Effects of the Presser and legislative criteria on classifying New Zealand's fitness to stand trial court outcomes

Khan V. Buchwald-Mackintosh^a, Mei Wah M. Williams^a and Joseph A. E. Sakdalan^b

^a*School of Psychology, Massey University, Level 3, North Shore Library Building, 229 Dairy Flat Highway 17, Albany village, North Shore, Auckland, New Zealand. Private Bag 102 904, North Shore Mail Center, Auckland 0745, New Zealand;* ^b*Regional Forensic Psychiatric Service, Avondale, Auckland, New Zealand*

New Zealand's legislation mandates the inclusion of several legislative criteria to assist in the determination of defendants' fitness to stand trial (FTST). However, the unlegislated Presser criteria have been described as useful to discriminate between defendants' fitness without formal research to identify them as such. The sample consisted of 252 defendants' FTST forensic assessments between 2005 and 2015 with a mean age of 30.1 years, where 87.7% were male. Defendants were primarily from Māori descent (37.3%), New Zealand European (34.1%) and Pacific Island descent (17.9%). The authors found that all Presser and legislative criteria significantly predicted whether the defendant was fit to stand trial (FST). Both the Presser and legislative criteria have excellent area under the curve (AUC) statistics. However, the Presser criteria, compared to the legislative criteria, improved specificity of court judgements. The Presser criteria may therefore assist in correct FTST court decisions, particularly for borderline-fit defendants.

Keywords: competence to stand trial; fitness to stand trial; forensic psychology; health assessors; legislative criteria; New Zealand legislation; Presser criteria; Pritchard criteria

Fitness to stand trial (FTST) is an inexact science. In New Zealand, a defendant whose capacity to function adequately in court is questioned is often referred to health assessors (psychologists, psychiatrists and/or specialist assessors) to undergo a formal assessment to determine whether they are fit to stand trial (FST). New Zealand health assessors are guided by the Criminal Procedure (Mentally Impaired Persons) Act (CP[MIP] Act) [Criminal Procedure (Mentally Impaired Persons) Act & New Zealand Statutes, 2003] to assess and report on FTST. FTST assessments are needed to identify whether a defendant is unable to participate in court proceedings or conduct a defence due to mental impairment. A mental impairment refers to

a defendant who has an incapacity (due to neurological deficits, mental illness or an intellectual disability) to participate in court (Visser, 2011). In New Zealand, diagnosis of a psychological or neurological disorder is not sufficient to rule that a defendant is unfit to stand trial (UFST). The CP[MIP] Act (2003) replaced Part 7 of the Criminal Justice Act (Criminal Justice Act & New Zealand Statutes, 1985) in recognition that dementia and acquired brain injury may also make it difficult for defendants to participate fully in court (Visser, 2011). These changes appear to reflect global recognition that various pathology and/or cognitive impairments may limit a defendant's ability to participate adequately in judicial settings. Ensuring defendants have

a fair trial and maintaining the integrity of the criminal justice system are central to the concept of FTST (Mossman et al., 2007).

Various terms have been adopted across countries for FTST capacities or judicial procedures (i.e. adjudicative competency in the United States), but these terms are generally based on the concept of preserving justice for defendants in the criminal justice system. This article refers to FTST for defendants who proceed through this alternative criminal justice process or have questionable ability to participate in court. This article also refers to health assessors as FTST assessors, regardless of the discipline of the professional or jurisdiction. Within a jurisdiction there are often criteria that set minimum standards for a ruling of FST. There are notable contrasts in the selection or wording of psycho-legal criteria mandated among different jurisdictions, but there are many fundamental similarities in the FTST criteria between jurisdictions abroad.

New Zealand's *legislative criteria* (see Appendix A) for FTST have their foundations in the Pritchard criteria, based on the English decision of *R v. Pritchard* (1836). The CP[MIP] Act (2003) does not limit additional criteria to be considered in a FTST hearing, but they are not statutory requirements to include to consider a defendant FST. New Zealand's legislative criteria also overlap with the Presser criteria (see Appendix B); however, there are some Presser criteria that are not considered within the scope of New Zealand's legislative criteria. States in Australia generally have a mix of the Presser criteria and the Pritchard criteria (*P v. Police*, 2006; *R v. Presser*, 1958). The Presser criteria were established in the Australian court decision of *R v. Presser* to help ascertain whether the defendant was FST. In this landmark case, the defendant had cognitive and intellectual impairments that were not considered within the scope of Australia's previous legislation (Kasinathan, Le, Barker, & Sharp, 2016). The United States mandates the Dusky Standard or the requirement for defendants to possess

factual and rational understanding, and the ability to consult with his or her lawyer. Ability to plead is the doctrine considered by England and Wales. Ability to plead best represents the Pritchard criteria but overall narrowly targets the ability of entering a plea (Rogers, Blackwood, Farnham, Pickup, & Watts, 2008).

Overseas evidence highlights some issues with assessing and reporting on FTST. Nicholson and Norwood (2000) found that health assessors do not always consider all relevant psycho-legal FTST criteria in their reports. These authors noted that in Utah 73% of assessors reported on the defendant's understanding of the legal process, whereas 85% of assessors reported on whether the defendant appreciated the charges against him or her. Not assessing or reporting some criteria threatens reliability when there is variation in the training received to carry out assessments and reports (Gowensmith, Murrie, & Boccaccini, 2012). Further, Large, Nielssen, and Elliott's (2009) research on inter-rater reliability showed there was only moderate agreement among health assessors' opinions of FTST. Given this, it is not uncommon for independent health assessors to assess, report and reach different conclusions for the same defendant.

In New Zealand, no structured FTST assessment or practice guideline has been developed to improve inter-rater reliability or objectivity in assessments, and no research has been published that identifies the extent of this problem. Improvements in the validity and reliability of FTST opinions have been found when using structured assessments such as the MacArthur Risk Competence Assessment Tool–Criminal Adjudication (Hoge, Bonnie, Poythress, Monahan, Eisenberg, & Feucht-Haviar, 1997). Structured assessments developed abroad would probably not have sufficient validity to be used within the New Zealand context due to the differences within legislation. The overall subjectivity involved in FTST assessment led Morris, Haroun, and Naimark (2004) to liken FTST decisions to coin flips in the courtroom.

In addition to this, New Zealand's legislative criteria, and the CP[MIP] Act (2003) in general, may be open to interpretation or have significant ambiguity (Brookbanks, 2014). Felthous (2011) also suggested that there are some difficulties with interpreting the Dusky Standard for FTST defendants in the United States. Any difficulties in interpretation may make it difficult to translate legislation into assessment practice. An assessment needs to be derived from legislation or case law for the courts to make a fair judgement of FTST. Abstract, vague or ambiguous wording in FTST legislation that do not specifically spell out the cognitive, functional and/or behavioural standards may lead to subjectivity in applying FTST law in psychology. Therefore, minimising the risk of subjectivity, or enhancing the ability to make a true finding of FTST, is needed to preserve justice for FTST defendants.

Compounding the problem of subjectivity for the courts is the notion that it may be difficult to discern whether a defendant is UFST or FST if they are borderline fit. In the trial of *P v. Police* (2006), the defendant could not be satisfactorily deemed FST or UFST with the current statutory legislative criteria because the defendant was on the cusp of fitness. As a result, the Presser criteria were used in conjunction with the legislative criteria to make a ruling for this defendant. The Presser criteria may set out what is expected for defendants to be found FST more concretely than New Zealand's legislative criteria. Including the Presser criteria and the legislative criteria may be an alternative to help preserve the integrity of New Zealand's court system.

The Presser criteria not captured by the CP[MIP] Act (2003) appear to target the extent of the defendants' knowledge around their rights and role in the court proceedings. Given greater concreteness, the Presser criteria may also be more readily translated into assessment practice and, therefore, could be considered standards that help ascertain defendants' fitness. Further, by assessing defendants using criteria set out in the CP[MIP] Act (2003) and

the extensions made by the Presser criteria, it may ensure that defendants deemed FST are able to participate meaningfully in court (Birgden & Thomson, 1999). By using the Presser criteria in court hearings, defendants would be trialled against minimum standards with less unfairness or prejudiced court procedures (*R v. Presser*, 1958; Samuels, O'Driscoll, & Allnutt, 2007). The Presser criteria's utility to be more discriminatory in ascertaining fitness and allowing defendants to have a fair trial does not have supporting empirical evidence (*P v. Police*, 2006).

New Zealand tends to have a low threshold for FTST leading to an over-estimation of defendants who are found FST rather than UFST (Samuels et al., 2007). This cautiousness may violate the doctrine of a fair trial process when UFST defendants who are deemed FST proceed through an unfair criminal trial. Torrey, Kennard, Eslinger, Lamb, and Pavle (2010) found that those who were assessed to have serious mental illness were three times as likely to be sentenced to imprisonment than to psychiatric hospitals. Parton, Day, and White (2004) also found that defendants who met the criteria for a diagnosis of intellectual disability frequently received prison sentences after a criminal court hearing, despite having difficulty understanding the trial process. A sentence of imprisonment for an UFST defendant considered FST by the court would be considered a miscarriage of justice. The most pertinent factor that rests within a FTST hearing is arguably whether the determination of a defendant as UFST or FST is correct. FTST hearings may become more ethical when criteria are included to reduce Type 1 and Type 2 error in court judgements.

Conversely, if the Presser criteria are not applicable to the outcome of FTST court decisions, then it may be unethical and an invasion of privacy to disclose unnecessarily information about defendants' capacities (Cram & Dobson, 1993; Groth-Marnat, 2009). Potentially private information may be exposed in court documents that could harm the defendant. Hence, assessors would need to

demonstrate caution and be able to demonstrate empirical support to assess the Presser criteria in exceptional cases when in general they do not assist in FTST findings. If the Presser criteria do not assist in determining court outcomes, then there may be a considerable reduction in time, resources and public spending on health assessor's assessments and reports. However, no empirical evidence to date in New Zealand supports whether defendants would benefit from being assessed against the Presser criteria standards in addition to the legislative criteria.

The current research attempts to provide empirical support to identify whether the inclusion of the Presser criteria improves health assessors' assessments in New Zealand. Forensic assessments highly influence the outcome in court (Skeem, Golding, Cohn, & Berge, 1998), yet research in New Zealand has not considered the extent to which the Presser criteria predict FTST court outcomes, despite common use of the Presser criteria by health assessors in New Zealand. More evidence may improve assessment and reporting procedures that would improve the court's ability to discriminate between FST and UFST defendants. Research to support scientific practice is needed for the courts to move beyond decisions that some researchers call a 'matter of chance' (Morris et al., 2004).

This research explores the relevance of the Presser criteria on court outcomes for defendants to have a fair trial, and to potentially reduce Type 1 and Type 2 error. The research potentially identifies whether unfit mentally impaired persons are deemed FST when they are not including the Presser criteria and aims to identify whether improvements in decision-making can be gained by comparing or combining the Presser criteria with the legislative criteria to predict court outcomes.

Method

Sample

Two hundred and fifty-two FTST court reports within the Auckland and Northland regions

between the years 2005 and 2015 comprised the data of this sample. The mean age for the defendants was 30.1 years, and the median age was 28 years with a standard deviation of 10.5 years. The majority of defendants in this sample were male (87.7%), whereas females accounted for a small proportion of the sample (12.3%). The sample contained a range of ethnicities including people of Māori descent (37.3%), New Zealand European (34.1%), Pacific Island descent (17.9%), Asian descent (3.6%) and other ethnicities (6.7%). Ethnicity was not recorded for less than 1% of the sample. A total of 95.6% of defendants did not have a tertiary qualification. Most defendants were unemployed (90.9%). Intellectual disability was diagnosed in 48.4% of defendants, 42.1% were diagnosed with mental illness, 7.1% had both a comorbid intellectual disability and mental health diagnosis, and 2.4% had neither diagnosis.

All participants had at least one FTST assessment by a health assessor, and 65.9% had a second assessment. Psychologists assessed and wrote reports 21.1% of the time whereas psychiatrists assessed and reported on the remainder of evaluations (78.9%). One hundred and thirty-seven (54.4%) defendants were declared UFST while 105 were found FST (41.7%), and 10 (4.0%) defendants had no recorded court outcome. The court outcome for defendants includes defendants where 55 were discharged, 52 defendants were discharged to an intellectual disability service, 33 were discharged to prison, 28 were discharged to a mental health service, 24 were withdrawn by leave, 20 were discharged into the community, 15 had no need for further involvement, five were found not guilty by reason of insanity, three received a hybrid penal forensic mental health sentence and one defendant was acquitted of the charge. Sixteen defendants had no recorded court outcome.

Data and sampling methodologies

The data were collected through archival data sources (court reports) written by health

assessors who were ordered by the court to assess and report on defendants according to Section 38 of the CP[MIP] Act (2003). These reports were available for research through the regional forensic psychiatric service in New Zealand. During the data collection, 203 variables were included in the database, and most variables were dichotomous (FST, UFST, etc.). The list of defendants from medical records was made available to staff at Waitemata District Health Board, and the medical records were sampled using two sampling methods.

The sampling method to obtain defendants' files from January 2005 to December 2012 was quota sampling, with a sample size of 190 defendants. These data were used for the study titled 'A Comparative Study on the Different Factors Associated with Fitness to Stand Trial between Mentally Disordered and Intellectually Disabled Defendants within the New Zealand Criminal Justice System' (Sakdalan & Egan, 2014). The remainder of the sample applied convenience sampling from January 2015 to February 2015. The former sample matched equal numbers of intellectual disability defendants with mental health defendants, while the latter sample had inclusion criteria and required that both a psychologist's and a psychiatrist's report existed for each of the 62 defendants. Different sampling strategies allowed separate research to be conducted on this database.

Of the 203 variables in the original database, seven variables were retained for these analyses. All variables in this study were dichotomous. To ensure row independence, only the latest FTST hearing's reports were retained when a defendant had a FTST hearing over more than one occasion. The first assessment for a particular court case was the only assessment included to maintain row independence. Missing values were imputed with the *k* nearest neighbours algorithm, and the average of five nearest data points was used for each missing value throughout (Little & Rubin, 2002). The researchers created random

variables for each variable with missing data. The variable with missing data was ordered according to the size of the values of their allocated random variable, before *k* nearest neighbours algorithm was ran. The researchers iterated this process until no missing values remained. This ensured that the original order or a single new order did not bias the data. The data were then imput. This study used an *a priori* critical value of $p = .05$ throughout.

The Health and Disability Ethic Committee (HDEC) in New Zealand gave ethical approval for a previous study to access the court reports for research purposes and a notice that the current study did not require a complete application. The research was deemed low risk by Massey University Human Ethics Committee. No contact with participants was required for this study. No personally identifiable information was available on this database. Only descriptive demographic information is included in the results.

Results

A chi-square analysis was applied to test the hypothesis that the outcome in court (FST, UFST) is dependent on each of the Presser and legislative criteria in univariate analyses. Odds ratios were included to demonstrate the direction of the effect and allow for statistical comparisons between variables. Assessment of defendants as meeting a criterion, or not meeting a criterion, predicts an increase in odds for FST or UFST, respectively. Either analysis results in the same statistic. The intercepts for all logistic analyses in this study were omitted. Court outcome was dependent on all univariate psycho-legal criteria examined. Therefore, the outcome in court significantly depends on the results of the assessment of each criterion across the sample. The legislative criteria (indicated by the first, second and third variables in Table 1) were highly significant; however, as they had higher larger sample size, this may have influenced the power of the analyses. Furthermore, the odds ratios for all

Table 1. Test of independence between court outcome and univariate psycho-legal criteria.

Variable	χ^2	<i>df</i>	<i>p</i>	OR	95% CI for OR
Ability to plead	88.02	1	<.001	19.50	[9.78, 38.89]
Ability to instruct counsel	68.85	1	<.001	13.75	[7.00, 27.00]
Ability to understand the nature, purpose and consequences	65.69	1	<.001	11.59	[6.16, 21.79]
Ability to exercise rights	23.90	1	<.001	22.50	[5.58, 90.66]
Understand charges	49.50	1	<.001	18.11	[6.84, 47.96]
Understand the substantial impact of evidence	23.19	1	<.001	11.36	[3.95, 32.67]
Understand the role of participants	41.23	1	<.001	7.13	[3.81, 13.36]

Note: OR = odds ratio, CI = confidence interval.

the assessed criteria were significantly different from one, indicating they significantly increase the odds for a finding of FST. Ability to plead increased the odds of the outcome in court significantly more than understanding the role of participants. There was more variation in most odds ratio confidence intervals for the Presser criteria extensions (the fourth, fifth, sixth and seventh variables in Table 1). This may be due to the smaller sample size of this group.

Table 2 displays the logistic models to test the hypothesis that the Presser criteria, when added to the legislative criteria model, do not improve on the model fit of the outcome in court. The legislative count variable and Presser count variable are variables that sum all legislative criteria and the overall criteria, respectively, that a defendant meets. The information added by including the Presser criteria did not significantly explain more variance for the increase in the degrees of freedom. Therefore, adding the Presser criteria to a model of the legislative criteria, $\chi^2(1) = 1.18$,

$p = .278$, did not significantly improve model fit for predicting the outcome in court. This may have been due to a sample size reduction, and, therefore, there may be a reduction in the power of the analysis. The Presser criteria had less residual deviance than the legislative criteria, which indicates that they have a better logistic model fit.

As in Table 3, the difference in the classification rate by using different groups of criteria is further emphasised using an area under the receiver operating characteristic curve (AUC) statistic to identify the ability for these criteria to classify the outcome in court. Table 3 suggests that both sets of criteria significantly modelled the specificity and sensitivity of the outcome in court better than chance. Using the Presser criteria achieved the largest AUC statistic. This may have been due to lower specificity made by the legislative criteria. The two sets of criteria had unequal sample sizes; therefore no inferences are made about significant differences even though the legislative criteria's AUC confidence interval did not

Table 2. Likelihood ratio test of legislative and Presser criteria predicting court outcome.

Model	Residual deviance	χ^2	<i>df</i>	<i>p</i>
1. Legislative criteria	25.80			
2. Presser criteria	24.62	1.18	1	.278

Note: The likelihood ratio test was conducted with a sample size of 35 participants.

Table 3. AUC statistics of grouped criteria predicting court outcome.

Model	N	AUC	SE	p	95% CI for AUC
Legislative criteria	192	.87	0.028	<.001	[.81, .92]
Presser criteria	38	.91	0.048	<.001	[.81, 1]

Note: N = sample size; AUC = area under the receiver operating characteristic curve; CI = confidence interval.

overlap with the AUC value for the Presser criteria. There was no a priori hypothesis to test the difference between the two AUC statistics using a statistical test.

Sensitivity and specificity analyses

The ability for both criteria to predict and classify court outcome is given below in a specificity and sensitivity analysis. The total sample size varied across criteria groups for the specificity and sensitivity analysis. Sensitivity in these analyses refers to defendants who are correctly classified by the model as FST when found FST by court as opposed to Type 2 error or a false negative (found UFST by the model when FST in court). Specificity indicates the model correctly classified defendants as UFST when defendants were ruled UFST by court as opposed to Type 1 error or a false positive (incorrectly classified as FST by the model when found UFST in court). The Presser criteria came to a sensitivity of .75 or 75.0% and specificity of .962 or 96.2% in correctly classifying the outcome in court. Sensitivity and specificity, when using the legislative criteria to predict court outcome, was .767 or 76.7% and .868 or 86.8%, respectively. This indicates that the Presser criteria have better specificity, but slightly poorer sensitivity in classifying court judgements.

Discussion

This research identified that all psycho-legal criteria surveyed significantly predicted the outcome in court. This study did not support the hypothesis that the Presser criteria significantly improve the model fit when added to the legislative criteria model. However, there

was an improved classification rate when using the Presser criteria compared to the legislative criteria.

The Presser criteria are important to aid in the discrimination of a defendant's FTST (*P v. Police*, 2006), and the current study supported this notion with an improved specificity for classifying court outcomes. Fewer defendants would be found FST when actually UFST when using the Presser criteria than when using the legislative criteria. In comparison to the Presser criteria, the legislative criteria classify fewer defendants as UFST when truly FST in court. The improvement of specificity may be most valuable in court, as incorrectly classifying defendants as FST when they are truly UFST may have more adverse implications for defendants than incorrectly classifying defendants as UFST when truly FST.

A preventable false-positive rate may have dramatic implications for defendants when health assessors neglect criteria that have a basis on fitness. Type 1 error has severe implications given the resulting outcome for the defendant. If defendants proceed to a criminal trial despite having mental impairment that impairs his or her ability to function in court, this could be considered a miscarriage of justice. Olley, Nicholls, and Brink (2009) note that correct determinations are critical given that a disproportionate number of prison inmates are diagnosed with a mental illness. Conversely, by using the extensions of the Presser criteria, a small but negligible number of defendants who are FST would be regarded as UFST by the court. This may incur minor cost increases in resources and public spending although not all defendants proceed to a secure psychiatric facility or residential treatment.

In this study, the Presser and legislative criteria were both excellent models of the outcome in court, in terms of their AUC statistics. This suggests that the two models are similar in their ability to classify court outcome. However, there was both overlap and a significant difference between the Presser criteria and the legislative criteria in the confidence intervals for their respective AUC statistics. There is good reason to suggest that this is due to differing sample sizes as health assessors assessed the Presser criteria less often in this study than the legislative criteria. Despite overlapping confidence intervals, sensitivity or specificity differences between the groups of criteria may still be significant, but inferences cannot be made regarding these differences with differing sample sizes.

This study found that the Presser criteria have a better fit in a logistic model than the legislative criteria but that the Presser criteria do not significantly improve on the model fit when added to a logistic model of the legislative criteria. By including the extensions of the Presser criteria, not enough additional variation explained court outcomes to be significant. Therefore, this result is inconclusive and requires a further follow-up study. This is the first known research to target this aspect of FTST and appears to be an important element to examine, especially in the New Zealand context.

All the psycho-legal criteria individually predicted the outcome in court in this study. Visser (2011) indicates that assessing the legislative criteria are not prescriptive and that assessing the Presser criteria is not mandatory. Using the Presser criteria as evidence for FTST, as described by Samuels et al. (2007), reduces the potential for prejudiced court proceedings, and therefore an unfair trial. This study highlights the need for health assessors to assess the Presser criteria to avoid false-positive court judgements. The Presser criteria have utility in the discrimination of FTST, which is particularly useful for defendants who are difficult to discern as UFST or FST.

Marcus, Poythress, Edens, and Lilienfeld (2010) assert that meeting FTST criteria, such as instructing counsel, is often viewed as taxonomic rather than resting on a continuum. Furthermore, borderline-fit defendants may be particularly difficult to assess with New Zealand's current legislation. An improvement in specificity and sensitivity may have more utility for people who are borderline fit than for those defendants who can be easily distinguished as UFST or FST. The problem of classifying borderline defendants can be exacerbated when it is difficult to translate legislation into assessment practice, especially since there is questionable objectivity in FTST assessments (Large et al., 2009; Morris et al., 2004). Using the Presser criteria may be valuable for borderline-fit individuals as suggested by Judge Baragwanath in the *P v. Police* (2006) trial. Using the Presser criteria does not appear to eliminate the potential for misclassification, but it does appear to reduce the frequency of it.

The New Zealand legislation does provide a legislative requirement for courts to assist in decisions for borderline-fit defendants. Judges must find, on a balance of probabilities, that a defendant is UFST (Visser, 2011). This suggests that a moderate degree of evidence is required. However, judicial officers prefer a determinate opinion on FTST or an opinion on the ultimate issue (Large et al., 2009), which contradicts applying the continuum approach. Since health assessors are to provide the most accurate prediction of a defendant's fitness, and since judges generally defer to the opinion of experts (Zapf, Hubbard, Cooper, Wheelles, & Ronan, 2004), using the Presser criteria and the legislative criteria combined may be considerably valuable for those on the cusp of fitness.

Research indicates that the Pritchard criteria are by no means comprehensive, yet a comprehensive assessment is needed to spell out the competence of individuals (Kalbeitzer & Benedetti, 2009; Rogers et al., 2008). In general, a health assessor should give enough time

and thoroughness to examine the functioning of the individual fully relevant to a jurisdiction's psycho-legal criteria (Mossman et al., 2007). New Zealand's adaptation of the Pritchard criteria into the legislative criteria as a sole means of deciding fitness is unsatisfactory to some researchers and professionals (Brookbanks & Mackay, 2010; *P v. Police*, 2006; Samuels et al., 2007). The Presser criteria may even be more effective in determining fitness than the legislative criteria given that they have a larger AUC value, less residual deviance, and better specificity. Conversely, the legislative criteria logistic model was not significantly different when adding the Presser criteria extensions and had marginally better sensitivity than the Presser criteria model. A larger number of predictor variables may account for the improvements seen in the Presser criteria's AUC value rather than due to the properties of the criteria themselves. This may explain why there was a non-significant difference in the log likelihood model, aside from a low sample size. However, given the results, this puts into question the suitability of using the legislative criteria alone compared to including all relevant criteria.

This study supports including the Presser criteria in reports on Section 38 of the New Zealand CP[MIP] Act (2003) to improve specificity. However, this study used judges' rulings as the gold standard, which inevitably may be subject to error. Currently, there is no better gold standard for a defendant's true fitness (Mossman et al., 2010). Since there are no practice guidelines and no known structured assessments to measure FTST in New Zealand, and the rates of inter-rater reliability of health assessors' opinions are only moderately agreeable (Large et al., 2009), this infers that practitioners' assessments may be subject to various degrees of subjectivity. Since FTST assessments are not likely to be completely objective assessments (Gowensmith et al., 2012), it becomes necessary to improve the means of correct court findings for defendants through watertight standards and assessments

that include all relevant information, such as the Presser criteria.

Current study's limitations

This study used non-probability sampling for the Auckland and Northland regions. Therefore, no inferences are made about a particular population of defendants. The subtle differences between the wording of the overlapping legislative criteria and Presser criteria were not distinguished in this study. This study had a reduced sample size for the log likelihood test due to some health assessors not reporting all psycho-legal criteria. This study relied on court judgements as the gold standard or measure of true fitness. There were unequal sample sizes for the two groups in the AUC analysis, and therefore no inferences could be made about significant differences. No multiple comparison statistical methods were employed in this research, and there may be significant results merely from conducting many hypothesis tests. Other variables such as intelligence or the type of mental impairment may also be relevant to the outcome in court but were not considered in this study. Lastly, the research did not consider valuable additional hypotheses a priori, and testing AUC differences between these criteria would be beneficial to explore in the future.

Future research

Firstly, more research could identify how the Presser criteria or legislative criteria affect the opinion of health assessors. Secondly, a replication of this study with a larger sample size may provide evidence that the Presser criteria improve the model fit of the legislative criteria. Thirdly, there is a need to develop practice guidelines and structured FTST assessments to improve specificity, sensitivity and interrater reliability in FTST examinations. Fourth, applying significance tests to specificity, sensitivity and AUC statistics would inferentially support the results of this research. Lastly, it

would be useful, particularly for theoretical considerations, to conduct research using ordination methods or dimension reduction methods to assist in delineating and exploring core factors of criteria that predict court outcomes in New Zealand.

Conclusions

Psycho-legal criteria form the basis of FTST decisions to ensure that defendants have a fair trial. Given the mediocre rates of assessing all of the Presser criteria in within assessments (Birgden & Thomson, 1999; Nicholson & Norwood, 2000), the subjectivity involved in the assessment process (Large et al., 2009; Morris et al., 2004) and the importance of the Presser criteria raised by researchers and court officials (*P v. Police*, 2006; Samuels et al., 2007), this research attempts to highlight areas that would improve FTST practice. By examining the effect of the Presser criteria and legislative criteria on court outcome, the authors found that the Presser criteria do predict the outcome in court. However, there was insufficient evidence to suggest that the Presser criteria significantly improved the model of the outcome in court when adding them to the legislative criteria model. When using the Presser criteria to measure court outcome, there was an improvement in specificity compared to using the legislative criteria, and sensitivity was marginally superior for the legislative criteria model. Despite this, the authors found comparable results between the two sets of criteria in an AUC analyses. This research found that the Presser criteria have some utility to discern fitness in addition to the legislative criteria. Hence, the use of the Presser criteria may improve the ability of health assessors to assist the court in determining fitness, particularly for 'borderline-fit' defendants. Enabling defendants to have a fair trial may be the most important aspect within FTST hearings, and this research attempts to support this notion examining evidence to

reduce the potential for unfair trials for New Zealand defendants in the future.

Declaration of conflicts of interest

Khan Vladimir Buchwald-Mackintosh has declared no conflicts of interest.

Mei Wah Williams has declared no conflicts of interest.

Joseph Allan Sakdalan has declared no conflicts of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

This study is in agreement with the ethical principles outlined in the declaration of Helsinki.

Low risk notification was given to commence this study by Massey University Human Ethics Committee.

Court cases

P v. Police 2 NZLR 528 (HC) (2006).

R v. Presser VR 45 at 48 (1958).

R v. Pritchard 7 C & P 303 (1836).

References

- Birgden, A., & Thomson, D. (1999). The assessment of fitness to stand trial for defendants with an intellectual disability: A proposed assessment procedure involving mental health professionals and lawyers. *Psychiatry, Psychology and Law*, 6(2), 207–214.
- Brookbanks, W. (2014). Special Hearings under New Zealand's Criminal Procedure (Mentally Impaired Persons) Act 2003. *Northern Ireland Legal Quarterly*, 65(2), 215–218.
- Brookbanks, W., & Mackay, R. D. (2010). Decisional competence and 'best interests': Establishing the threshold for fitness to stand trial. *Otago Law Review*, 12(2), 265–284.
- Cram, S. J., & Dobson, K. S. (1993). Confidentiality: Ethical and legal aspects for Canadian psychologists. *Canadian Psychology*, 34(3), 347–363.
- Criminal Justice Act, New Zealand Statutes. (1985). Retrieved from <http://www.legislation.govt.nz/act/public/1985/0120/latest/whole.html>

- Criminal Procedure (Mentally Impaired Persons) Act, New Zealand Statutes. (2003). Retrieved from <http://www.legislation.govt.nz/act/public/2003/0115/latest/DLM223818.html>
- Felthous, A. R. (2011). Competence to stand trial should require rational understanding. *Journal of the American Academy of Psychiatry and the Law Online*, 39(1), 19–30.
- Gowensmith, W. N., Murrie, D. C., & Boccaccini, M. T. (2012). Field reliability of competence to stand trial opinions: How often do evaluators agree, and what do judges decide when evaluators disagree? *Law and Human Behavior*, 36(2), 130–139.
- Groth-Marnat, G. (2009). *Handbook of psychological assessment* (5th ed.). Hoboken, NJ: John Wiley & Sons.
- Hoge, S. K., Bonnie, R. J., Poythress, N., Monahan, J., Eisenberg, M., & Feucht-Haviar, T. (1997). The MacArthur adjudicative competence study: Development and validation of a research instrument. *Law and Human Behavior*, 21(2), 141–179.
- Kalbeitz, R., & Benedetti, R. (2009). Assessment of competency to stand trial in individuals with mental retardation. *Journal of Forensic Psychology Practice*, 9(3), 237–248.
- Kasinathan, J., Le, J., Barker, A., & Sharp, G. (2016). Presser - the forgotten story. *Australas Psychiatry*, 24(5), 478–482.
- Large, M., Nielssen, O., & Elliott, G. (2009). Reliability of psychiatric evidence in serious criminal matters: Fitness to stand trial and the defence of mental illness. *Australian and New Zealand Journal of Psychiatry*, 43(5), 446–452.
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). New York, NY: John Wiley and Sons.
- Marcus, D. K., Poythress, N. G., Edens, J. F., & Lilienfeld, S. O. (2010). Adjudicative competence: Evidence that impairment in “rational understanding” is taxonic. *Psychological Assessment*, 22(3), 716–722.
- Morris, G. H., Haroun, A. M., & Naimark, D. (2004). Health Law in the Criminal Justice System Symposium: Competency to stand trial on trial. *Houston Journal of Health Law and Policy*, 4, 193–238.
- Mossman, D., Bowen, M. D., Vanness, D. J., Bienenfeld, D., Correll, T., Kay, J., ... & Lehrer, D. S. (2010). Quantifying the accuracy of forensic examiners in the absence of a “gold standard”. *Law and Human Behavior*, 34(5), 402–417.
- Mossman, D., Noffsinger, S. G., Ash, P., Frierson, R. L., Gerbasi, J., Hackett, M., ... & Wall, B. W. (2007). AAPL practice guideline for the forensic psychiatric evaluation of competence to stand trial. *Journal of the American Academy of Psychiatry and the Law Online*, 35, s3–s72.
- Nicholson, R. A., & Norwood, S. (2000). The quality of forensic psychological assessments, reports, and testimony: Acknowledging the gap between promise and practice. *Law and Human Behavior*, 24(1), 9–44.
- Olley, M. C., Nicholls, T. L., & Brink, J. (2009). Mentally ill individuals in limbo: Obstacles and opportunities for providing psychiatric services to corrections inmates with mental illness. *Behavioral Sciences & the Law*, 27(5), 811–831.
- Parton, F., Day, A., & White, J. (2004). An empirical study on the relationship between intellectual ability and an understanding of the legal process in male remand prisoners. *Psychiatry, Psychology and Law*, 11(1), 96–109.
- Rogers, T. P., Blackwood, N. J., Farnham, F., Pickup, G. J., & Watts, M. J. (2008). Fitness to plead and competence to stand trial: A systematic review of the constructs and their application. *The Journal of Forensic Psychiatry & Psychology*, 19(4), 576–596.
- Sakdalan, J. A., & Egan, V. (2014). Fitness to stand trial in New Zealand: Different factors associated with fitness to stand trial between mentally disordered and intellectually disabled defendants in the New Zealand criminal justice system. *Psychiatry, Psychology and Law*, 21(5), 658–668.
- Samuels, A., O'Driscoll, C., & Allnutt, S. (2007). Fitness issues in the context of judicial proceedings. *Australasian Psychiatry*, 15(3), 212–216.
- Skeem, J. L., Golding, S. L., Cohn, N. B., & Berge, G. (1998). Logic and reliability of evaluations of competence to stand trial. *Law and Human Behavior*, 22(5), 519–547.
- Torrey, E. F., Kennard, A. D., Eslinger, D., Lamb, R., & Pavle, J. (2010). *More mentally ill persons are in jails and prisons than hospitals: A survey of the states*. Arlington, VA: Treatment Advocacy Center.
- Visser, S. (2011). Assessment and care of offenders with mental impairments. In F. W. Seymour, S. Blackwell & J. Thorburn (Eds.), *Psychology and the Law in Aotearoa New Zealand* (pp. 193–204). Wellington, New Zealand: New Zealand Psychological Society.
- Zapf, P. A., Hubbard, K. L., Cooper, V. G., Wheelles, M. C., & Ronan, K. A. (2004).

Have the courts abdicated their responsibility for determination of competency to stand trial to clinicians? *Journal of Forensic Psychology Practice*, 4(1), 27–44.

Appendix A: Legislative criteria

Unfit to stand trial, in relation to a defendant:

1. means a defendant who is unable, due to mental impairment, to conduct a defence or to instruct counsel to do so; and
2. includes a defendant who, due to mental impairment, is unable:
 - a. to plead;
 - b. to adequately understand the nature or purpose or possible consequences of the proceedings;
 - c. to communicate adequately with counsel for the purposes of conducting a defence. (CP[MIP] Act, 2003)

Appendix B: Presser criteria

1. Pleading to the charge and exercising his right of challenge;
2. Understanding what it is that he has been charged with;
3. Understanding that the proceedings before the court would be an inquiry as to whether or not he did what he was charged with;
4. Following in general terms, the course of the proceedings before the court;
5. Understanding the substantial effect of any evidence given against him;
6. Making a defence to, or answering, the charge;
7. Deciding what defence he would rely on;
8. Giving instructions to his legal representative (if any); and
9. Making his version of facts known to the court and to his legal representative, if any.