University of Kent

Kent Academic Repository

Cloonan-Thomas, Sinead, McEwan, Troy, Ogloff, James R. P., Ducat, Lauren and Tyler, Nichola (2026) *Predictive validity of the HCR-20³ for arson recidivism*. Psychology, Crime & Law . pp. 1-28. ISSN 1068-316X.

Downloaded from
https://kar.kent.ac.uk/116354/ The University of Kent's Academic Repository KAR

The version of record is available from
https://doi.org/10.1080/1068316X.2026.2731481

This document version
Publisher pdf

DOI for this version

Licence for this version
CC BY-NC-ND (Attribution-NonCommercial-NoDerivatives)

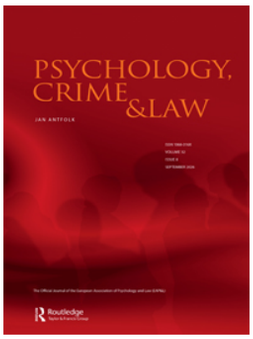
Additional information

Versions of research works

Versions of Record
If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

Author Accepted Manuscripts
If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in ***Title of Journal*** , Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

Enquiries
If you have questions about this document contact ResearchSupport@kent.ac.uk. Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies).



Routledge
Taylor & Francis Group

Psychology, Crime & Law

ISSN: 1068-316X (Print) 1477-2744 (Online) Journal homepage: www.tandfonline.com/journals/gpcl20

Predictive validity of the HCR-20^{V3} for arson recidivism

Sinead Cloonan-Thomas, Troy McEwan, James R. P. Ogloff, Lauren Ducat & Nichola Tyler

To cite this article: Sinead Cloonan-Thomas, Troy McEwan, James R. P. Ogloff, Lauren Ducat & Nichola Tyler (23 Sep 2026): Predictive validity of the HCR-20^{V3} for arson recidivism, Psychology, Crime & Law, DOI: 10.1080/1068316X.2026.2731481

To link to this article: https://doi.org/10.1080/1068316X.2026.2731481

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

View supplementary material

Published online: 23 Sep 2026.

Submit your article to this journal

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=gpcl20

PSYCHOLOGY, CRIME & LAW
https://doi.org/10.1080/1068316X.2026.2731481

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Predictive validity of the HCR-20^{V3} for arson recidivism

Sinead Cloonan-Thomas [a], Troy McEwan [a,b,c], James R. P. Ogloff [a,b], Lauren Ducat[a,b] and Nichola Tyler [a]

[a]Centre for Forensic Behavioural Science, Swinburne University of Technology and Forensicare, Melbourne, Australia; [b]Victorian Institute of Forensic Mental Health (Forensicare), Melbourne, Australia; [c]Centre of Research and Education in Forensic Psychology, University of Kent, Canterbury, UK

ABSTRACT

In the absence of a valid risk assessment tool for adult firesetting, practitioners often use general or violence risk assessment instruments to assess risk of future firesetting. This study aimed to examine the predictive validity of the Historical Clinical Risk Management-20 Version 3 (HCR-20^{V3}) for arson recidivism. A pseudo-prospective design was used, with risk assessments scored from file review in a sample of 75 participants assessed for firesetting at a community forensic mental health service in Victoria, Australia between 2003 and 2015. Participants were followed up using police records for a median of 5 years (range: 8 months to 12.6 years; interquartile range: 2 to 7 years). Seven participants (9.3%) received a further arson-related charge. Time-dependent analyses indicated that in the first two years post-assessment, a HCR-20^{V3} Future Violence/Case Prioritisation judgement of high risk showed good predictive validity for future arson. These findings offer preliminary support for the use of the HCR-20^{V3} Future Violence/Case Prioritisation judgement to identify individuals who may be at an elevated risk of arson reoffending in the first two years post-assessment. However, significant methodological limitations mean results must be interpreted with caution and require replication. Furthermore, application of the HCR-20^{V3} to firesetting requires informed adaptation and consideration of fire-related factors.

ARTICLE HISTORY
Received 19 March 2026
Accepted 25 August 2026

KEYWORDS
Arson; firesetting; risk assessment; HCR-20^{V3}; fire interest

Deliberate firesetting, legally referred to as arson, can cause severe harm including injury and death, as well as significant property and environmental destruction (Arson Prevention Forum, 2017; Campbell, 2021; Tyler et al., 2019). Given the potential harm associated with firesetting, people with arson convictions are commonly perceived as dangerous and likely to reoffend (Hardcastle et al., 2011; Nichols, 2022) and subsequently have trouble accessing employment and housing (Allender et al., 2005; Atkin & Armstrong, 2011; Ellison et al., 2013; Haslewood-Pócsik et al., 2008). A recent meta-analysis of apprehended firesetting adults and young people (i.e. those whose firesetting led to criminal justice

CONTACT Nichola Tyler ntyler@swin.edu.au Centre for Forensic Behavioural Science, Swinburne University of Technology, PO Box 218, Hawthorn, VIC 3122, Australia

Supplemental data for this article can be accessed online at https://doi.org/10.1080/1068316X.2026.2731481.

© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

contact), found that the base rate of arson reoffending was 8–10% and the rate of self-reported firesetting reoffending was 17–20%, indicating that as many as 1 in 5 people with a history of firesetting go on to repeat this behaviour (Sambrooks et al., 2021).

The potential high harms associated with firesetting, and the need for empirically informed decision-making, mean that it is important to be able to assess the risk of repeat firesetting in valid and reliable ways. Effective risk assessment is essential to informing case prioritisation, risk management, and tailored interventions (Andrews & Dowden, 2007; Bonta & Andrews, 2023; Logan, 2014). However, little academic attention has been given to firesetting risk assessment and there are no validated risk assessment instruments available that are specific to this purpose (Gannon et al., 2022). As a result, practitioners lack valid guidance to aid effective assessment and risk management of future firesetting risk (Gannon & Pina, 2010).

Firesetting-specific approaches to risk assessment

There have been some attempts to progress firesetting risk assessment knowledge and practice. For example, actuarial models based predominantly on static risk factors (i.e. unchangeable factors; young age at firesetting onset, history of contact with psychiatric services) have shown some validity for predicting future firesetting (Ducat et al., 2015; Edwards & Grace, 2013). However, these models are limited by the low base rate of reoffending in their respective samples and the absence of independent validation.

Firesetting-specific interview guides or aide memoires have also been developed to assist with information gathering and case formulation; namely the Northgate Firesetter Risk Assessment (NFRA; Taylor & Thorne, 2005, 2013, 2019), the St Andrew's Fire and Arson Risk Instrument (SAFARI; Long et al., 2014), and the Pathological Firesetter Interview (PSFI; Taylor et al., 2004). Although these instruments may aid the assessment and formulation processes, they are not intended to be used as standalone, predictive risk assessment tools. They are yet to be tested for this purpose, or for their stated purpose of improving the quality of firesetting assessment.

Potential causes and correlates of firesetting

Existing theory and research provide some indication of risk factors that may have utility when attempting to assess firesetting risk. The Multi-Trajectory Theory of Adult Firesetting (M-TTAF; see Gannon et al., 2012, 2022 for a full overview) is the most comprehensive multifactorial theory that attempts to explain adult firesetting and highlights potential risk markers and risk factors. The M-TTAF is a two-tiered model that conceptualises firesetting as a diverse behaviour that can arise from multiple developmental, psychological, and environmental factors rather than a single underlying cause. Tier 1 of the model integrates static and dynamic risk and protective factors such as caregiver environment, cultural forces, offence-supportive cognition, inappropriate fire interest, self-regulation issues, and communication problems.

Tier 2 of the M-TTAF demarcates five trajectories of firesetting based on critical risk factors, clinical features, and motivations: *antisocial, grievance, fire interest, emotionally expressive* or *need for recognition*, and *multi-faceted*. These trajectories highlight the diversity of firesetting functions and motivations. For instance, the *fire interest* subtype

describes those who are likely to exhibit an inappropriate fascination with fire, hold fire-supportive attitudes, and set fires for sensory stimulation or thrill-seeking reasons, while the *emotionally expressive* subtype are likely to experience difficulties with communication, self-regulation, and mental health, and use fire as a method to communicate distress (e.g. self-harm). This heterogeneity poses significant challenges for risk assessment and intervention, including understanding causal mechanisms underlying firesetting, selection of appropriate tools to guide risk assessment and case prioritisation, and identifying treatment targets.

Many of the risk factors and risk markers within the M-TTAF are consistent with findings from the empirical literature, although the limited evidence base precludes identification of causal risk and protective factors (Cloonan-Thomas et al., 2026). Studies of repeat firesetting have identified a number of general criminogenic factors associated with recidivism, including difficulties with education and employment (Dickens et al., 2009; Field, 2016; Rice & Harris, 1991), a history of offending and antisocial behaviour (Dickens et al., 2009; Ducat et al., 2015; Edwards & Grace, 2013; Field, 2016; Rice & Harris, 1991), substance use issues (Ducat et al., 2015; Repo et al., 1997; Repo & Virkkunen, 1997; Sambrooks et al., 2021), mental health difficulties (Dickens et al., 2009; Ducat et al., 2015; Field, 2016; Lindberg et al., 2005), personality disorder diagnosis (Dickens et al., 2009; Ducat et al., 2015; Repo et al., 1997; Rice & Harris, 1991; Wyatt et al., 2019) and relationship and communication problems (Dickens et al., 2009; Field, 2016; Ó Ciardha, Alleyne, et al., 2015). Many of these align with the 'central eight' risk/need factors for general offending (Bonta & Andrews, 2023).

In addition to risk factors for general offending, specific fire-related factors have been associated with firesetting (Gannon et al., 2013, 2023; Ó Ciardha, Barnoux, et al., 2015; Tyler et al., 2015) and repeat firesetting (Dickens et al., 2009; Ó Ciardha, Barnoux, et al., 2015; Rice & Harris, 1991; Sambrooks et al., 2024; Tyler et al., 2015). These include fire-related cognitive schema (i.e. knowledge structures comprising beliefs, attitudes, and scripts about fire), including fire interest (i.e. inappropriate fascination with fire) and fire scripts (i.e. cognitive rules about fire use). In addition to cognition, other fire-related factors include a history of formally recorded firesetting incidents (i.e. arson-related charges or convictions), childhood antisocial fire exposure, family history of firesetting, and child and adolescent fire play and firesetting (Barrowcliffe et al., 2022; Barrowcliffe & Gannon, 2015, 2016; Ducat et al., 2015; Edwards & Grace, 2013; Field, 2016; Horsley et al., 2022; Rice & Harris, 1996; Sambrooks et al., 2021; Tyler et al., 2015). Fire-related factors appear to be unique to those who set fires, suggesting these factors may be particularly relevant for firesetting risk assessment, formulation, and treatment (Gannon et al., 2013).

Use of existing risk assessment tools for firesetting

In the absence of a firesetting-specific risk assessment tool, some authors have recommended that practitioners use well-validated violence or general risk assessment tools to guide risk assessment of future firesetting (Gannon & Pina, 2010; McEwan et al., 2012). The Historical Clinical Risk Management-20 Version 3 (HCR-20^{V3}; Douglas et al., 2013; Webster et al., 1995, 1997) is a structured professional judgement (SPJ) tool designed to assess risk of violence recidivism in adults. The HCR-20^{V3} is one of the most extensively researched and widely used violence risk assessment tools worldwide

(Singh et al., 2014), and has been found to have good psychometric properties, predictive validity, and clinical value (Brookstein et al., 2020; Douglas et al., 2017; Douglas & Shaffer, 2021; Doyle et al., 2014). Arson and firesetting are included in the HCR-20^{V3}'s definition of violence, when the act of firesetting is aimed at causing harm to another person (Douglas et al., 2013). However, firesetting without violent intent (e.g. to damage property) is not considered within the tool's definition of violence. Therefore, researchers have tentatively suggested the HCR-20^{V3} may be suitable for assessing firesetting risk in cases where harmful intent is present (Edwards, 2020; Gannon et al., 2022; Gannon & Pina, 2010). While the HCR-20^{V3} may capture several empirically supported correlates that are shared by violence and firesetting, the tool does not explicitly assess fire-related factors (e.g. fire interest, fire scripts) that may contribute to firesetting risk.

Given that not all firesetting is motivated by intent to harm (Gannon et al., 2012, 2022) and that firesetting and general offending share a number of historical and dynamic risk factors, some researchers have suggested the Level of Service instruments (Andrews et al., 2004, 2008) may also be suitable for firesetting risk assessment (Doley et al., 2011; McEwan et al., 2012). Similarly to the HCR-20^{V3}, however, these instruments were not specifically designed to assess firesetting and do not provide guidance on the assessment of fire-related factors meaning they may neglect to capture information that is important for informing firesetting risk assessment, formulation, and risk management. Furthermore, while both violence and general offending risk assessment tools are used in practice, they have not been validated for use with firesetting as an outcome.

The current study

In the absence of any research testing recommendations about the use of violence risk assessment tools for police-recorded firesetting (i.e. arson), this study had three aims. First, to evaluate whether judgements made using the HCR-20^{V3} can provide a valid assessment of arson risk. Second, to compare the predictive validity of HCR-20^{V3} judgements for arson reoffending to the predictive validity of the same judgements for violent reoffending in the same sample. Third, to investigate whether the addition of fire-related factors (i.e. history of firesetting, fire interest) incrementally improves the predictive validity of HCR-20^{V3} judgements for arson.

These aims give rise to three hypotheses: (1) judgements informed by the HCR-20^{V3} will provide a valid risk assessment for arson reoffending; (2) the predictive validity of HCR-20^{V3} judgements will be stronger for future violent offending compared to arson offending; and (3) fire-related factors will add significantly to the predictive ability of HCR-20^{V3} judgements when predicting future arson.

Method

Participants

Data were drawn from a community forensic mental health service in Victoria, Australia. Offering a statewide assessment and treatment service, the program takes referrals from the courts, community, mental health, and correctional services for individuals who are exhibiting problem behaviours such as violence, problematic sexual behaviour, stalking,

and firesetting. Individuals included in this study were referred to the service and assessed for the problem behaviour of firesetting (which may or may not have involved criminal charges for arson). Participants were identified from the intake database by searching for referrals involving firesetting between 31 October 2003 and 24 July 2015.

A total of 129 referrals were identified, but 54 were excluded from the current study due to the referral being withdrawn or not accepted, the referral not relating to the primary problem of firesetting, because the file could not be located, or because reoffending data was not available. Three participants who did not have available reoffending data were included in the sample. These participants had no evidence of a formal offending history (i.e. were referred by mental health or community services and there was no reference to past offending in the file) and so there would be no expectation that they would be matched to a criminal history. They were included in the 'non-recidivism' group for predictive analyses.

Sample description and index offence characteristics

The sample used in analyses comprised 75 firesetting participants (61 male, 81.3%; 14 female, 18.7%) with ages ranging from 17–64 years ($M = 32$, $SD = 11.44$). Participants were referred from probation services (56.0%), mental health services (22.7%), courts or parole board (14.6%), or other sources (e.g. self-referral or community-based services; 6.7%). Demographic information is shown in Table 1.

In the firesetting that led to referral and assessment, 34.7% ($n = 26$) of participants targeted property (e.g. residential, business, own home), 13.3% ($n = 10$) targeted a car or other vehicle, 12.0% ($n = 9$) targeted small or miscellaneous items (e.g. rubbish bin, mattress), 9.3% ($n = 7$) targeted bush or vegetation, and 4.0% ($n = 3$) used fire for self-

Table 1. Demographic characteristics for the final sample ($N = 75$).

Demographic characteristics	*N* (%)	Missing *N* (%)
Male	61 (81.3)	0 (0.0)
Age at index offence	Range = 17–64 $M = 32.27$, $SD = 11.44$	8 (10.7)
Single at assessment	45 (60.0)	5 (6.7)
Completed secondary schooling	19 (25.3)	2 (2.7)
Employed at index offence	17 (22.7)	13 (17.3)
Offending History		
Any criminal history[a]	60 (80.0)	0 (0.0)
Prior violent offence charge	35 (46.7)	0 (0.0)
Prior non-violent offence charge	53 (70.7)	1 (1.3)
Prior property damage charge	34 (5.3)	0 (0.0)
Prior arson charge	17 (22.7)	1 (1.3)
Prior firesetting history[b]	42 (56.0)	1 (1.3)
Clinical Factors		
Any mental health diagnosis	58 (77.3)	0 (0.0)
Diagnosed psychotic disorder	26 (34.7)	0 (0.0)
Diagnosed personality disorder	17 (22.7)	10 (13.3)
Fire Factors		
Fire interest	12 (16.0)	19 (25.3)
Juvenile firesetting	14 (18.7)	19 (25.3)
History of fire play	15 (20.0)	24 (32.0)

[a]Any charge, conviction, or self-reported offending.
[b]Any documented history of firesetting (i.e. charge, conviction, self-report, other report).

immolation purposes. In 44.0% ($n = 33$) of cases, the victim of the index arson was known to the perpetrator (e.g. targeting the property or vehicle of an ex-partner or family member). In 16.0% ($n = 12$) of cases, there was stated intent for the firesetting to endanger the life of the victim and in 30.7% ($n = 23$) of cases, lives were endangered by location (i.e. the location or circumstances of the fire created a reasonably foreseeable risk of harm to others). Firesetting resulted in serious injury or death in 5.3% ($n = 4$) of cases.

Of the total sample, 38 participants (50.7%) were charged with a further offence following the date of assessment, with 31 (41.3%) receiving a non-violent charge and 21 (28.0%) receiving a violent charge. The arson recidivism rate was substantially lower, with seven participants (9.3%) subsequently charged with arson. Of the arson recidivism group, 71.4% ($n = 5$) were male with a mean age of 40.5 years ($SD = 6.72$; range = 32–49).

Measures

Demographic, offence history, and clinical variables

Data extracted from clinical files included information about sex, age, marital and occupation status, education level, family of origin details, and mental health and substance use history. Information on prior offending history was coded from formal records available to the original assessing clinician (i.e. charges and convictions) and recorded self-report, including previous arson and arson-related offending, as well as violent and non-violent offending.

Historical Clinical Risk Management – 20 Version 3 (HCR-20^{V3}; Douglas et al., 2013)

The Historical Clinical Risk Management – 20 Version 3 (HCR-20^{V3}; Douglas et al., 2013) is a SPJ guideline for assessing risk of violent recidivism in males and females aged 18 years and above. The HCR-20^{V3} consists of 20 items across three scales: Historical (H-scale), including 10 historical factors (e.g. history of problems with violence, employment, substance use); Clinical (C-scale) consisting of five dynamic factors (e.g. recent problems with insight, symptoms of major mental illness, instability); and Risk Management (R-scale), consisting of contextual dynamic risk factors relevant to future interventions and supports (e.g. future problems with professional services, living situation, personal support).

The presence of each risk factor was scored on a three-point scale (0 = absent, 1 = partially present, 2 = present) and its relevance to future risk management was rated as low, moderate, or high. Risk factor presence and relevance were then used to inform a categorical summary risk judgement for Future Violence/Case Prioritisation, classified as low, moderate, or high risk. Summary risk judgements for Severity (i.e. the potential seriousness of harm) and Imminence (i.e. the temporal proximity of violent behaviour) were classified as low, moderate, or high. An 'imminent' timeframe was considered 'hours to days or days to weeks' in accordance with the HCR-20^{V3} manual (Douglas et al., 2013, p. 63).

In addition to the summary risk judgements, total scores were calculated for research purposes to determine whether the presence of risk factors (rather than the raters' structured judgement) demonstrated predictive validity for future firesetting. Total scores were calculated by summing the presence ratings for each risk factor to create a continuous

score ranging from 0–40. Risk scenarios and formulations were constructed informally by the raters to inform summary risk judgements but were not recorded for research purposes. HCR-20^{V3} judgements have been found to have good predictive validity for violent reoffending in Australian forensic psychiatric samples (Brookstein et al., 2020; Shepherd et al., 2017). Descriptive information for HCR-20^{V3} items, scale scores, and judgements for the full sample are available in Table 2.

Missing Data. Items on the HCR-20^{V3} were omitted where there was insufficient file information to make a rating (per Douglas et al., 2013). Risk factors pertaining to attitudes, ideation, and emotional/behavioural stability had a substantial number of omissions (>10%) due to a lack of available file information. While the HCR-20 Version 2 specifies a threshold for the number of items that could be omitted before the test was deemed invalid (no more than four items, with no more than two on the Historical scale and no more than one on the Clinical and Risk Management scales; Webster et al., 1997), the HCR-20^{V3} manual does not set a numerical limit on item omissions. Rather, it suggests that the measure becomes invalid when missing information prevents a defensible risk judgement from being made (Douglas et al., 2013). In the current sample, the number

Table 2. HCR-20^{V3} descriptive information ($N = 75$).

Scale/Risk Factor	Scores *N* (%)			Omissions *N* (%)
	No	Partial	Yes	
Historical				
H1. Violence	0 (0.0)	24 (32.0)	51 (68.0)	0 (0.0)
H2. Other antisocial behaviour	15 (20.0)	25 (33.3)	35 (46.7)	0 (0.0)
H3. Relationships	2 (2.7)	14 (18.7)	57 (76.0)	2 (2.7)
H4. Employment	13 (17.3)	24 (32.0)	38 (50.7)	0 (0.0)
H5. Substance use	17 (22.7)	11 (14.7)	45 (60.0)	2 (2.7)
H6. Major mental disorder	13 (17.3)	13 (17.3)	48 (64.0)	1 (1.3)
H7. Personality disorder	26 (34.7)	24 (32.0)	15 (20.0)	10 (13.3)
H8. Traumatic experiences	18 (24.0)	15 (20.0)	36 (48.0)	6 (8.0)
H9. Violent attitudes	38 (50.7)	19 (25.3)	9 (12.0)	9 (12.0)
H10. Treatment or supervision response	28 (37.3)	15 (20.0)	32 (42.7)	0 (0.0)
Clinical				
C1. Insight	9 (12.0)	21 (28.0)	44 (58.7)	1 (1.3)
C2. Violent ideation or intent	51 (68.0)	3 (4.0)	6 (8.0)	15 (20.0)
C3. Symptoms of major mental disorder	37 (49.3)	14 (18.7)	24 (32.0)	0 (0.0)
C4. Instability	19 (25.3)	16 (21.3)	26 (34.7)	14 (18.7)
C5. Treatment or supervision response	37 (49.3)	20 (26.7)	17 (22.7)	1 (1.3)
Risk Management				
R1. Professional services and plans	42 (56.0)	18 (24.0)	11 (14.7)	4 (5.3)
R2. Living situation	26 (34.7)	23 (30.7)	20 (26.7)	6 (8.0)
R3. Personal support	17 (22.7)	27 (36.0)	29 (38.7)	2 (2.7)
R4. Treatment or supervision response	14 (18.7)	38 (50.7)	23 (30.7)	0 (0.0)
R5. Stress coping	3 (4.0)	23 (30.7)	43 (57.3)	6 (8.0)
Total Scores	**Range**	***Mean***	***SD***	**Omissions *N* (%)**
H-Scale Total Score	3–19	12.21	3.67	0 (0.0)
C-Scale Total Score	0–10	4.11	2.54	0 (0.0)
R-Scale Total Score	0–10	5.08	2.54	0 (0.0)
Overall Total Score	4–37	21.40	7.37	0 (0.0)
Judgements	**Judgement Level *N* (%)**			**Omissions *N* (%)**
	Low	Moderate	High	
Future Violence/Case Prioritisation	22 (29.3)	35 (46.7)	18 (24.0)	0 (0.0)
Serious Physical Harm	41 (54.7)	27 (36.0)	6 (8.0)	1 (1.3)
Imminent Violence	52 (69.3)	18 (24.0)	4 (5.3)	1 (1.3)

of omissions per participant ranged from 0–7 and only three cases (4.0%) had more than four omissions. A Future Violence/Case Prioritisation judgement was made for all participants, and all but one were given Severity and Imminence judgements. The HCR-20^{V3} scale scores and total scores were computed using all available data without prorating, as this provides the most comparable level of information available to the clinician using the guideline.

Fire-related variables

Information about firesetting was collected using a data collection form established for a related project (Nanayakkara et al., 2020). Eighty-seven variables were collected from participant files, including history of firesetting and fire play, apparent motivation for the firesetting that led to referral (e.g. revenge, boredom, excitement), contextual and behavioural variables associated with the firesetting that led to referral (e.g. fire target, use of an accelerant) and clinical variables (e.g. active symptoms of psychosis at the time of firesetting).

Prior history of firesetting was operationalised as any documented evidence of firesetting prior to the index arson, regardless of whether the firesetting resulted in criminal justice contact or not. This item was coded dichotomously (yes/no) and included any recorded arson or arson-related convictions or charges, self-reported firesetting, and firesetting reported by others (e.g. correctional or residential staff, carers).

A single item measured participants' interest in fire and related paraphernalia based on self- or other-report. Indicators of fire interest included: likes watching fires, likes watching videos about fires, thinks about fire regularly, collects fire-related items such as matches or fire brigade items. This item was coded as a binary (yes/no) variable with items coded as 'yes' if evidence of any self- or other-reported fire interest was present in the file.

Definitions of recidivism

Recidivism in the current study was defined by the presence of an authorised police charge for a relevant offence, noting it is possible that charges may have been withdrawn or did not result in a conviction. This study was interested in the outcomes of arson and violent recidivism. Arson recidivism was defined as any subsequent arson (e.g. arson causing death, criminal damage by fire, intentionally causing a bushfire) or arson-related (e.g. lighting a fire during a prohibited period, failing to extinguish a fire, leaving a fire unattended) charge (for a complete list of charges see Supplementary Materials, Table 1). Violent reoffending was defined as a subsequent violent charge against another person, including offences in the categories of murder, manslaughter, assault, and aggravated robbery.

Procedure

This study employed a pseudo-prospective design. The HCR-20^{V3} was scored from clinical file information between January and May 2016 by two raters who were blind to reoffending outcomes and to the outcomes of other risk assessment instruments used at the time of assessment. Both raters were early career psychologists with doctoral training in forensic psychology and specific training in the HCR-20^{V3}. To establish interrater reliability, 11 cases (15% of the sample) were randomly selected to be independently rated by both

Table 3. Interrater reliability of the HCR-20^{V3}.

HCR-20^{V3} Measure	ICC	95% CI
H-Scale	.94	.80–.98
C-Scale	.84	.51–.95
R-Scale	.82	.46–.95
Total Scale Score	.98	.93–.99
Future Violence/Case Prioritisation Judgement	.69	.18–.91
Severity Risk Judgement	.81	.45–.94
Imminence Risk Judgement	.90	.69–.97

Note: ICC = Intraclass Correlation Coefficient; CI = Confidence Interval.

raters. Table 3 presents intra-class correlation coefficients (ICC) for the HCR-20^{V3} scale scores, total score, and risk judgements. The H-scale and total scores showed excellent agreement between the two raters while the C-scale and R-scale scores demonstrated good agreement. The Severity ($\kappa = .79$) and Imminence ($\kappa = .81$) judgements showed good interrater reliability. Weighted kappa for the Future Violence/Case Prioritisation judgement ($\kappa = .52$) was the lowest of the three summary risk judgements, reflecting fair agreement between the raters despite high agreement on the total score. The Future Violence/Case Prioritisation judgements showed disagreement in three cases: two Low-Moderate disagreements and one Moderate-High disagreement. No Low-High discrepancies were observed, indicating modest differences in clinical judgement, rather than substantial disagreement. Discrepancies between the two raters were not resolved; rather, the original rater's scores were retained for analysis.

Raters accessed clinical file material up until the point of the original assessment, including both the assessing psychologist's or psychiatrist's notes in addition to referral material (depending on the referrer this could include a copy of the formal criminal history, offence details, judges' sentencing comments, other material relied upon by the courts, and area mental health service assessments). Therefore, dynamic risk factors were rated based on information available to the original assessing clinician at the time of the original assessment. Once all data was collected, identifying information was removed from the file except for a unique identifier used to link with police charge data.

Data for this study were collected with a waiver of consent due to the use of a pre-existing dataset and the impracticality of finding participants to obtain their consent to have their data employed in this study. Ethical approval was granted by the Swinburne University Human Research Ethics Committee (SUHREC; Ref 2015315; Ref 20247726-18406), the Victoria Police Human Research Ethics Committee (VPHREC; Ref 213/14) and the Forensicare Operational Research Committee.

Linkage procedures

Recidivism data were obtained from the Law Enforcement Assistance Program (LEAP) maintained by Victoria Police. LEAP is a state-wide operational policing database that stores particulars of all crimes involving police response. It is a restricted online site that is updated in real time by Victoria Police members who enter details of all contact between police and members of the community.

The data linkage procedure first involved assigning a unique identifier code to each participant. Data were then shared with Victoria Police who extracted exact and probable matches from LEAP using combinations of identifying information (first name, surname,

aliases, date of birth, age range and gender). Offending after the index assessment date was extracted for the current study on 30 June 2016. Information extracted included the charge date and time, the date the offence was committed, the location of the offence, and how the individual was processed (i.e. summonsed to appear, arrested, cautioned).

Once the data were extracted, all data were de-identified for confidentiality purposes and returned to the researchers with the unique identifier code only. Where there were matches, LEAP records were obtained for the period of June 2004 to June 2016. Follow-up time was from the date of assessment to 30 June 2016, with follow-up time ranging from 8 months to 151 months or 12.6 years ($M = 64.87$ months, 4.96 years; $SD = 42.86$ months, 3.62 years). Actual time at risk during the follow-up period was not controlled for, as information regarding periods of reduced offending risk (e.g. incarceration) was unavailable.

Statistical analyses

Statistical analyses were conducted using IBM SPSS v31 and RStudio v2025.09 with scripts written in R-Project v4.5.2. The demographic, clinical, offending and fire-specific variables contained some missing data which were determined to be missing at random ($p = .61$). Accordingly, listwise deletion was applied so that cases were excluded from analyses where data were missing but retained in all other analyses (Tabachnick & Fidell, 2019).

Interrater reliability was assessed using intraclass correlation coefficients (ICC; two-way random, absolute agreement, single measures; Koo & Li, 2016; Shrout & Fleiss, 1979) for continuous variables. For ordinal variables (e.g. summary risk judgements), Cohen's weighted kappa with linear weights was also employed as the recommended statistic for ordinal data (Cohen, 1968; Gwet, 2014). ICCs were additionally calculated for ordinal variables to allow for comparison with previous studies using this approach (e.g. Brookstein et al., 2020). Kendall's tau was used to examine correlations between arson recidivism and HCR-20^{V3} items, scale scores, and summary risk judgements.

We examined the ability of the HCR-20^{V3} total score and each of the summary risk judgements (Future Violence/Case Prioritisation, Severity, Imminence) to discriminate between those with and without arson reoffending using the receiver operating characteristic (ROC) and corresponding area under the curve (AUC; Singh et al., 2015). Discrimination analyses focused primarily on the Future Violence/Case Prioritisation risk judgement as this is the judgement most commonly used to direct resource allocation and inform risk management decisions (Douglas et al., 2013).

Following examination of the AUCs for the total follow-up period, time-dependent AUC analyses were conducted for the HCR-20^{V3} total score and Future Violence/Case Prioritisation judgement at one, two, three, four, and five years following assessment to understand the discriminative ability of the tool over time. Time-dependent AUCs were not calculated for violent reoffending because dates of subsequent charges were not available.

As there are no universal benchmarks for the interpretation of AUCs (Challinor et al., 2021), the following thresholds were adopted based on the extant literature on forensic risk assessment: <.65 = small effect, .65–.70 = moderate effect; and >.70 = large effect (Brookstein et al., 2020; Dolan & Doyle, 2000). In addition to evaluating discrimination, sensitivity (the proportion of people who are correctly classified as being at increased

risk among all those who reoffend) and specificity (the proportion of people correctly classified as being at lower risk among all those who do not reoffend) statistics were calculated for each risk category to provide a full picture of classification accuracy.

Predictive power is often assessed using calibration, which can be measured by comparing recidivism rates in each risk category to rates in a large, comparable, independent sample rated with the same tool (Hanson, 2017), but such a sample is not available for arson reoffending. Alternatively, calibration can be established by comparing recidivism rates in each risk category to expected values from a logistic function. This is appropriate for an actuarial risk assessment tool constructed using logistic regression. However, SPJ guidelines like the HCR-20^{V3} are constructed rationally and there is no reason to think that results might be consistent with a logistic function. Positive and negative predictive power (PPV and NPV) are a more appropriate predictive measure for SPJ guideline evaluation. They serve a similar function (e.g. PPV allows comparison of the probability of correctly identifying someone as reoffending given their risk categorisation with the probability of correct identification given the base rate of reoffending), but they are more meaningful for clinicians as they include information about expected probability of reoffending within the group assigned to the risk category. Therefore, predictive power of the Future Violence/Case Prioritisation judgement was evaluated using PPV (the probability that someone classified as being at increased risk will reoffend) and NPV (the probability that someone classified as being at decreased risk will not reoffend). Calculation of sensitivity, specificity, PPV and NPV requires a dichotomous outcome, therefore, the risk categories were grouped as 'low versus (moderate & high)' and '(low & moderate) versus high' to give a full picture of classification and prediction at different cutoff points.

Finally, to examine the effects of time on reoffending, and provide context for interpreting AUC values, Kaplan-Meier survival analysis was undertaken, stratified by Future Violence/Case Prioritisation and Imminence risk levels. Log rank tests were performed to assess differences in the pattern of arson recidivism between low, moderate, and high risk judgements. Due to the small sample size and long follow-up period (12.6 years), the follow-up time was truncated when 20% of the sample was lost (8.42 years) to increase the stability of the results (Gebski et al., 2018; Pocock et al., 2002).

Analyses of fire-related factors (i.e. history of firesetting and fire interest) and subsequent arson offences used Fisher's Exact test to examine univariate relationships due to the small sample size and low expected cell frequencies. Odds ratios were derived as a measure of effect size, applying the Haldane-Anscombe correction for zero event instances (Lawson, 2004). We originally intended to conduct a hierarchical logistic regression to assess whether having a history of firesetting and fire interest incrementally added to the predictive validity of the HCR-20^{V3}. However, due to the limited sample size and very low proportion of arson recidivists reporting fire interest, this analysis could not be pursued.

Given the small number of arson recidivists in this study, results are likely to be uncertain, as demonstrated by wide ranges in the 95% confidence intervals provided for all predictive analyses. However, 95% confidence intervals alone may not clearly demonstrate the uncertainty in results; therefore, we examined the robustness of the results by moving one reoffending participant from the high to moderate risk group and rerunning all analyses. This sort of robustness testing is less common in risk assessment literature,

but relatively standard in other fields as a way of demonstrating how certain or uncertain a finding is to changes in the underlying data (see for example Greenland, 1996)

Results

HCR-20^{V3} and arson recidivism

The prevalence of arson and violence recidivism within each HCR-20^{V3} risk category is shown in Table 4. Approximately 28% ($n = 26$) of participants had a Future Violence/Case Prioritisation risk judgement of low, 47% ($n = 44$) were judged to be at moderate risk, and 25% ($n = 24$) were judged to be at high risk of future violence.

Table 5 presents Kendall's tau correlations between arson recidivism and the HCR-20^{V3} risk factors, scale scores, total score, and summary risk judgements. Risk factors *H6. History of Problems with Major Mental Disorder* and *C3. Recent Problems with Symptoms of Major Mental Disorder* showed positive, significant correlations with firesetting recidivism. This indicates that lifetime diagnosis of a major mental disorder (e.g. psychotic disorder, major mood disorder, other disorders that crossly impair intellectual, cognitive, or social/interpersonal functioning) and active symptoms of mental disorder in the months preceding the assessment were both associated with repeat arson. No significant correlations were found between arson recidivism and any of the HCR-20^{V3} scale scores or judgements.

Kendall's tau correlations showed that the HCR-20^{V3} total scale score was positively and significantly correlated with the three risk judgements: Future Violence/Case Prioritisation ($\tau = .68$, $p < .001$), Severity ($\tau = .44$, $p < .001$), and Imminence ($\tau = .54$, $p < .001$). These results suggest good internal consistency, with higher total scores corresponding with higher risk judgements.

Predictive validity of the HCR-20^{V3} for arson and violent recidivism

Table 6 summarises the results of ROC analyses examining the ability of the HCR-20^{V3} total score and summary risk judgements to discriminate between those with and without arson or violent reoffending. For arson, all AUCs were non-significant, and the effect ranged from small to moderate, indicating that the HCR-20^{V3} total score and summary risk judgements did not discriminate between those with and without arson reoffending

Table 4. HCR-20^{V3} risk judgements for arson and violent recidivists.

HCR-20^{V3} risk judgement	Arson recidivism (n = 7) *n* (%)	Violent recidivism *n*(n = 21) *n* (%)
Future Violence/Case Prioritisation		
Low/Moderate	3 (42.9)	10 (47.6)
High	4 (57.1)	11 (52.4)
Severity[a]		
Low	2 (28.6)	6 (30.0)
Moderate/High	5 (71.4)	14 (70.0)
Imminence[a]		
Low	4 (57.1)	9 (45.0)
Moderate/High	3 (42.9)	11 (55.0)

Note: Risk categories were combined to protect participant confidentiality due to small cell sizes.
[a]Violent recidivism: missing, $n = 1$ (4.8%).

Table 5. Kendall's tau correlations between arson recidivism and HCR-20^{V3} items, scale scores, and risk judgements.

Variable	Arson Recidivism
Historical Items	
H1. Violence	.12
H2. Other antisocial behaviour	.13
H3. Relationships	−.10
H4. Employment	.08
H5. Substance use	−.03
H6. Major mental disorder	.22*
H7. Personality disorder	.19
H8. Traumatic experiences	−.05
H9. Violent attitudes	.20
H10. Treatment/supervision response	.03
Clinical Items	
C1. Insight	.06
C2. Violent ideation or intent	−.12
C3. Symptoms of major mental disorder	.25*
C4. Instability	.12
C5. Treatment/supervision response	.09
Risk Management Items	
R1. Professional services/plans	.01
R2. Living situation	.03
R3. Personal support	−.11
R4. Treatment/supervision response	.09
R5. Stress coping	.03
Scale Scores	
H-Scale	.14
C-Scale	.12
R-Scale	.01
Total Scale Score	.09
Risk Judgements	
Future Violence/Case Prioritisation	.20
Severity	.13
Imminence	.10

*$p < .05$.

over the entire follow-up period. Robustness analyses showed a minor reduction in AUC estimates, demonstrating some uncertainty in the results given the small number of arson recidivists.

For violent recidivism, the Future Violence/Case Prioritisation risk judgement effectively discriminated between individuals with and without violent recidivism across the

Table 6. Discrimination performance of the HCR-20^{V3} for arson and violent recidivism ($N = 75$).

Risk Judgement	Arson Recidivism AUC (95% CI)	Violent Recidivism AUC (95% CI)
Future Violence/Case Prioritisation	.69 (.47–.91)	.77** (.65–.89)
	.64 (.43–.86)	*.75** (.64–.87)*
Severity[a]	.62 (.42–.81)	.67* (.53–.80)
	.56 (.34–.77)	*.66* (.52–.79)*
Imminence	.58 (.35–.82)	.67* (.53–.82)
	.56 (.16–.58)	*.67* (.52–.81)*
Total Score	.61 (.35–.87)	.78** (.65–.90)

Notes: AUC = Area Under the Curve; 95% CI = 95% Confidence Interval. The *italicised* results show the results of robustness testing conducted to demonstrate the level of uncertainty in the results, given the small number of recidivists. For these analyses, one recidivist was reassigned from the High to Moderate risk category.

*$p < .05$.

**$p < .01$.

[a]Given no arson recidivists were classified as High risk for Severity, one participant was instead reassigned from the Moderate to Low risk category for robustness testing.

entire follow-up period, with a large effect size, as did the total score (see Table 6). Effect sizes for Severity and Imminence judgements were lower but statistically significant. Robustness analyses revealed little change in the AUC estimates, showing greater certainty in the results, likely due to the greater number of violent recidivists in the sample.

Time-dependent discrimination of the HCR-20^{V3} for arson recidivism

Time-dependent AUC analyses were conducted for the HCR-20^{V3} total score and Future Violence/Case Prioritisation judgement to examine predictive validity for arson recidivism at various points in the first five years of follow-up (see Table 7). Over the first two years both the total score and Future Violence/Case Prioritisation judgement demonstrated statistically significant discriminatory effects, which remained during robustness testing. Across this time the total score showed a somewhat stronger discriminatory effect than the Future Violence/Case Prioritisation judgement. Beyond the first two years, a decline in discriminatory ability was observed for both, with widening confidence intervals indicating increased uncertainty around the results. The Future Violence/Case Prioritisation judgement showed a steady decline over time, and while AUCs remained significant until the fifth year, robustness testing showed that results were quite vulnerable to small changes in the underlying data, meaning the size of discriminatory effects after two years should be interpreted very cautiously, even when statistically significant. The discriminatory ability of the HCR-20^{V3} total score also declined over time, with a marked reduction evident in the third year and a non-significant AUC of .53 by the end of the fifth year.

Prediction indices for arson and violent recidivism

Table 8 presents classification and prediction indices (sensitivity, specificity, PPV, and NPV) for arson and violent recidivism based on the HCR-20^{V3} Future Violence/Case Prioritisation

Table 7. Time-dependent AUCs for arson recidivism over five time points for the HCR-20^{V3} future violence/case prioritisation judgement and total score.

Time at Risk	AUC	95% CI
Future Violence/Case Prioritisation Judgement		
12 months (1 year)	.90	.85–.95
	.78	*.59–.98*
24 months (2 years)	.79	.63–.96
	.70	*.51–.89*
36 months (3 years)	.71	.52–.90
	.64	*.46–.82*
48 months (4 years)	.70	.51–.89
	.63	*.45–.81*
60 months (5 years)	.57	.30–.85
	.52	*.27–.77*
Total Score		
12 months (1 year)	.95	.90–1.00
24 months (2 years)	.87	.74–1.00
36 months (3 years)	.69	.38–1.00
48 months (4 years)	.69	.38–.99
60 months (5 years)	.53	.19–.87

Notes: AUC = Area Under the Curve; 95% CI = 95% Confidence Interval. The *italicised* results show the results of robustness testing conducted to demonstrate the level of uncertainty in the results, given the small number of recidivists. For these analyses, one recidivist was reassigned from the High to Moderate risk category.

judgement at both moderate and high risk cut-offs. For arson recidivism, using a moderate risk cut-off produced high sensitivity but low specificity values, indicating most recidivists were correctly identified but at the cost of a 69% false positive rate. Conversely, using a high risk cut-off reduced sensitivity to 57%, identifying fewer than two-thirds of arson recidivists, but improved specificity markedly (79%), leading to a dramatically reduced false positive rate.

PPV and NPV must be interpreted in relation to the base rate of arson (9.3%). When using a moderate risk cut-off, PPV showed only a modest improvement in identifying arson recidivists over chance (11.3% *vs* 9.3%). However, when a high risk cut-off was used, PPV increased, with the probability of recidivism more than twice the base rate (22.2% *vs* 9.3%). Conversely, at a high risk cut-off, NPV suggested a small improvement in identifying individuals who did not go on to reoffend using arson, with 94.7% of those classified as low risk not going on to reoffend compared with 90.7% based on the base rate alone. Given the small number of arson recidivists, these results must be interpreted with caution, as demonstrated by the robustness testing, where movement of a single case from high risk to moderate risk substantially changed the classification and prediction metrics (see Table 8).

When considering violent recidivism, a similar pattern emerges, where a moderate risk cut-off yielded high sensitivity but low specificity, with moderate/high judgements correctly capturing 95% of violent recidivists at the cost of a 61% false positive rate. When a high risk cut-off was used, specificity increased and sensitivity decreased. At a high risk cut-off, the sensitivity value was slightly lower for violent (52.4%) compared to arson (57.1%) recidivism, indicating that the HCR-20^{V3} is capturing approximately half of those who reoffended violently in the high risk group.

At a moderate risk threshold for violent recidivism, PPV exceeded the base rate, with 37.7% of individuals classified as moderate or high risk subsequently reoffending violently compared to 28% based on the base rate alone (approximately 1.3 times higher). At a high risk cut-off, the probability of someone in the high risk group reoffending violently was 61.1%, which is more than twice the base rate, representing a comparable relative increase above the base rate observed for arson (see Table 8). At a moderate risk cut-off, NPV for violence was equivalent to the NPV for arson. At a high risk cut-off, NPV for

Table 8. Prediction-related indices for arson and violent recidivism based on the HCR-20^{V3} future violence/case prioritisation judgement ($N = 75$).

Outcome	Reoffended *n* (%)	Sensitivity	Specificity	PPV	NPV
Arson Recidivism					
Low vs. (Moderate & High)	7 (9.3)	85.7%	30.9%	11.3%	95.5%
(Low & Moderate) vs High	7 (9.3)	57.1%	79.4%	22.2%	94.7%
		42.9%	*79.4%*	*16.7%*	*93.1%*
Violent Recidivism					
Low vs (Moderate & High)	21 (28.0)	95.0%	38.9%	37.7%	95.5%
(Low & Moderate) vs High	21 (28.0)	52.4%	87.0%	61.1%	82.5%
		47.6%	*87.0%*	*55.6%*	*81.0%*

Notes: PPV = Positive Predictive Value; NPV = Negative Predictive Value. The *italicised* results show the results of robustness testing conducted to demonstrate the level of uncertainty in the results, given the small number of recidivists. For these analyses, one recidivist was reassigned from the High to Moderate risk category. There is no robustness testing for the Low vs (Moderate & High) rows because movement of one recidivist from High to Moderate risk does not change the classification of the case in this dichotomised analysis.

violent recidivism was 82.5%, representing a modest improvement over the base rate of non-recidivism (72.0%), with a greater proportion of individuals classified as low risk not going on to reoffend. Robustness testing for violence showed a notable decline in the classification and prediction metrics. While this was less pronounced than for arson, it nonetheless indicates uncertainty in the results due to the small sample size.

Time-dependent prediction indices for arson recidivism

Table 9 presents prediction and classification indices for arson recidivism based on the HCR-20^{V3} Future Violence/Case Prioritisation judgement at five annual follow-up points. At a moderate risk cut-off, the indices were consistently poor for arson across all time points. Although sensitivity and NPV were 100% for the first four years (i.e. all recidivists were classified as moderate or high risk), specificity and PPV remained low over time. For instance, between 88% and 94% of individuals classified as moderate or high risk did not subsequently reoffend.

In contrast, at a high risk cut-off, all classification and prediction indices were strong for the first two years of follow-up. In the first two years, sensitivity was between 75–100% and specificity was between 77–80%, demonstrating a good balance between identifying recidivists as high risk and non-recidivists as low or moderate risk. PPV and NPV also remained high relative to the base rate. However, these results should be interpreted with caution as robustness analyses showed a substantial reduction, particularly for sensitivity and PPV values, indicating these findings are highly sensitive to small changes in risk categorisation. Furthermore, as the follow-up time increased, predictive performance declined at both cut-off points.

Table 9. Prediction-related indices for arson recidivism based on the HCR-20^{V3} future violence/case prioritisation judgement at five time points.

Outcome/time at risk	Reoffended *n* (%)	Sensitivity	Specificity	PPV	NPV
Low vs. (Moderate and High)					
12 months (1 year)	3 (4.0)	100%	30.0%	5.6%	100%
24 months (2 years)	4 (5.3)	100%	28.6%	7.4%	100%
36 months (3 years)	5 (6.7)	100%	23.9%	9.3%	100%
48 months (4 years)	5 (6.7)	100%	23.3%	9.2%	100%
60 months (5 years)	7 (9.3)	81.0%	22.5%	12.1%	90%
(Low and Moderate) vs High					
12 months (1 year)	3 (4.0)	100%	80.0%	17.2%	100%
		66.7%	*80.0%*	*12.2%*	*98.3%*
24 months (2 years)	4 (5.3)	74.5%	76.8%	15.4%	98.1%
		49.6%	*76.8%*	*10.8%*	*96.4%*
36 months (3 years)	5 (6.7)	55.3%	76.1%	15.3%	95.6%
		36.9%	*76.1%*	*10.7%*	*93.9%*
48 months (4 years)	5 (6.7)	55.3%	74.4%	14.4%	95.5%
		36.9%	*74.4%*	*10.1%*	*93.8%*
60 months (5 years)	7 (9.3)	53.3%	77.5%	23.8%	92.6%
		41.9%	*77.5%*	*19.7%*	*91.0%*

Notes: PPV = Positive Predictive Value; NPV = Negative Predictive Value. The *italicised* results show the results of robustness testing conducted to demonstrate the level of uncertainty in the results, given the small number of recidivists. For these analyses, one recidivist was reassigned from the High to Moderate risk category. There is no robustness testing for the Low vs (Moderate & High) rows because movement of one recidivist from High to Moderate risk does not change the classification of the case in this dichotomised analysis.

Survival analyses for arson offending

Time to arson offending was between 8 and 151 months at risk ($M = 64.87$, $SD = 42.86$). Kaplan-Meier survival analysis (see Figure 1) was used to observe differences in offending patterns over time, associated with the HCR-20^{V3} Future Violence/Case Prioritisation judgement. The overall mean survival time for those with an arson charge was 90.81 months (95% CI = 82.70, 97.66), while across the Future Violence/Case Prioritisation groups, average survival times to arson charge were: low ($M = 91.75$ months), moderate ($M = 81.10$ months), and high ($M = 75.85$ months). In the first two years of follow-up, the survival patterns show clear delineation between the three groups, with those in the high risk group reoffending earlier and more frequently than those classified as low or moderate risk, aligning with the stronger discriminatory performance observed in the AUC analyses during the first two years. However, over time the distinction between the three risk groups weakens, with the low and moderate risk groups converging at approximately 60 months, indicating limited differentiation between these categories in the longer term. Consistent with these observations, the log-rank test indicated that the overall differences in survival distributions did not reach statistical significance, $\chi^2(2) = 5.364$, $p = .068$. Robustness analyses indicated that the overall trend was unchanged; however, the differences between the groups lessened, $\chi^2(2) = 2.334$, $p = .311$.

A survival analysis was also conducted for the HCR-20^{V3} Imminence judgement, given the centrality of time to reoffending for this risk construct (Supplementary Materials, Figure 1).

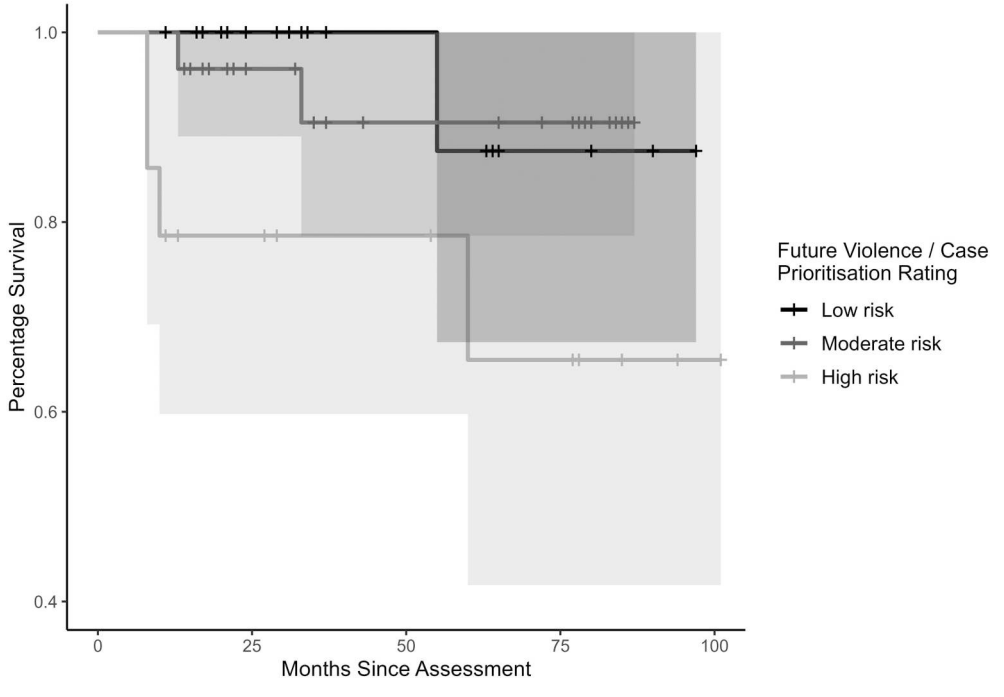


Figure 1. Kaplan-Meier survival curve for arson offending by HCR-20^{V3} future violence/case prioritisation rating.

Note: The survival curve was truncated at 8.42 years (101 months) to improve estimate stability, as fewer than 20% of the sample remained beyond this point.

Overall, the curve showed that no participants reoffended within an imminent timeframe, with the first reoffending episode occurring eight months after the assessment date. Participants in the moderate and high risk groups for Imminence appeared to reoffend more quickly compared to those in the low-risk group, with the following average survival times for each group: low ($M = 88.34$), moderate ($M = 85.67$), and high ($M = 65.75$); but differences between groups were not statistically significant, $\chi^2(2) = 1.368$, $p = .505$.

Fire-related factors

Univariate analyses of any documented history of firesetting showed that all seven recidivists had set a fire prior to the index arson, with Fisher's Exact test indicating that arson recidivists were significantly more likely to have a firesetting history than non-recidivists ($p = .017$). Using the Haldane-Anscombe correction (due to the zero cell), the estimated odds ratio was 16.38 (95% CI = 0.90–298.30). The width of the confidence interval indicates substantial imprecision arising from the small sample size. While the odds ratio suggests higher odds of arson recidivism among those with a prior firesetting history, the magnitude of this association could range from meaningless to very large and cannot be estimated with confidence.

Of the total sample, there was evidence of fire interest in 16% ($n = 12$) of cases, identified either through self- or other-report of thoughts and emotions, or possessing fire-related paraphernalia. Fewer than three firesetting recidivists reported fire interest (exact number obscured for privacy), and Fisher's Exact test revealed no significant relationship between fire interest and arson recidivism ($p = .711$).

Due to the complete separation in the data (i.e. the absence of any arson recidivists without a prior history of firesetting) and the very low number of recidivists reporting fire interest, the intended logistic regression analysis exploring the incremental predictive validity of these variables was not conducted as estimates would be unstable and unlikely to produce meaningful results.

Discussion

This study represents the first examination of the ability of the HCR-20^{V3} (Douglas et al., 2013) to predict future police-recorded firesetting (i.e. arson). Results provide preliminary evidence that in the short to medium term (approximately one to two years post-assessment) a HCR-20^{V3} Future Violence/Case Prioritisation high risk judgement has some ability to identify those at an elevated risk of arson reoffending. Due to the small sample size and low base rate of arson reoffending (9.3%), the results of this study should be interpreted with great caution and practitioners should be circumspect in their application to practice. Additionally, the findings highlight the urgent need for further research on effective tools for assessing risk of future firesetting.

Predictive validity of HCR-20^{V3} judgements for future arson compared to violence

Time-dependent predictive analyses yielded preliminary evidence that in the short to medium term (approximately one to two years post-assessment) the HCR-20^{V3} Future Violence/Case Prioritisation risk judgement may demonstrate predictive validity for future

arson, consistent across a range of discrimination and prediction metrics (Ogonah et al., 2023; Singh, 2013). In contrast, analyses extending across the total follow-up period (12.6 years), indicated that the Future Violence/Case Prioritisation judgement was not associated with future arson. When examining violent recidivism for the total follow-up period, the Future Violence/Case Prioritisation judgement demonstrated stronger, statistically significant discriminative performance, producing AUC results consistent with the wider violence literature (e.g. Brookstein et al., 2020; Doyle et al., 2014; Persson et al., 2017). This indicates that the Case Prioritisation/Future Violence judgement demonstrated comparable predictive performance for violence in this sample to that observed in previous samples.

To illustrate the predictive performance of the Case Prioritisation/Future Violence judgement for future arson over time, the survival curve (Figure 1) demonstrated clear separation between the low, moderate, and high risk groups for the first two years, indicating meaningful short to medium term risk stratification. Over time, high risk judgements remained associated with higher reoffending rates, whereas the moderate and low risk groups converged, suggesting limited differentiation between these categories. These findings are consistent with the classification metrics, which showed good sensitivity and specificity in the first two years of follow-up when the high risk group was distinguished from the low and moderate risk groups, though a lack of robustness of these results should be noted given the small number of cases with future arson.

For the total follow-up period, the absence of a high risk Future Violence/Case Prioritisation judgement was effective at ruling out those who did not reoffend, with 95% of participants classified as moderate or low risk not receiving a further arson-related charge. Conversely, participants judged as high risk were more likely to reoffend, relative to the base rate, showing meaningful stratification of high risk classifications compared to the low and moderate risk classifications. However, only one in five participants in the high risk group went on to reoffend, indicating that a substantial proportion of high risk classifications were false positives. Classification and prediction metrics for violent recidivism were comparable to those identified for arson when interpreted in the context of their respective reoffending base rates. Taken together, these findings suggest the HCR-20^{V3} Future Violence/Case Prioritisation judgement was more effective, for both arson and violence, at ruling out individuals who would not reoffend than at identifying those who would, a pattern that is consistent with the broader violence risk assessment literature (Brookstein et al., 2020; Douglas et al., 2003; Picchioni et al., 2024).

These findings provide partial support for Hypotheses 1 and 2. The HCR-20^{V3} Future Violence/Case Prioritisation judgement showed some utility for predicting future arson in the short to medium term but, as hypothesised, the overall discriminative performance was better for violent recidivism. Discrimination and classification indices were similar for both reoffending outcomes. Given that violence and arson are rare events, prediction is inherently challenging (Fazel et al., 2012; Silva, 2020), and these results must be interpreted cautiously in light of the study's methodological limitations.

Fire-related factors: firesetting history and fire interest

Data limitations meant that the intended analyses examining the incremental predictive validity of firesetting history and fire interest could not be meaningfully conducted, and Hypothesis 3 could not be tested. Univariate findings indicated that all arson recidivists in

this sample had a prior history of firesetting, consistent with previous research that indicates those with a firesetting history are significantly more likely to reoffend using fire than those without (Sambrooks et al., 2021). This highlights that firesetting history is likely important to consider in the assessment and formulation of firesetting behaviour, however, further research is required to determine the predictive utility of this variable.

Only a very small proportion of arson recidivists were assessed as having an interest in fire or fire-related paraphernalia, and no significant association with arson recidivism was observed. Previous research has yielded mixed results regarding fire interest as a correlate of repeat firesetting, depending on the data collection and measurement methods used (Dickens et al., 2009; Rice & Harris, 1991; Tyler et al., 2015), resulting in a lack of clarity surrounding the role of fire interest in repeat firesetting. Not all firesetting is driven by fire interest, and though it may be a salient factor in some cases, firesetting pathways are heterogeneous, and a range of other motivations may underlie firesetting (Gannon et al., 2012, 2022). Moreover, emerging research on fire scripts provides an explanation for why individuals may engage in firesetting in the absence of fire interest (Butler & Gannon, 2021; Gannon et al., 2023).

Future research is required to determine whether fire-related factors meaningfully improve risk assessment and case formulation beyond established general and violence risk assessment tools, using adequately powered samples and more complete and comprehensive measurement of these factors. To achieve this, routine assessment and documentation of empirically supported fire factors for clinical and research purposes is required, for instance, through the use of the Four Factor Fire Scales (Ó Ciardha, Barnoux, et al., 2015; Ó Ciardha, Tyler, et al., 2015) or the Firesetting Questionnaire (Gannon et al., 2023).

Implications

The findings of this study provide a tentative indication that the HCR-20^{V3} Future Violence/Case Prioritisation judgement may perform better than chance in identifying individuals at an increased risk of arson reoffending in the first one to two years post-assessment. In accordance with Risk-Need-Responsivity principles, individuals judged as high risk may warrant greater investment in tailored, formulation-informed risk management in the short to medium term (Bonta & Andrews, 2023). Conversely, those judged as low or moderate risk may require less intensive intervention, allowing resources to be prioritised to those classified as high risk. However, it is important to recognise that most people in the high risk group will not reoffend, and so intervention needs to reflect this, focusing on positive and supportive strategies addressing perceived causes of past firesetting, rather than restrictive strategies premised on the idea that any individual assessed as high risk is very likely to reoffend. Ongoing reassessment and dynamic risk management remain essential given the instability of predictive performance of the HCR-20^{V3} for arson over time.

The Imminence judgement demonstrated poor predictive accuracy for arson recidivism. No participants were found to reoffend imminently, which may have been due to participants being on probation or community corrections orders during this timeframe or because steps were taken to manage risk after the assessment. No arson recidivists were judged as high risk of Severity (i.e. likely reflecting the nature of referrals seen within a community-based service) and data were not available to explore indicators of severity in the follow-up offending to ascertain the validity of Severity judgements.

While the HCR-20^{V3} total score demonstrated comparable predictive validity to the Future Violence/Case Prioritisation judgement in the first two years, its performance declined markedly from the third year onwards, indicating overall poorer performance relative to the Future Violence/Case Prioritisation judgement. Future research should seek to examine the HCR-20^{V3} total score and summary risk judgements in larger samples to better understand their predictive utility for firesetting.

The tentative finding that the HCR-20^{V3} Future Violence/Case Prioritisation judgement outperformed the total score is consistent with contemporary conceptualisations of SPJ that emphasise formulation-driven integration of risk and protective factors, which may not be fully captured by total scores in isolation (Heilbrun et al., 2017). The relatively stronger performance of the Future Violence/Case Prioritisation judgement may, in part, reflect rater expertise, as the primary rater in this study had specialised firesetting knowledge, meaning their risk judgements may have incorporated factors relevant to firesetting that are not explicitly captured by the HCR-20^{V3}. Whilst such informed development of risk judgements is consistent with a SPJ approach, it is possible that results may not generalise to the wider population of practitioners without commensurate levels of firesetting expertise. Additionally, risk judgements in this study were made without formally completing the HCR-20^{V3} steps of case formulation and development of risk scenarios. While some level of informal formulation is necessary to arrive at a summary risk judgement, how adherence to the full HCR-20^{V3} protocol would affect the predictive validity of judgements is unknown. Nevertheless, these findings underscore the importance of practitioner competence when assessing risk of firesetting and caution against direct translation of violence-focused risk judgments to firesetting (Gannon et al., 2022).

Given that the current study involved a small, highly selective sample, we compared our sample to a more representative cohort of 1,052 individuals convicted of arson between 2000 and 2009 in the same jurisdiction (Ducat et al., 2015) to contextualise how our sample may differ from the broader firesetting population in an Australian context. The current sample had comparatively higher rates of mental health problems, including psychotic disorders (34.7% vs 7.1%) and personality disorders (22.7% vs 10.5%), and higher rates of prior offending, including prior violent charges (46.7% vs 37.8%) and prior arson charges (22.7% vs 6.7%). Therefore, the current sample may represent a more complex, antisocial, and/or aggressive subgroup of the firesetting population, characterised by greater psychiatric morbidity and more extensive offending histories than would be observed in a more representative firesetting sample.

While violent intent for arson was not comprehensively captured in the data, the victim of the firesetting leading to the original assessment was known to the perpetrator in 44% of cases and there was potential for the fire to cause harm (i.e. stated intent to endanger the victim or potential endangerment of life due to the location of the fire) in 47% of cases. As such, it is plausible that firesetting in this sample may have been more violent or aggressive in nature (i.e. firesetting with the intention to cause harm or instil fear in the victim) compared to the broader firesetting population (Gannon et al., 2023). In turn, the performance of the Future Violence/Case Prioritisation judgment in this sample may not reflect performance in firesetting samples predominantly characterised by non-violent intent. Nonetheless, approximately half of the sample did not display overt intent to harm, indicating the HCR-20^{V3} may retain some predictive utility for arson even in the absence of identified violent intent; however, future research is required to clarify this.

One of the ongoing challenges of firesetting risk assessment and intervention is the heterogeneity of the firesetting population, including myriad motivations for setting fires (Gannon et al., 2012). Future research should consider whether general and violent risk assessment tools, such as the HCR-20^{V3}, can effectively predict firesetting with various motivations (e.g. thrill-seeking, fire interest, financial gain, crime concealment; Gannon et al., 2012), including those driven by both violent and non-violent intent.

Limitations

The limitations of this study are highlighted throughout the paper, and readers are urged to exercise caution when interpreting the results of this small-scale preliminary study. Notably, this study is limited by its small sample size and low base rate of further arson, which reduced statistical power and increased vulnerability to bias and uncertainty in the results. The use of police charge data to measure reoffending in this study likely means that the base rate for arson recidivism is an under-estimation of the true base rate of reoffending. Furthermore, time at risk was not controlled for due to limited information on periods during which risk may have been reduced (e.g. incarceration), nor was post-assessment engagement in risk management or intervention programs. Therefore, estimates of reoffending rates and time to reoffending should be interpreted with these limitations in mind. The addition of self-report data on firesetting recidivism in future research would likely provide a more accurate estimate of reoffending (Sambrooks et al., 2021). Research in related areas of offending, such as stalking, highlights that the operationalisation of recidivism outcomes can meaningfully alter conclusions about tool performance (McEwan et al., 2020), therefore broader approaches to outcome measurement (e.g. self-report measures of reoffending) are necessary to more accurately capture firesetting behaviour, improve base rate estimations, and strengthen predictive analyses.

As commonly observed in file review studies, the quality of the data was restricted to the information available on file. There were several gaps where data was not available, or lacked standardised measurement (e.g. fire interest), leading to considerable rates of missing data for some variables that likely biased the reporting of these factors. The small sample size and large amounts of missing data meant that some of the intended analyses could not be completed. Additionally, given the sample was drawn from a forensic mental health service that specialises in the assessment and treatment of problem behaviours, the profile of the sample does not reflect the broader firesetting population, which limits the generalisability of these results. This emphasises the need for further risk assessment validation research across diverse firesetting populations and settings (e.g. custodial, psychiatric, community) to understand the consistency and generalisability of these findings.

Conclusions and future directions

Taken together, the findings of this study provide tentative preliminary support for short to medium term use of the HCR-20^{V3} Future Violence/Case Prioritisation risk judgement to identify those who may be at high risk of arson reoffending. However, caution is warranted against directly applying the HCR-20^{V3} to firesetting without informed adaptation of the tool, including understanding the nature of the firesetting and the fire-related factors associated with the behaviour.

These findings align with current recommendations that firesetting risk assessment should be conducted using a SPJ approach, guided by a firesetting-specific theory (e.g. M-TTAF; Gannon et al., 2012, 2022). Use of the HCR-20^{V3} in conjunction with the firesetting-specific M-TTAF framework encourages systematic consideration of firesetting-specific risk factors alongside the broader violence risk domains captured by the HCR-20^{V3}. In the absence of a validated firesetting-specific risk instrument, this combined approach represents a pragmatic and evidence-informed process for firesetting assessment, integrating best-practice risk assessment with contemporary firesetting research. However, findings from the present study indicate that only the high risk category on the HCR-20^{V3} Future Violence/Case Prioritisation judgement meaningfully predicted future arson in the short to medium term, and even then, only one in five people in this group reoffended. Accordingly, even when structured tools are used and informed by firesetting research, substantial limitations in the prediction and assessment of arson risk remain, and these limitations should be reflected in assessment findings and risk management recommendations.

At present, our understanding of firesetting risk assessment is poor, and significant work is needed to align firesetting research and assessment practices with other areas of offending. Future studies examining the predictive validity of general and violence risk assessment tools for firesetting, even those involving small samples, would contribute pertinent research to the evidence base and enable the eventual synthesis of findings through meta-analysis.

Acknowledgements

The authors would like to thank Dr Vindya Nanayakkara, Dr Innes Seric, and Dr Benjamin Spivak for their contributions to this paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was supported by the Commonwealth through an Australian Government Research Training Program Scholarship.

Data availability statement

Data cannot be shared due to legal and ethical requirements associated with sensitive and health data. However, descriptive and analytic output is available from the corresponding author upon request.

ORCID

Sinead Cloonan-Thomas http://orcid.org/0000-0003-2263-1076
Troy McEwan http://orcid.org/0000-0001-6935-8420
James R. P. Ogloff http://orcid.org/0000-0002-3137-5556
Nichola Tyler http://orcid.org/0000-0002-3717-1941

References

Allender, P., Brown, G., Bailey, N., Colombo, T., Helen, P., & Abril, S. (2005). *Prisoner resettlement and housing provision: A good practice ideas guide*. Coventry University.

Andrews, D. A., Bonta, J., & Wormith, S. (2004). *The Level of Service/Case Management Inventory (LS/CMI)*. Multi-Health Systems.

Andrews, D. A., Bonta, J., & Wormith, S. J. (2008). *The Level of Service/Risk Need Responsivity Inventory (LS/RNR): Scoring guide*. Multi-Health Systems.

Andrews, D. A., & Dowden, C. (2007). The risk-need-responsivity model of assessment and human service in prevention and corrections: Crime prevention jurisprudence. *Canadian Journal of Criminology and Criminal Justice*, *49*(4), 439–464. https://doi.org/10.3138/cjccj.49.4.439

Arson Prevention Forum. (2017). *State of the nation 2017*. https://www.stoparsonuk.org/arson/documents/Arson-Prevention-Forum-Booklet.pdf

Atkin, C. A., & Armstrong, G. S. (2011). Does the concentration of parolees in a community impact employer attitudes toward the hiring of ex-offenders? *Criminal Justice Policy Review*, *24*(1), 71–93. https://doi.org/10.1177/0887403411428005

Barrowcliffe, E. R., & Gannon, T. A. (2015). The characteristics of un-apprehended firesetters living in the UK community. *Psychology, Crime & Law*, *21*(9), 836–853. https://doi.org/10.1080/1068316X.2015.1054385

Barrowcliffe, E. R., & Gannon, T. A. (2016). Comparing the psychological characteristics of un-apprehended firesetters and non-firesetters living in the UK. *Psychology, Crime & Law*, *22*(4), 382–404. https://doi.org/10.1080/1068316X.2015.1111365

Barrowcliffe, E. R., Tyler, N., & Gannon, T. A. (2022). Firesetting among 18–23 year old un-apprehended adults: A UK community study. *Journal of Criminological Research, Policy and Practice*, *8*(3), 140–154. https://doi.org/10.1108/JCRPP-06-2021-0026

Bonta, J., & Andrews, D. A. (2023). *The psychology of criminal conduct* (7th ed.). Routledge.

Brookstein, D. M., Daffern, M., Ogloff, J. R. P., Campbell, R. E., & Chu, C. M. (2020). Predictive validity of the HCR-20^{V3} in a sample of Australian forensic psychiatric patients. *Psychiatry, Psychology and Law*, *28*(3), 325–342. https://doi.org/10.1080/13218719.2020.1775152

Butler, H., & Gannon, T. A. (2021). Do deliberate firesetters hold fire-related scripts and expertise? A quantitative investigation using fire service personnel as comparisons. *Psychology, Crime & Law*, *27*(4), 383–403. https://doi.org/10.1080/1068316X.2020.1808978

Campbell, R. (2021, September). *Intentional structure fires*. National Fire Protection Association. https://www.nfpa.org/News-and-Research/Data-research-and-tools/US-Fire-Problem/Intentional-fires

Challinor, A., Ogundalu, A., McIntyre, J. C., Bramwell, V., & Nathan, R. (2021). The empirical evidence base for the use of the HCR-20: A narrative review of study designs and transferability of results to clinical practice. *International Journal of Law and Psychiatry*, *78*, Article 101729. https://doi.org/10.1016/j.ijlp.2021.101729

Cloonan-Thomas, S., McEwan, T., Ducat, L., & Tyler, N. (2026). *Identifying risk and protective factors for adult firesetting within a public health framework: A narrative review* [Manuscript submitted for publication]. Centre for Forensic Behavioural Science, Swinburne University of Technology.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin*, *70*(4), 213–220. https://doi.org/10.1037/h0026256

Dickens, G., Sugarman, P., Edgar, S., Hofberg, K., Tewari, S., & Ahmad, F. (2009). Recidivism and dangerousness in arsonists. *The Journal of Forensic Psychiatry & Psychology*, *20*(5), 621–639. https://doi.org/10.1080/14789940903174006

Dolan, M., & Doyle, M. (2000). Violence risk prediction: Clinical and actuarial measures and the role of the Psychopathy Checklist. *The British Journal of Psychiatry*, *177*(4), 303–311. https://doi.org/10.1192/bjp.177.4.303

Doley, R., Fineman, K., Fritzon, K., Dolan, M., & McEwan, T. E. (2011). Risk factors for recidivistic arson in adult offenders. *Psychiatry, Psychology and Law*, *18*(3), 409–423. https://doi.org/10.1080/13218719.2011.559155

Douglas, K. S., Hart, S. D., Webster, C. D., & Belfrage, H. (2013). *HCR-20V3: Assessing risk for violence – user guide*. Mental Health, Law, and Policy Institute, Simon Fraser University.

Douglas, K. S., Ogloff, J. R., & Hart, S. D. (2003). Evaluation of a model of violence risk assessment among forensic psychiatric patients. *Psychiatric Services*, *54*(10), 1372–1379. https://doi.org/10.1176/appi.ps.54.10.1372

Douglas, K. S., & Shaffer, C. S. (2021). The science of and practice with the HCR-20 V3 (Historical-Clinical-Risk Management – 20, Version 3). In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (2nd ed., pp. 253–293). Routledge.

Douglas, T., Pugh, J., Singh, I., Savulescu, J., & Fazel, S. (2017). Risk assessment tools in criminal justice and forensic psychiatry: The need for better data. *European Psychiatry*, *42*, 134–137. https://doi.org/10.1016/j.eurpsy.2016.12.009

Doyle, M., Power, L. A., Coid, J., Kallis, C., Ullrich, S., & Shaw, J. (2014). Predicting post-discharge community violence in England and Wales using the HCR-20^{V3}. *The International Journal of Forensic Mental Health*, *13*(2), 140–147. https://doi.org/10.1080/14999013.2014.906517

Ducat, L., McEwan, T. E., & Ogloff, J. R. (2015). An investigation of firesetting recidivism: Factors related to repeat offending. *Legal and Criminological Psychology*, *20*(1), 1–18. https://doi.org/10.1111/lcrp.12052

Edwards, E. R. (2020). *Arson risk assessment: The development of an actuarial model for arson recidivism* [Unpublished doctoral dissertation]. University of Canterbury.

Edwards, M. J., & Grace, R. C. (2013). The development of an actuarial model for arson recidivism. *Psychiatry, Psychology and Law*, *21*(2), 218–230. https://doi.org/10.1080/13218719.2013.803277

Ellison, M., Fox, C., Gains, A., & Pollock, G. (2013). An evaluation of the effect of housing provision on re-offending. *Safer Communities*, *12*(1), 27–37. https://doi.org/10.1108/17578041311293125

Fazel, S., Singh, J. P., Doll, H., & Grann, M. (2012). Use of risk assessment instruments to predict violence and antisocial behaviour in 73 samples involving 24827 people: Systematic review and meta-analysis. *BMJ*, *345*(2), e4692. https://doi.org/10.1136/bmj.e4692

Field, O. H. (2016). *Risk factors for arson recidivism in adult offenders* [Unpublished doctoral dissertation]. University of Birmingham.

Gannon, T. A., Ó Ciardha, C., Doley, R., & Alleyne, E. (2012). The multi-trajectory theory of adult firesetting (M-TTAF). *Aggression and Violent Behavior*, *17*(2), 107–121. https://doi.org/10.1016/j.avb.2011.08.001

Gannon, T. A., Ó Ciardha, C., Barnoux, M. F. L., Tyler, N., Mozova, K., & Alleyne, E. K. A. (2013). Male imprisoned firesetters have different characteristics than other imprisoned offenders and require specialist treatment. *Psychiatry: Interpersonal and Biological Processes*, *76*(4), 349–364. https://doi.org/10.1521/psyc.2013.76.4.349

Gannon, T. A., Olver, M. E., Alleyne, E., Butler, H., Lister, V., Ó Ciardha, C., Sambrooks, K., & Tyler, N. (2023). The development and validation of the Firesetting Questionnaire. *Journal of Psychopathology and Behavioral Assessment*, *45*(1), 27–47. https://doi.org/10.1007/s10862-022-10011-x

Gannon, T. A., & Pina, A. (2010). Firesetting: Psychopathology, theory and treatment. *Aggression and Violent Behavior*, *15*(3), 224–238. https://doi.org/10.1016/j.avb.2010.01.001

Gannon, T. A., Tyler, N., Ó Ciardha, C., & Alleyne, E. (2022). *Adult deliberate firesetting: Theory, assessment, and treatment*. Wiley.

Gebski, V., Garès, V., Gibbs, E., & Byth, K. (2018). Data maturity and follow-up in time-to-event analyses. *International Journal of Epidemiology*, *47*(3), 850–859. https://doi.org/10.1093/ije/dyy013

Greenland, S. (1996). Basic methods for sensitivity analysis of biases. *International Journal of Epidemiology*, *25*(6), 1107–1116. https://doi.org/10.1093/ije/25.6.1107

Gwet, K. L. (2014). *Handbook of inter-rater reliability* (4th ed.). Advanced Analytics.

Hanson, R. K. (2017). Assessing the calibration of actuarial risk scales: A primer on the E/O index. *Criminal Justice and Behavior*, *44*(1), 26–39. https://doi.org/10.1177/0093854816683956

Hardcastle, L., Bartholomew, T., & Graffam, J. (2011). Legislative and community support for offender reintegration in Victoria. *Deakin Law Review*, *16*(1), 111–132. https://doi.org/10.21153/dlr2011vol16no1art96

Haslewood-Pócsik, I., Brown, S., & Spencer, J. (2008). A not so well-lit path: Employers' perspectives on employing ex-offenders. *The Howard Journal of Criminal Justice*, *47*(1), 18–30. https://doi.org/10.1111/j.1468-2311.2008.00501.x

Heilbrun, K., Yasuhara, K., Shah, S., & Locklair, B. (2017). Approaches to violence risk assessment: Overview, critical analysis, and future directions. In K. S. Douglas & R. K. Otto (Eds.), *Handbook of violence risk assessment* (pp. 29–65). Routledge.

Horsley, F. K., James, T. K., Baker, N., Broughton, R., Hampton, X., Knight, A., Langford, I., Pomfrey, E., & Unsworth, L. (2022). Exploring explicit and implicit correlates of early anti-social fire exposure and everyday fire use in adulthood. *Journal of Criminological Research, Policy and Practice*, *8*(3), 169–182. https://doi.org/10.1108/JCRPP-06-2021-0036

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Lawson, R. (2004). Small sample confidence intervals for the odds ratio. *Communications in Statistics – Simulation and Computation*, *33*(4), 1095–1113. https://doi.org/10.1081/SAC-200040691

Lindberg, N., Holi, M. M., & Virkkunen, T. M. (2005). Looking for pyromania: Characteristics of a consecutive sample of Finnish male criminals with histories of recidivist fire-setting between 1973 and 1993. *BMC Psychiatry*, *5*(1), 47. https://doi.org/10.1186/1471-244X-5-47

Logan, C. (2014). The HCR-20 Version 3: A case study in risk formulation. *International Journal of Forensic Mental Health*, *13*(2), 172–180. https://doi.org/10.1080/14999013.2014.906516

Long, C. G., Banyard, E., Fulton, B., & Hollin, C. R. (2014). Developing an assessment of fire-setting to guide treatment in secure settings: The St Andrew's Fire and Arson Risk Instrument (SAFARI). *Behavioural and Cognitive Psychotherapy*, *42*(5), 617–628. https://doi.org/10.1017/S1352465813000477

McEwan, T. E., Doley, R. M., & Dolan, M. (2012). Bushfire and wildfire arson: Arson risk assessment in the Australian context. In G. L. Dickens, P. A. Sugarman, & T. A. Gannon (Eds.), *Firesetting and mental health: Theory, research and practice* (pp. 206–223). RCPsych Publications.

McEwan, T. E., Harder, L., Brandt, C., & de Vogel, V. (2020). Risk factors for stalking recidivism in a Dutch community forensic mental health sample. *International Journal of Forensic Mental Health*, *19*(2), 127–141. https://doi.org/10.1080/14999013.2019.1661885

Nanayakkara, V., Ogloff, J. R. P., McEwan, T. E., & Davis, M. R. (2020). Applying classification methodology to high-consequence firesetting. *Psychology, Crime & Law*, *26*(7), 710–732. https://doi.org/10.1080/1068316X.2020.1733568

Nichols, T. (2022). *Public perceptions and attitudes towards people with arson convictions: A preliminary examination and comparison with sexual convictions* [Unpublished master's thesis]. Victoria University of Wellington.

Ó Ciardha, C., Alleyne, E. K. A., Tyler, N., Barnoux, M. F. L., Mozova, K., & Gannon, T. A. (2015). Examining the psychopathology of incarcerated male firesetters using the Millon Clinical Multiaxial Inventory-III. *Psychology, Crime & Law*, *21*(6), 606–616. https://doi.org/10.1080/1068316X.2015.1008478

Ó Ciardha, C., Barnoux, M. F. L., Alleyne, E. K. A., Tyler, N., Mozova, K., & Gannon, T. A. (2015). Multiple factors in the assessment of firesetters' fire interest and attitudes. *Legal and Criminological Psychology*, *20*(1), 37–47. https://doi.org/10.1111/lcrp.12065

Ó Ciardha, Caoilte, Tyler, Nichola, & Gannon, Theresa A (2015). A Practical Guide to Assessing Adult Firesetters' Fire-Specific Treatment Needs Using the Four Factor Fire Scales. *Psychiatry*, *78*(*4*), 293–304. https://doi.org/10.1080/00332747.2015.1061310

Ogonah, M. G. T., Seyedsalehi, A., Whiting, D., & Fazel, S. (2023). Violence risk assessment instruments in forensic psychiatric populations: A systematic review and meta-analysis. *The Lancet Psychiatry*, *10*(10), 780–789. https://doi.org/10.1016/S2215-0366(23)00256-0

Persson, M., Belfrage, H., Fredriksson, B., & Kristiansson, M. (2017). Violence during imprisonment, forensic psychiatric care, and probation: Correlations and predictive validity of the risk assessment instruments COVR, LSI-R, HCR-20^{V3}, and SAPROF. *International Journal of Forensic Mental Health*, *16*(2), 117–129. https://doi.org/10.1080/14999013.2016.1266420

Picchioni, M., Ruiz, R., de Girolamo, G., Iozzino, L., Zamparini, M., Wancata, J., Unger, A., Heitzman, J., Markewitz, I., Dressing, H., & Large, M. M. (2024). The predictive validity and temporal characteristics of the HCR-20v3 for inpatient violence in forensic inpatient settings. An international study. *Psychiatry Research*, *339*, Article 116079. https://doi.org/10.1016/j.psychres.2024.116079

Pocock, S. J., Clayton, T. C., & Altman, D. G. (2002). Survival plots of time-to-event outcomes in clinical trials: Good practice and pitfalls. *The Lancet*, *359*(9318), 1686–1689. https://doi.org/10.1016/S0140-6736(02)08594-X

Repo, E., & Virkkunen, M. (1997). Criminal recidivism and family histories of schizophrenic and non-schizophrenic fire setters: Comorbid alcohol dependence in schizophrenic fire setters. *Journal of the American Academy of Psychiatry and the Law*, *25*(2), 207–215.

Repo, E., Virkkunen, M., Rawlings, R., & Linnoila, M. (1997). Criminal and psychiatric histories of Finnish arsonists. *Acta Psychiatrica Scandinavica*, *95*(4), 318–323. https://doi.org/10.1111/j.1600-0447.1997.tb09638.x

Rice, M. E., & Harris, G. T. (1991). Firesetters admitted to a maximum security psychiatric institution. *Journal of Interpersonal Violence*, *6*(4), 461–475. https://doi.org/10.1177/088626091006004005

Rice, M. E., & Harris, G. T. (1996). Predicting the recidivism of mentally disordered firesetters. *Journal of Interpersonal Violence*, *11*(3), 364–375. https://doi.org/10.1177/088626096011003004

Sambrooks, K., Olver, M. E., Page, T. E., & Gannon, T. A. (2021). Firesetting reoffending: A meta-analysis. *Criminal Justice and Behavior*, *48*(11), 1634–1651. https://doi.org/10.1177/00938548211013577

Sambrooks, K., Tyler, N., & Gannon, T. A. (2024). Single versus multiple firesetting: An examination of demographic, behavioural and psychological factors. *Psychiatry, Psychology and Law*, *31*(6), 1152–1177. https://doi.org/10.1080/13218719.2023.2272904

Shepherd, S. M., Campbell, R. E., & Ogloff, J. R. P. (2017). The utility of the HCR-20 in an Australian sample of forensic psychiatric patients. *Psychiatry, Psychology and Law*, *25*(2), 273–282. https://doi.org/10.1080/13218719.2017.1364676

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, *86*(*2*), 420–428. https://doi.org/10.1037//0033-2909.86.2.420

Silva, E. (2020). The HCR-20 and violence risk assessment – will a peak of inflated expectations turn to a trough of disillusionment? *BJPsych Bulletin*, *44*(6), 269–271. https://doi.org/10.1192/bjb.2020.14

Singh, J. P. (2013). Predictive validity performance indicators in violence risk assessment: A methodological primer. *Behavioral Sciences & The Law*, *31*(1), 8–22. https://doi.org/10.1002/bsl.2052

Singh, J. P., Desmarais, S. L., Hurducas, C., Arbach-Lucioni, K., Condemarin, C., Dean, K., Doyle, M., Folino, J. O., Godoy-Cervera, V., Grann, M., Ho, R. M. Y., Large, M. M., Nielsen, L. H., Pham, T. H., Rebocho, M. F., Reeves, K. A., Rettenberger, M., de Ruiter, C., Seewald, K., & Otto, R. K. (2014). International perspectives on the practical application of violence risk assessment: A global survey of 44 countries. *International Journal of Forensic Mental Health*, *13*(3), 193–206. https://doi.org/10.1080/14999013.2014.922141

Singh, J. P., Yang, S., & Mulvey, E. P. (2015). Reporting guidance for violence risk assessment predictive validity studies: The RAGEE statement. *Law and Human Behavior*, *39*(1), 15–22. https://doi.org/10.1037/lhb0000090

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson.

Taylor, J. L., & Thorne, I. (2005). *Northgate firesetter risk assessment* [Unpublished manual]. Tyne & Wear NHS Foundation Trust.

Taylor, J. L., & Thorne, I. (2013). Pathological firesetting by adults: Assessing and managing risk within a functional analytic framework. In C. Logan & L. Johnstone (Eds.), *Managing clinical risk: A guide to effective practice* (pp. 142–164). Routledge.

Taylor, J. L., & Thorne, I. (2019). Assessing firesetters with intellectual disabilities. *Journal of Intellectual Disabilities and Offending Behaviour*, *10*(4), 102–118. https://doi.org/10.1108/JIDOB-10-2019-0020

Taylor, J. L., Thorne, I., & Slavkin, M. L. (2004). Treatment of firesetting behaviour. In W. Lindsay, J. Taylor, & P. Sturmey (Eds.), *Offenders with developmental disabilities* (pp. 221–241). Wiley.

Tyler, N., Gannon, T. A., Dickens, G. L., & Lockerbie, L. (2015). Characteristics that predict firesetting in male and female mentally disordered offenders. *Psychology, Crime & Law*, *21*(8), 776–797. https://doi.org/10.1080/1068316X.2015.1054382

Tyler, N., Gannon, T. A., Ó Ciardha, C., Ogloff, J. R., & Stadolnik, R. (2019). Deliberate firesetting: An international public health issue. *The Lancet Public Health*, *4*(8), e371–e372. https://doi.org/10.1016/S2468-2667(19)30136-7

Webster, C. D., Douglas, K. S., Eaves, D., & Hart, S. D. (1997). *HCR-20: Assessing risk for violence (version 2)*. Simon Fraser University.

Webster, C. D., Eaves, D., Douglas, K. S., & Wintrup, A. (1995). *The HCR-20 scheme: The assessment of dangerousness and risk*. Simon Fraser University and Forensic Psychiatric Services Commission of British Columbia.

Wyatt, B., Gannon, T. A., McEwan, T. E., Lockerbie, L., & O'Connor, A. (2019). Mentally disordered firesetters: An examination of risk factors. *Psychiatry*, *82*(1), 27–41. https://doi.org/10.1080/00332747.2018.1534520