and their neighbours \cite{8}, random samples of incarcerated offenders \cite{9}, and complete cohorts of homicide offenders \cite{10,11}. The results using official criminal records of convictions for violent crimes \cite{3} and self and collateral reports of aggressive behaviour \cite{4} concur in observing an increased prevalence of violent crimes among people with schizophrenia.

Violent crimes by persons with severe mental illness pose a significant human and financial burden on society, and especially on the health and criminal justice systems. If the factors associated with this violence could be identified,
management and treatment programmes could be established to prevent the violent behaviour. A number of different strategies and instruments have been developed in an effort to assess the risk of violent behaviour among persons with severe mental illness [12–15]. Most are based on a wealth of research showing that both historical and current factors play a role. Historical factors that have been associated with violent behaviour by persons with severe mental illness include antisocial behaviour, criminal offending, and substance misuse among parents, conduct problems in childhood and early adolescence, substance misuse in adolescence, and aggressive behaviour towards others or violent crime. Clinical factors that have been linked to violent behaviour in this population include symptoms of psychosis, depression, and non-compliance with medication. Other current factors include substance misuse and inadequate community care [16,17].

However, predicting the likelihood of violent behaviour by persons with severe mental illness has proven very difficult. At the best, previous predictions have produced area under ROC curve (AUC, $\hat{A}$) estimates below 0.89 for learning and 0.71 for test sets in selected subgroups or 0.96 for complete datasets [18] and 0.80 for a stratum including only mentally ill subjects [19]. Unlike previous studies, the present study tackled this specific prediction problem using Bayesian statistics and included the use of informative priors and a new type of data triangulation process in the model construction. In addition, design was uncommon for psychiatric domain – independent datasets were used for model learning and tests. However, the use of the Bayesian approach in psychiatric research is not new. In 1963, Overall [20] reported modelling for 13 different types of psychotic symptoms. The Bayesian approach has lead to findings related to aetiology as well as patient’s treatment and diagnosis setting [21–25].

The present study aimed at identifying risk and protective factors associated with admission to a forensic psychiatric hospital and to create a robust merger model.

Unlike in the most of previous studies of violent crimes among people with severe mental illness, the modelling carried out here was based on a multinational data, and both Bayesian and frequentist (logistic regression) methods were used and compared.

2. Methods

We have previously shown that the forensic patients recruited in the different countries were very similar on clinical measures and on indicators of childhood problems and family characteristics [26]. We have not identified any major differences in environmental factors or treatment practices across countries that were associated with forensic admission. Only minor differences were identified. For example, there was more use of illicit drugs in some countries than in others.

2.1. Design

Forensic hospitals were identified in four countries that were responsible for large catchment areas, in which almost all, if not all, persons with severe mental illness who were accused of a crime underwent a pre-trial psychiatric assessment, and if it was judged that they had committed a crime, they were sent to the forensic hospital. The sites were Southern British Columbia in Canada (here forth Sweden), Finland, the state of Hessen in Germany (here forth Germany), and Southern Sweden (here forth Sweden). In these sites, a person with severe mental illness is admitted to a forensic hospital after committing a serious violent crime and after his mental status has been assessed. Among subjects with mental illness, being admitted to a forensic psychiatric hospital is thus a proxy for having committed a violent offence.

In all study sites, there is a national, universally accessible, health system. Within each site, the forensic sample included consecutive discharges with a diagnosis of a major mental disorder. The control group included patients with severe mental illness being discharged from general psychiatric hospitals within the same geographic regions as the forensic hospitals. The patients were intensively assessed in the two weeks preceding discharge by experienced forensic psychiatrists who were trained to use standardized and validated diagnostic and symptom protocols and followed for two years, with interviews at six-month intervals.

2.2. Participants

Patients in the forensic hospitals with diagnoses of severe mental illness who were scheduled for discharge were invited to participate in the study. After complete description of the study to the patients, written informed consent was obtained. A diagnostic interview was then completed. If the diagnosis of severe mental illness was confirmed, the participant was included in the study and the other interviews and assessments were completed and information was collected from files. At study entry, participants were asked for permission to contact their mothers or an older sibling in order to collect information about them as children and about other family members. If the participant agreed, the collateral was invited to participate. The same procedure was followed with the patients recruited in the general hospitals.

Almost all (95.8%) the patients were male with schizophrenia diagnosis (68.2%). 186 (60.4%) of the patients were recruited from the forensic hospitals and 122 from general hospitals. Of the 308 patients, 78 were recruited in Finland, 31 in Sweden, 79 in Germany, and 120 in Canada, between 1998 and 2000 [26,27] (cf. Table 1). Within each country, the study was approved by the appropriate ethical committee.
Table 1
Variables. Status and age on first hospitalization, and principal diagnosis were not included, because they were established after the subject has turned the age of 18.

<table>
<thead>
<tr>
<th>Defined separately for</th>
<th>Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>Full time work, military service, full brothers, half brothers, full sisters, half sisters, number of siblings, mothers age at birth, fathers age at birth, behaviour problems in school, behaviour problems in home, behaviour problems in community, attention problems, hyperactivity problems, depression problems, substance abuse periods, duration of substance abuse, anxiety problems, other problems, problems generally, mental therapy before the age of 18, psychotherapy before the age of 18, psychotropics before the age of 18, counselling before the age of 18, care before the age of 18, institutional before the age of 18, performance in elementary school, performance in secondary school, change from elementary to secondary school, the highest level of education</td>
</tr>
<tr>
<td>Brother, sister, family</td>
<td>Incarcerated</td>
</tr>
<tr>
<td>Father, mother, biological mother, biological father</td>
<td>Occupation, education</td>
</tr>
<tr>
<td>Family, father, mother, biological mother, biological father</td>
<td>Subject living with</td>
</tr>
<tr>
<td>Father, mother, family, brother, sister, biological mother, biological father</td>
<td>Sanity, suicide attempts, convictions, convictions for violence</td>
</tr>
</tbody>
</table>

2.3. Analyses

In Bayesian analysis, the typical idea is to take advantage of existing information as a priori probability. Using Bayesian theorem, a priori can be combined with likelihood in order to produce a posterior probability. This study, however, assumed a different objective view for psychiatric domain – the priors were derived from actual observations in a model, not from beliefs or prior studies. We saw that the elements related to psychiatric patient’s background may contribute to subject’s tendency of forensic admission. This study contributed to understanding the aggregate outcome of societal, mental and somatic protective and risk factors: do forensic or general patients differ regarding to known background path?

2.3.1. Classification

Our aim was a classification task in which members in given datasets were classified into two groups based on the pre-admission properties. Here, the key task was the prediction of an event before its possible occurrence: the predicted event occurs in the future relative to the information available at the time of prediction and time is inherent to the concept of prognosis and distinguishes it from diagnosis. Prognosis can be defined as the prediction of the future course and outcome of disease processes, which may either concern their natural course or their outcome after treatment [28]. The main reason for doing a classification task – and not a regression task – was that the data set included hundreds of potential predictors and the outcome of interest and most of the predictors were dichotomous.

Firstly, we searched the country-specific factors among the factors presented in Table 1. This was done to find robust country-independent predictors and, by doing this, substitution method could be used in testing (i.e., independent test sets). In these non-informative country-specific models, no informative prior values were used. The prior probabilities for class variable (forensic or general) were set 0.5, prior evidence weights were set close to zero (0.001) and 1/k was used to calculate the marginal probabilities, where k is the number of potential values for the predictor variable.

Raw “as was observed” data was mainly used in these classifications. Interaction terms were included only for the most plausible interactions (e.g., biological father’s substance use and biological father’s convictions). We see that interactions can be very hard to understand from the clinical perspective, they cannot be directly observed, they are sometimes artificial and/or arbitrary, and the dataset included an extensive amount of predictors (i.e., deriving interactions would have been time costly and the potential gains would have been hypothetical). However, some composite variables (i.e., plausible variables merged from two or more variables) were also included.

For the very first time, the aim in this study was to find robust predictors and to generalize them to different country-specific sets. Country-independent predictions which would also be plausible were seen as more an important motive than mere predictive performance (i.e., the predictors can always get better and better, but generalized predictors and plausible explanations for them are more interesting in this domain). Thus, secondly, to produce generalized results, we used only the most influential predictors (most common factors with at least three hits out of four in our four country-specific models) found using the greedy hill-descending algorithm and assessed their impact using the regular NBC classifying.

Thirdly, based on the data triangulation of both the qualitative assessment of model robustness and the quantitative assessment of parameter candidates, we entered the Finnish data into this empty qualitative model with the found predictors and observed outcomes and probabilities to establish the likelihood for fusion (cf. RPV from Chapter 3.1 for more detailed description). We then utilized a type of Naive Bayesian Fusion (NBF) and combined priors and related weights from the German country specific model to the Finnish likelihood model in order to obtain a merger model.

In the fusion, we used the German prior probabilities (direct probabilities) and German prior weights (i.e., number of observations in the particular predictor conditional to class variable) obtained from the country specific model, and gathered the marginal probabilities for the fusion using the reversed version of German model (cf. Table 4 for details related to the
probabilities used). Unlike in previous studies, we also weighted the marginal probabilities with an evidence ratio (i.e., the relative accuracy between the German and Finnish country-specific models). Equal credibility (“degree of belief”, here one example in the German set matched one example in the Finnish set) for both datasets was assumed because (as will be seen) the number of observations in both sets was relatively similar; both seemed to follow the whole data’s profile and their RPV was quite similar. In simple terms, we used Finnish model as likelihood, German model as priors and merged them to a generalized model structure (i.e., a merger model was done by combining the Finnish and German data). Lastly, thorough testing was done.

The merger model was formed with a Naive Bayes classifier (NBC, P-Course [29–31], PREQ) algorithm. P-Course project was initiated by the need of valid and relatively easy-to-use classification tool for the medical domain. Method’s most important key features are the well-operating greedy hill-descending algorithm, graphical user interface, utility matrix for class variable, and its ability to utilize empirical prior information (i.e., reverse or direct probabilities and weights) in the model construction.

The algorithm has been found to be well-operating in the following tasks: variable selection (i.e., teaching is quite easy) is efficient in exploratory missions and, thus, the predictive performance tends to be good in exploratory analyses. In addition, NBC is known to perform well when variables are independent, to have a low level of bias, and to score well on measures of discrimination but not on calibration. Another possible methods for this type of prediction task include for example logistic regression (LR, typically with, e.g., stepwise variable selection or Bayesian model averaging, BMA [32,33]), and classification and regression tree (CART) analysis [34]. For information regarding P-Course, see sources utilizing this NBC [29–31,35–37].

In this study, the country-specific models built with P-Course without the priors were compared to confirmatory LR model’s accuracy without priors. This was done to ensure the predictive performance of NBC and robustness of findings also in this confirmatory setting, because NBC has presented its variable selection and prediction performance in exploratory missions [29,31]. Thus, LRs were thought with the same predictors. This is a fair game – when doing this, the result for LR is independent of the program used (i.e., the performance of search algorithm does not affect the results).

In addition, LR with commonly used stepwise variable selection method was used to reflect the variable selection performance of NBC and LR. Stata 10 IC statistical software 2 was used in LR modelling.

2.3.2. Measures

We assessed the relations between class variable and predictors with posterior odds (PO), which equal the product of the prior odds and the likelihood ratio. The POs were

\[
PO = \frac{P_{PC}}{P_{NPC}},
\]

in which \(P_{PC}\) presents the predicted class and \(P_{NPC}\) the non-predicted class. POs are not directly dependent on data size and give an idea of the predictor’s strength.

The inner validity (precision, also complement to the accuracy; error rate) of the model refers to the capability to predict the observations in the learning set precisely and the outer validity (reproducibility, “reliability”) refers to the model’s capability to predict observations in the test set or future dataset not included in the model learning. There is a trade-off between precision and reliability, and we wanted to search a model which has acceptable performance in both of them. We assessed these with the substitution method: naturally independent datasets were utilized for the learning and testing of model. We considered substitution with independent data set a more conservative and real life orientated method compared to simple portioning, in which the dataset is divided into independent learning and test sets that are eventually dependent.

The leave-one-out cross-validation (LOOCV) was also used. For a dataset with \(N\) examples, \(N\) experiments were performed and for each experiment, \(N - 1\) examples were used for learning and the remaining 1 example for testing (i.e., the test example is labelled by an NBC classifier trained on all other learning examples). The mean error \(\bar{E}\) in LOOCV was estimated as

\[
\bar{E} = \frac{1}{N} \sum_{i=1}^{N} E_i.
\]

The credibility intervals (CrI) for accuracies and POs were estimated using the Jeffreys interval – a Bayesian CrI based on Jeffreys prior (a variant of Beta 0.5, 0.5 distribution) [38]. Unlike conventional confidence intervals (CI), CrI directly indicates the limits for the probability of finding the mean within the given limits. In addition, Jeffreys interval does not violate the rules of probability (i.e., the lower level of CrI does not fall below 0 and the higher level of CrI does not exceed 1).

Analyzed complementary outcomes are also important in terms of specificity (\(spe\), predicted general examples of all observed forensic cases; true negative rate) and sensitivity (\(sen\), predicted forensic cases of all observed forensic cases; true positive rate). However, these do not validly represent models’ performance as single measurements. Thus, we utilized the diagnostic odds ratio (DOR) [39,40] estimated as

\[
DOR = \frac{\text{sen}/(1 - \text{sen})}{(1 - \text{spe})/\text{spe}}
\]

1 URL: http://p-course.hiit.fi/p-course/.

2 The performance of LR’s variable selection may be software-dependent. Thus, these results apply only for Stata 10.
to indicate the discriminative power as the ratio of the odds of a positive forensic prediction result among forensic patients to the odds of positive general prediction results among the general patient population. The 95% CIs for the DORs can be estimated conventionally from

\[
95\%CI = \ln \text{DOR} \pm 1.96 \cdot \text{se}(\ln \text{DOR})
\]

with the antiln back-transformation [40]. The standard error (se) for \( \ln \text{DOR} \) is

\[
\text{se}(\ln \text{DOR}) = \sqrt{\frac{1}{\text{true positives}} + \frac{1}{\text{true negatives}} + \frac{1}{\text{false positives}} + \frac{1}{\text{false negatives}}}
\]

Generally, AUC \( \hat{A} \) may be preferred to accuracy [41], if a classifier like NBC is able to produce probability estimations for class prediction cases. \( \hat{A} \) is equivalent to the probability that a randomly chosen negative example (general patient; i.e., non-forensic patient) will have less of an estimated probability of belonging to the positive class (forensic patient) than a randomly chosen positive case [42]. \( \hat{A} \)s can be estimated straightforwardly for learning sets.

However, “true” probabilities for examples in test sets are unknown. Thus, in \( \hat{A} \) estimation we took the highest class probability as the correct prediction, if it was the observed class [42,43] following a simple method:

\[
\hat{A} = \frac{S_0 - n_0(n_0 + 1)/2}{n_0n_1}
\]

\[
S_0 = \sum_{i=1}^{n} r_i,
\]

where \( n_0 \) is the number of positive cases and \( n_1 \) is the number of negative cases, and \( r_i \) is the rank of \( i \)th positive example in the ranked list. The standard error (se) [42] for AUC was estimated as

\[
\text{se}(\hat{A}) = \sqrt{\frac{\hat{\theta}(1 - \hat{\theta}) + (n_0 - 1)}{n_0n_1} + (n_1 - 1)Q_1 - \hat{\theta}^2}
\]

with

\[
\hat{\theta} = \frac{S_0}{n_0n_1} \quad \text{and} \quad Q_0 = 1/6(2n_0 + 2n_1 + 1)(n_0 + n_1 + 1)(n_0 + n_1) - Q_1,
\]

where

\[
Q_1 = \sum_{j=1}^{n_0} (r_j - 1)^2.
\]

In addition, the Stata 10 IC software was used to estimate accurate \( \hat{A} \)s for learning models.

In close relation to the \( \hat{A} \), Gini coefficient indicates the degree of concentration of a variable in a distribution of its elements (Gini compares the Lorenz curve of a ranked empirical distribution with the diagonal) [43]:

\[
\text{Gini} = 2\hat{A} - 1.
\]

When compared to simple accuracy, \( \hat{A} \) and Gini can produce different signals as they depend on the ranking of the observations.

Lastly, logarithmic loss function (log score) is a way of presenting losses from actual outcomes in prediction tasks [44]: the log score is a measure for prediction distribution (the closer to zero, the better). Log score is a strictly proper scoring rule, which reaches minimum only when true probabilities are predicted. Log score also penalizes the inability to produce faithful probabilities.

3. Results

60.4% (\( n = 186 \)) of the subjects were forensic patients (mean age 40.8 years; 96.8% males) and 39.6% (\( n = 122 \)) were general psychiatric patients (mean age 37.0 years; 94.3% males). The greedy NBC algorithm identified seven common factors from the four national datasets: Violent crime prior to the crime that leads to admission to the forensic hospital, biological father with substance misuse, biological father with criminal convictions, male gender, presence of half-brothers, no use of psychotropics before age 18, and sexual abuse before age 18.

Following variables were present in all explorations using the independent country-specific data sets: Violent crime prior to the crime that leads to admission to the forensic hospital, biological father with substance misuse, and biological father with criminal convictions. The presence of half-brothers did not have significant predictive value (measured as the logscore during hill-descending variable selection) in Canada, no use of psychotropics before age 18 did not have significant predictive value in Sweden, and sexual abuse before age 18 and gender (no females in the dataset) did not have significant predictive value in Germany. However, the forced inclusion of half-brothers, no use of psychotropics before age 18 (also a plausible variable), sexual abuse before the age of 18 and gender (also a plausible variable) to the model did not worsen the accuracy of country-specific models and, thus, they were included to the models.
3.1. Non-merger models

Country-specific information was compared with the obtained model structure. The Finnish dataset presented the highest inner validity (91.0%), while the German set displayed the lowest (84.2%), owing to the data and the chosen qualitative model structure (Table 2). The Swedish model was not significantly different from the Swedish default model. In the country-specific models, NBCs LOOCV and accuracy results were not significantly different and in the whole sample, LOOCV and the accuracy results were equivalent. This indicates that these predictive models are quite robust in terms of resampling.

Generally, when compared to NBC, confirmatory LR produced weaker scores for smaller datasets and little higher scores for larger datasets. This was a result which has been reported also by Ng and Jordan [45] and Blomstedt et al. [31] – LR may perform better than NBC in larger sets and worse than NBC in smaller sets. Also during the exploratory LR modelling, the datasets were found to be complex from the variable selection view. Stepwise LR did not perform particularly well in the variable selection, if the number of potential predictors exceeded approx. 60. In contrast, all feasible (approx. 280) variables were used during NBC explorations. With a lower number of potential predictors stepwise LR learned the models well and, thus, a number of different predictions had to be done. However, the best performing LR models found were over-fitted to the learning set: they did not make particularly good predictions in the independent test sets (Table 3).

Consequently, explorative LR with stepwise variable selection seemed to result to more precise estimations in the learning sets. Meanwhile, explorative LR did not perform particularly well in the test sets (i.e., a reliability test). In contrast, NBC's greedy algorithm resulted quite directly to a feasible amount of (robust) predictors and performed surprisingly well in the reliability tests.

Fig. 1 depicts the data triangulation of predictive performance (qualitative and quantitative robustness and efficiency of the model) using NBC models. In the figure, trend lines present the country-specific models; the horizontal axis gives the name of the utilized test set for the country-specific models, and the vertical axis gives the RPV (model's relative predictive value). In contrast, Fig. 2 presents RPVs for the LR models.

The general quality of the model (i.e., robustness based on the testing in different data sets) was depicted as the smoothness of the trend line and quantitative performance was depicted as the level of the trend. The odds for failure avoidance were

$$\text{RPV} = \frac{1 - \text{Accuracy}_0}{1 - \text{Accuracy}_1}$$

where RPV represent the odds for the model's success in terms of the comparison of predictive uncertainty related to compared models, Accuracy$_0$ presents the probability of the model and Accuracy$_1$ the related default probability (i.e., the % of

| Table 2 |
|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|
| Variables                | Distributions % (absolute frequencies) | Variables                | Distributions % (absolute frequencies) | Variables                | Distributions % (absolute frequencies) |
| Predicted class          | Canada (n = 120)           | Finland (n = 78)          | Germany (n = 79)            | Sweden (n = 31)           | Total (n = 308)                 |
| Forensic patient          | 50.0 (60)                  | 71.8 (56)                 | 63.3 (50)                  | 64.5 (20)                 | 60.4 (186)                     |
| Violent crime prior to the crime that leads to admission to the forensic hospital | | | | |
| Male                      | 51 (48)                    | 67 (31)                   | 61 (33)                    | 75 (21)                   | 60 (133)                       |
| Substance misuse by biological father | 99 (117)                   | 88 (69)                   | 100 (79)                   | 95 (30)                   | 96 (295)                       |
| Conviction(s) of biological father | 35 (35)                    | 27 (18)                   | 21 (15)                    | 21 (5)                    | 28 (73)                        |
| Half-brothers             | 7 (6)                      | 7 (4)                     | 7 (4)                      | 9 (2)                     | 7 (16)                         |
| Psychotropics before age of 18 | 66 (20)                    | 80 (9)                    | 70 (7)                     | 59 (4)                    | 69 (40)                        |
| Sexual abuse before age of 18 | 26 (28)                    | 25 (15)                   | 12 (9)                     | 29 (7)                    | 22 (59)                        |
| Accuracy (NBC)            | 84.2                       | 91.0                      | 86.1                       | 90.3                      | 86.4                           |
| 95% CrI$^{b}$ for the prediction score | 76.9, 89.9                  | 84.8, 96.7                | 77.2, 92.4                 | 76.4, 97.2                | 82.2, 89.9                     |
| LOOCV                     | 79.2                       | 89.7                      | 84.8                       | 87.1                      | 86.4                           |
| 95% CrI for LOOCV         | 71.3, 85.7                 | 81.6, 95.0                | 75.7, 91.4                 | 72.2, 95.5                | 82.2, 89.9                     |
| AUC                       | 84.2                       | 91.0                      | 83.5                       | 81.8                      | 84.3                           |
| Accuracy (LR), prediction score$^{c}$ | 41.1, 58.9                 | 61.2, 80.9                | 52.3, 73.3                 | 50.0, 79.5                | 54.9, 65.7                     |
| Jackknifing (LR)          | 85.0                       | 89.7                      | 84.8                       | 83.9                      | 87.0                           |
| 95% CrI for the default score | 0.413                      | 0.259                     | 0.303                      | 0.283                     | 0.383                          |
| Default log score         | 0.693                      | 0.595                     | 0.657                      | 0.650                     | 0.693                          |
| Accuracy (LR)             | 85.0                       | 88.5                      | 83.5                       | 80.6                      | 86.7                           |

$^{a}$ Variables were not present in some country-specific predictions and were thus forced to the model. This had no effect on the accuracy of models.

$^{b}$ All CrIs here are 95%CrIs from Jeffreys interval with Beta (0.5, 0.5) a priori distribution.

$^{c}$ Missing values were input for LR models (accuracy and Jackknifing) using NBC.
Table 3
The results of exploratory logistic regression analysis using independent nation-specific data sets.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Distributions</th>
<th>(absolute frequencies)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicted class</td>
<td>Canada (n = 120)</td>
<td>Finland (n = 78)</td>
</tr>
<tr>
<td>Forensic patient</td>
<td>50.0 (60)</td>
<td>71.8 (56)</td>
</tr>
<tr>
<td>Predictors in LR</td>
<td>Probability distributions for forensic patients %</td>
<td></td>
</tr>
<tr>
<td>Violent crime prior to the crime that leads to admission to the forensic hospital</td>
<td>51 (48)</td>
<td>67 (31)</td>
</tr>
<tr>
<td>Male</td>
<td>88 (69)</td>
<td>95 (30)</td>
</tr>
<tr>
<td>Substance misuse by biological father</td>
<td>27 (18)</td>
<td>21 (15)</td>
</tr>
<tr>
<td>Conviction(s) of biological father</td>
<td>7 (6)</td>
<td>7 (4)</td>
</tr>
<tr>
<td>Sexual abuse before age 18</td>
<td>26 (28)</td>
<td>25 (15)</td>
</tr>
<tr>
<td>Homicide status on first hospitalization</td>
<td>16 (13)</td>
<td>12 (9)</td>
</tr>
<tr>
<td>Subject native</td>
<td>58 (45)</td>
<td>55 (45)</td>
</tr>
<tr>
<td>Alcohol abuse/dependence</td>
<td>44 (53)</td>
<td>41 (53)</td>
</tr>
<tr>
<td>Drug abuse/dependence</td>
<td>11 (13)</td>
<td>17 (13)</td>
</tr>
<tr>
<td>Biological father born in site country</td>
<td>73 (88)</td>
<td>73 (88)</td>
</tr>
<tr>
<td>Substance misuse by biological father</td>
<td>11 (13)</td>
<td>17 (13)</td>
</tr>
<tr>
<td>Mental therapy before the age of 18</td>
<td>22 (68)</td>
<td>22 (68)</td>
</tr>
<tr>
<td>Accuracy (default is in parenthesis)</td>
<td>Canada</td>
<td>Finland</td>
</tr>
<tr>
<td>Prediction score in learning set</td>
<td>87.5 (50.0)</td>
<td>92.3 (71.8)</td>
</tr>
<tr>
<td>Prediction score using Canadian test set</td>
<td>na</td>
<td>76.9 (50.0)</td>
</tr>
<tr>
<td>Prediction score using Finnish test set</td>
<td>65.8 (71.8)</td>
<td>na</td>
</tr>
<tr>
<td>Prediction score using German test set</td>
<td>62.5 (63.3)</td>
<td>75.6 (63.3)</td>
</tr>
<tr>
<td>Prediction score using Swedish test set</td>
<td>58.3 (64.5)</td>
<td>66.7 (64.5)</td>
</tr>
<tr>
<td>Prediction score using all data</td>
<td>72.6 (60.4)</td>
<td>79.2 (60.4)</td>
</tr>
</tbody>
</table>

Fig. 1. NBC models’ robustness profiles and models’ relative predictive values (RPV). The used model is indicated with the legend markers, the input dataset is given in the horizontal axis and RPV is given in the vertical axis. The value of default RPV is 1.0 and misclassification is RPV times more likely, if the a priori best guess instead of the model is used in predictions.

Fig. 2. LR models’ robustness profiles and models’ relative predictive values (RPV). Explorative LR models seem to be more volatile than NBC models presented in the Fig. 1 and the average RPVs for LR models are lower compared to NBC (i.e., when using NBC, the odds for model success were higher).
observations in the biggest class). The interpretation is that a misclassification is RPV times more likely if the default is used instead of the model in the given prediction task.

As can be seen from Fig. 1 and based on the NBC modelling, the Finnish dataset seemed to present the average qualitative structure of the whole dataset, with high discriminative performance and robust quantitative probability structure as its trend line closely followed the trend of whole dataset and never dropped. The models with German and Swedish data did almost as well, while the Canadian model indicated volatility. However, explorative LR models were volatile in terms of RPV (Fig. 2). When observing the average RPVs of NBC models, they were higher than those observed during the LR modelling. Here, the odds for the model’s success were higher using NBC than using LR.

3.2. Merger model

The Finnish dataset \((n = 78)\) was utilized as the likelihood in NBF due to its robustness. The German dataset \((n = 79)\) was used to form informative priors, because the German model seemed to be robust and differed from the Finnish dataset in that it presented a higher degree of males while indicating lower level of violence, half-brothers, psychotropics, and sexual abuse. Bayesian inference was used to assess the evidence of predictors and related POs. POs were adjusted to equal levels a priori, i.e., the mentally ill subject has an a priori 50% chance (Beta 0.5, 0.5) to become a forensic patient.

Table 4 reveals quite strong POs for patients not being admitted to forensic care if they have a background free of violent crimes. Gender and sexual abuse as a child or adolescent provided suggestive POs, while no psychotropics before the age of 18, no half-brothers as well as no substance misuse by the biological father provided weak POs, and no criminal conviction of the biological father had no significant effect on POs for protecting against admission to a forensic hospital.

When evaluating the POs of becoming a forensic psychiatric patients, the most relevant risk factors for admission to forensic psychiatry included violent crimes, conviction(s) of the biological father and no psychotropics before age 18 with quite strong POs, substance misuse by the biological father with suggestive POs, and being male as well as no sexual abuse as a child or adolescent with weak POs.

Model’s performance was very good (Table 5). Considering the inner validity, accuracy is between 86.1% and 91.0%, and the LOOCV is 85.9%. The model produces high reproducibility (test score) between 82.5% and 87.1%. Model’s log score is low. The DOR is high for the Finnish data (102.0) and for the German data (93.1). The test DOR is good in the Canadian dataset (30.3) and in the Swedish dataset (80.0). Test Â is very good for Canadian (91.8) and Swedish (83.9) test data substitutions as well as for model learning (91.0; 83.5).

4. Discussion

This case study explored protective and risk factors for a forensic admission among the mentally ill in multinational setting, aiming at creating a robust merged model. This exploration was carried out in a Bayesian framework, enabling us to merge multinational reference datasets into a fusion model. We also introduced a new, simple and illustrative method (relative predictive value, RPV) that represents the odds for the model’s success in terms of the comparison of predictive

<table>
<thead>
<tr>
<th>Predictive factors</th>
<th>Predicted class</th>
<th>Evidence strength</th>
<th>Priors</th>
<th>Inversed</th>
<th>Priors</th>
<th>Inversed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>PO</td>
<td>95%CrI</td>
<td>%</td>
<td>PO</td>
<td>95%CrI</td>
</tr>
<tr>
<td>Violent crime prior to the crime that leads to admission to the forensic hospital</td>
<td>87</td>
<td>13.6</td>
<td>8.9</td>
<td>22.7</td>
<td>91</td>
<td>19.6</td>
</tr>
<tr>
<td>Female</td>
<td>60</td>
<td>3.0</td>
<td>2.2</td>
<td>4.2</td>
<td>10</td>
<td>2.1</td>
</tr>
<tr>
<td>No substance misuse by biological father</td>
<td>36</td>
<td>1.1</td>
<td>0.8</td>
<td>1.6</td>
<td>92</td>
<td>2.4</td>
</tr>
<tr>
<td>No conviction(s) of biological father</td>
<td>34</td>
<td>1.0</td>
<td>0.7</td>
<td>1.4</td>
<td>99</td>
<td>3.2</td>
</tr>
<tr>
<td>No half-brothers</td>
<td>37</td>
<td>1.2</td>
<td>0.9</td>
<td>1.6</td>
<td>94</td>
<td>3.2</td>
</tr>
<tr>
<td>Psychotropics &lt; age of 18</td>
<td>38</td>
<td>1.2</td>
<td>0.9</td>
<td>1.7</td>
<td>97</td>
<td>6.6</td>
</tr>
<tr>
<td>Sexual abuse as a child or adolescent</td>
<td>49</td>
<td>2.0</td>
<td>1.4</td>
<td>2.7</td>
<td>16</td>
<td>1.5</td>
</tr>
<tr>
<td>Risk factors</td>
<td>Forensic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>%</td>
<td>PO</td>
<td>95%CrI</td>
<td>%</td>
<td>PO</td>
<td>95%CrI</td>
</tr>
<tr>
<td>Violent crime prior to the crime that leads to admission to the forensic hospital</td>
<td>96</td>
<td>11.8</td>
<td>5.9</td>
<td>30.3</td>
<td>94</td>
<td>8.1</td>
</tr>
<tr>
<td>Male</td>
<td>69</td>
<td>1.1</td>
<td>0.8</td>
<td>1.5</td>
<td>97</td>
<td>1.7</td>
</tr>
<tr>
<td>Substance misuse by biological father</td>
<td>84</td>
<td>2.6</td>
<td>1.7</td>
<td>4.1</td>
<td>21</td>
<td>1.3</td>
</tr>
<tr>
<td>Conviction(s) of biological father</td>
<td>95</td>
<td>9.4</td>
<td>4.7</td>
<td>20.0</td>
<td>5</td>
<td>1.7</td>
</tr>
<tr>
<td>Half-brothers</td>
<td>88</td>
<td>3.6</td>
<td>2.3</td>
<td>5.9</td>
<td>23</td>
<td>1.5</td>
</tr>
<tr>
<td>No Psychotropics &lt; age of 18</td>
<td>94</td>
<td>7.7</td>
<td>4.3</td>
<td>16.5</td>
<td>24</td>
<td>1.5</td>
</tr>
<tr>
<td>No sexual abuse as a child or adolescent</td>
<td>69</td>
<td>1.1</td>
<td>0.8</td>
<td>1.5</td>
<td>92</td>
<td>1.4</td>
</tr>
</tbody>
</table>

All CrIs here are 95%CrIs from Jeffreys interval with Beta (0.5, 0.5) a priori distribution.
Also, we may state that if the setting would not have been controlled, model’s predictive performance might have been higher.

crime among patients with severe mental illness [46]. The present study failed to replicate these findings. Fullam et al. [54] summarized the earlier results of the violence research and suggested a relationship between psychopathy and violence in other hand, obvious and largely reported reasons including sex, age, and diagnosis were matched between the groups may have been weak signals behind factors with known relevance for admission to a forensic psychiatric hospital. On the other hand, sex, age, and diagnosis were matched between the groups and from this perspective we had the potential to explore new predictors for admission to a forensic psychiatric hospital. Also, we may state that if the setting would not have been controlled, model’s predictive performance might have been higher.

However, all the predictors in the model were observed before the first (index) admission to the hospital. Unlike in most previous studies in this domain, the Swedish and Canadian test sets did not originate from the same distribution as the learning sets – they were independent. All predictors included in the merger model were found to be good predictors and they are also plausible from the clinical point of view.

4.1. Clinical findings

The present study measured a wide range of social and community problems, for example, academic failure, non-completion of military service, poor work adjustment, and found them to be insignificant predictors of admission to a forensic psychiatric hospital due to the commission of a violent crime. By contrast, a record of conviction for a criminal offence in the past (before the crime that led to admission) was found to be the strongest predictor of admission. In a meta-analysis [46] and one study that compared different methods for predicting violent crime [47], the prediction methods ignored the greedy NBC. Despite this, both the meta-analysis and the study showed that criminal history variables such as prior criminal convictions, being the victim of violence, having a history of aggressive behaviour towards others, and having been sentenced to prison were strong predictors of violent crime among offenders with severe mental illness. The results of the present study concurred in showing that violent crime prior to the index hospitalization predicted admission to a forensic hospital.

The impact of the biological father’s antisocial behaviour measured as a conviction for a criminal offence and substance abuse was strongly associated with admission to a forensic hospital, as has been previously observed by Putkonen et al. [48,49]. This association may be explained in at least three ways. First, the impact of biological father’s antisocial behaviour may reflect an inter-generational transmission of a genetic vulnerability to antisocial and/or violent behaviour [50]. Second, it may result from inadequate and inappropriate parenting that characterizes parents who themselves present antisocial behaviours [51]. Third, it may result from an interaction between a specific genetic vulnerability and specific environmental factors, such as that observed between the presence of the low activity allele of the promoter region of the MAOA gene and physical maltreatment in childhood leading to conduct disorder in childhood and violent offending in adulthood among men [52]. The present study did not include genetic data for fathers which constituted the second major drawback. However, we controlled for societal reasons to specify possible genetic relations. If the data would have included genetic measures, model’s predictive performance may have been higher.

In an additional analyses controlling for gene-related factors such as foster care that could have resulted from the father’s antisocial behaviour, we found that the phenomenon appeared to be more biological than social as the presence of foster care had only an insignificant impact on the risk of admission to a forensic hospital. However, this does not rule out the influence of a dysfunctional family [51] resulting from the presence of a father who was actively engaging in antisocial behaviour. The observed impact on half-brothers may reflect such a situation. Evidence shows that parents who are themselves antisocial provide less than optimal parenting to their children [51] and that poor parenting is associated with the development of antisocial and aggressive behaviour in children [53]. To sum up, the strength of the association with the biological father’s antisocial behaviour was stronger than the associations with many other psychosocial factors.

A meta-analysis reported that poor work adjustment and a diagnosis of antisocial personality were predictors of violent crime among patients with severe mental illness [46]. The present study failed to replicate these findings. Fullam et al. [54] summarized the earlier results of the violence research and suggested a relationship between psychopathy and violence in

Table 5

<p>| Merger model’s (NBC) discriminative performance. |</p>
<table>
<thead>
<tr>
<th>Set</th>
<th>A %</th>
<th>95% CrI</th>
<th>Sen</th>
<th>Spe</th>
<th>AUC %</th>
<th>AUC se %</th>
<th>AUC %</th>
<th>AUC se %</th>
<th>Gini</th>
<th>DOR</th>
<th>95% CI</th>
<th>Log score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fin</td>
<td>91.0</td>
<td>83.2, 95.9</td>
<td>0.91</td>
<td>0.91</td>
<td>98.6</td>
<td>1.3</td>
<td>91.0</td>
<td>3.7</td>
<td>0.97</td>
<td>102.0</td>
<td>18.3, 569.3</td>
<td>0.30</td>
</tr>
<tr>
<td>Can test</td>
<td>82.5</td>
<td>75.0, 88.5</td>
<td>0.92</td>
<td>0.73</td>
<td>91.8</td>
<td>2.5</td>
<td>82.5</td>
<td>3.4</td>
<td>0.84</td>
<td>30.3</td>
<td>10.3, 89.0</td>
<td>0.48</td>
</tr>
<tr>
<td>Ger</td>
<td>86.1</td>
<td>77.2, 92.4</td>
<td>0.98</td>
<td>0.66</td>
<td>81.5</td>
<td>4.4</td>
<td>83.5</td>
<td>4.5</td>
<td>0.63</td>
<td>93.1</td>
<td>11.1, 777.8</td>
<td>0.31</td>
</tr>
<tr>
<td>Swe test</td>
<td>87.1</td>
<td>72.2, 95.5</td>
<td>0.98</td>
<td>0.67</td>
<td>83.9</td>
<td>6.6</td>
<td>81.8</td>
<td>7.6</td>
<td>0.68</td>
<td>80.0</td>
<td>3.8, 1682.4</td>
<td>0.65</td>
</tr>
<tr>
<td>All</td>
<td>86.0</td>
<td>81.8, 89.6</td>
<td>0.94</td>
<td>0.74</td>
<td>93.5</td>
<td>1.4</td>
<td>84.3</td>
<td>2.2</td>
<td>0.87</td>
<td>44.7</td>
<td>21.6, 92.9</td>
<td>0.41</td>
</tr>
</tbody>
</table>

a All CrIs here are 95%CrIs from Jeffreys interval with Beta (0.5, 0.5) a priori distribution.
b AUC for the unknown test probabilities (in Canadian and Swedish sets).

c AUC and related standard errors counted from predictions for the known probabilities (Finland and Germany).
d Gini coefficient.
e 95%CIs for the DOR.
f The LOOCV result was 85.9% (95Crl 79.9, 90.7).
patients with schizophrenia. Again, the present study failed to replicate these results. However, the present study showed that both female gender and childhood sexual abuse can be protective factors. The finding that female gender was a protective factor is not surprising given that males typically have higher rates of involvement in physical aggression and violence [55] and delinquent behaviours [56] than females. The impact of sexual abuse as a child or juvenile as protective factor was not strong and remains unexplained.

Patient’s substance abuse is a well-known risk factor for criminal and violent recidivism [48,57,58]. Among patients with severe mental illness, abuse of illicit drugs in the previous year has been found to increase the risk of violence by 2.5 times among the psychotic patients [47]. In the present study, both the patients admitted to the forensic hospitals and those admitted to the general psychiatric hospitals had high levels of substance abuse. Previous studies have shown, however, that after taking account of a childhood history of conduct problems, substance misuse no longer predicts aggressive behaviour towards others or violent crime [59–61]. This finding was supported by this study.

The present study, however, found that the use of psychotropic medication before the age of 18 was a protective factor. This may be a manifestation of the importance of early diagnosis and/or the result of effectiveness and compliance related to the actual pharmaceutical treatment. Other forms of treatment before the age of 18 had no robust impact regarding the admission to forensic care in this sample of patients with severe mental illness.

4.2. Modelling findings

Protective and risk factor exploration and predictions were carried out in a Bayesian framework, which enabled us to merge datasets in order to construct a robust model. Our study was data driven and we assumed an objective view. This was due to fact that we utilized matched case-control datasets, and knowledge related to the protective and risk factors for admissions to forensic psychiatry were limited and no logically coherent informative priors could be found in the literature. Thus, we assumed a pragmatic approach by merging the evidence from national datasets with Bayesian methods, established a merger model and did extensive external tests for the model.

Previously, the NBC method used in this study has outperformed or equalled LR [29,31,36] and BMA [31]. However, the predictive performance depends on various factors and the performance may vary from a prediction to another prediction. Thus, we did confirmatory LR analysis for the country-specific data and found that LR produced somewhat weaker scores for smaller datasets and little higher scores for larger datasets. Consequently, the present study concurred with previous results [31,36,45].

However, stepwise LR did not perform particularly well in the variable selection during the exploratory LR modelling, if the number of potential predictors exceeded approx. 60. This has been observed before in a context where classification trees and regressions were compared [62]. Here, in contrast to LR, all feasible (approx. 280) variables were used during NBC explorations. In addition, the LR models were over-fitted to the learning set; they did not make particularly good predictions in the independent test sets.

The predictive performance of NBC can be improved considerably with informative priors [29,31]. In these situations, NBC’s performance may even exceed the performance of LR and BMA [31]. The present study also showed that the model merging with independent a prior and likelihood data does not necessarily worsen the predictive performance of the model. On the other hand, the merging has the potential to result to generalized results and simple models.

The data triangulation process of predictive model is not a simple task, for example, multiple information sources must be merged through the fusion. Fortunately, DOR is a prevalence independent indicator which is able to combine the strengths of sensitivity and specificity as a single measure for discriminatory power. The DOR introduces some positive advantages over accuracy comparison and, e.g., tolerates the two-way assessments of confidence intervals with relatively straightforward interpretation. The DOR is helpful when the balance between false positive and false negative rate is not highly important and when multiple models are being compared.

AUC analyses may be preferred to accuracy as a measure of predictive performance as this procedure can provide information which is fairly independent of the base rates of violence in different populations. Our model was able to produce very good AUC estimates when compared to previous studies of violence predictions. The high AUC levels may be a manifestation of approach’s performance in AUC metrics, NB’s efficiency in utilizing all data and justified simplifications made for the AUC estimations.

Our study indicated that the measures given by AUC or Gini, accuracy, DOR, and log score are different due to their different perspectives for the classification and, thus, they should be used together. AUC can be extremely suitable for substitution situations where the order of observations is relevant for comparison (i.e., the situation is related to learning). In this study, the DORS and AUCs as well as Gini coefficients indicated that the actual discriminative classification modelling between general and forensic patients was robust and concentrated in all datasets.

The DOR and AUC do not enable straightforward interpretation for the qualitative properties of the model. Thus, we introduced a new, simple and illustrative method for the data triangulation of qualitative and quantitative assessment through the relative predictive value (RPV). The qualitative and quantitative observations can be simultaneously done in terms of failure avoidance when using the informative model instead of default. The most important benefits of RPV include both statistical and clinical ones: the possibility to do understandable qualitative and quantitative assessment for the modelling, and a measure which is also feasible for the medical domain in terms of model comparison (i.e., clinicians are typically hesitant to take risk and they want to avoid the risk of misprognosis). RPVs also indicated that NBC models were here quite robust in
comparison to the explorative LR models. When observing the average RPVs of NBC models, they were higher than those observed during the LR modelling: the odds for the model’s success were higher using NBC than using LR.

We see logical rationality in the means of using the data or other sources of information which follows the aggregate trend of multiple datasets. This most evidential likelihood (i.e., data or other information set with the highest predictive value among all options) has the potential to give a robust quantitative probability structure for the explored qualitative classification model structure. If model averaging or merging is of interest, datasets must be logically consistent and they must be weighted for their relative plausibility derived from their accuracy.

The third and fourth major drawbacks of the analysis were also related to the case-control setting. The study lacked the size of population at risk for the studied groups and was unable to determine the incidence of violence for the predictions. Thus, we used a uniform a priori distribution (Jeffreys Beta 0.5, 0.5) for POs. In addition, total control over matching was not achieved in the data: age and gender distributions were little different in the datasets and some control subjects were missing. However, we were able to adjust the model structure for gender differences and the minor age structure difference seemed to be irrelevant as predictor. To sum up, the main limitations likely had only an insignificant impact on the predictions as long as it is remembered that these results apply for Canadian, Finnish, German and Swedish settings, and the data were matched for sex, age, and diagnosis.

Generally, as observations, parameters, and models are uncertain, there exist several ways to explain data with the parameters and models. Of all the plausible explanations, the simplest can be considered best, yielding the best predictions (this is the well-known Occam’s Razor principle). Philosophically speaking, our NB fusion approach leaned against the Razor: factors not common among national datasets may be a result of white noise or due to controversies in datasets, insignificant cultural reasons or country-specific factors. Here the quite complex modelling approach was beneficial: we did not lose in terms of log score and gained some in terms of accuracy, RPV and AUC.

5. Conclusions

This study introduced a number of applications which have not been used in the previous studies in the psychiatric domain. The exploration was carried out in a Bayesian framework, which enabled us to merge datasets for a robust fusion model. For the data triangulation, we introduced a new, simple and illustrative method called relative predictive value. The established merger model indicated both high predictive capabilities and robustness. The approach produced generalized predictors of admission to forensic psychiatric hospitals, notably antisocial behaviour in the biological father.

Author contributions

Equal contribution of both first authors (ESO and TRI). ESO contributed to the concept, modelling, statistics, design, management, and drafting; and TRI contributed to the concept, design, management, and revision. JTI, SHO, MER; and ORY contributed to design, management, and revision.

Role of funding source

Funding was provided by the Finnish Funding Agency for Technology and Innovation, and ESiOR Oy, Kuopio, Finland which had no direct involvement with the article.

Acknowledgement

ESO wishes to thank Jussi Lahtinen (CoSCo, HIIT, Finland) for co-operation with the P-Course software development.

References


Erkki Soini is the founder member, CEO and Vice President at the ESiOR Oy, part-time lecturer at the University of Kuopio and a founder member of DIADEM research cluster. He has authored 12 papers in peer-reviewed journals, given 26 congress presentations, authored three books, and won a best student poster award of methods and concepts: the study was related to Bayesian analysis in proactive modelling. The main topics in his scientific work in medicine and biomedicine are econometrics, health and pharmacoeconomics including, e.g., cost, quality of life, cost-effectiveness, cost-utility, net benefit, budget impact, data-mining and modelling studies.

Tarja Rissanen is Unit Manager at the Niuvanniemi Hospital and doctoral student at the Department of Health Policy and Management, University of Kuopio. She has authored three papers in peer-reviewed journals and six presentation abstracts about mental health and abuse services. The dissertation will deal with the care of treatment resistant minors.

Jari Tiitonen is Professor and Chairman at the Department of Forensic Psychiatry, University of Kuopio. He is also Research Professor at the National Public Health Institute. He has authored 202 papers in peer-reviewed journals. The main topics in his scientific work in psychiatry are psychopharmacology, functional brain imaging, autoradiography studies on dopamine and serotonin receptors and transporters in the human brain, and epidemiology of mental disorders and crime.
Sheilagh Hodgins is Professor at the Institute of Psychiatry King's College London. She conducts research on the causes of antisocial and violent behaviour and effective strategies for reducing these behaviours. Data for the current analyses were based on the study she lead that compared outcomes among matched samples of forensic and general psychiatric treatment recruited in Canada, Finland, Germany, and Sweden.

Markku Eronen is the Medical Director of Vanha Vaasa Hospital, and a lecturer at University of Turku. His research interests include biological and epidemiological studies on mental disorders, aggression and violent behavior. He has about 40 scientific publications on these subjects. He is a member of the board of Finnish Psychiatric Association.

Olli-Pekka Ryynänen is Professor at the Department of Public Health and General Practice at the University of Kuopio, Finland.