

## Structured professional judgment in the new Italian forensic treatment model: the validation of the Italian version of the DUNDRUM TOOLKIT

### Il giudizio professionale strutturato nel nuovo modello trattamentale forense italiano: la validazione della versione italiana del DUNDRUM TOOLKIT

Lia Parente\* | Fulvio Carabellese\* | Eliseo Secli | Monica Rutigliano | Donatella La Tegola  
Luigi Buongiorno | Enrico Zanalda | Marco Zuffranieri | Roberto Catanesi | Gabriele Mandarelli  
Giulia Petroni | Viola Ferrante | Giuseppe Nicolò | Giuseppe Nese | Corrado Villella  
Harry G. Kennedy | Mary Davoren | Felice Carabellese\*

#### OPEN ACCESS

#### Double blind peer review

**How to cite this article:** Parente L. et al. (2025). Structured professional judgment in the new Italian forensic treatment model: the validation of the Italian version of the DUNDRUM TOOLKIT. *Rassegna Italiana di Criminologia*, XIX, 1, 048-064 <https://doi.org/10.7347/RIC-012025-p048>

**Corresponding Author:** Felice F. Carabellese email: [felicefrancesco.carabellese@uniba.it](mailto:felicefrancesco.carabellese@uniba.it)

**Copyright:** © 2025 Author(s). This is an open access, peer-reviewed article published by Pensa Multimedia and distributed under the terms of the Creative Commons Attribution 4.0 International, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. *Rassegna Italiana di Criminologia* is the official journal of Italian Society of Criminology.

© The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors

**Received:** 23.09.2024

**Accepted:** 25.01.2025

**Published:** 31.03.2025

Pensa MultiMedia

ISSN 1121-1717 (print)

ISSN 2240-8053 (on line)

[doi10.7347/RIC-012025-p048](https://doi.org/10.7347/RIC-012025-p048)

#### Abstract

**Objective:** The purpose was to demonstrate the effectiveness of the DUNDRUM Toolkit in Italy. The research focused on evaluating the disposition of forensic patients in REMS and CRAPs, hypothesizing that the correspondence between standardized assessments and juridical decisions was not that satisfying.

**Methods:** The DUNDRUM Toolkit was translated, after adapting it to national standards and tested in different regions. A total of 192 forensic patients from 9 REMS and several CRAPs in Italy were evaluated, comparing current levels of therapeutic safety with those ones considered as most appropriate. The evaluation used DUNDRUM-1 to determine pre-treatment safety needs. The progress in the treatment and recovery was examined through DUNDRUM-3 and DUNDRUM-4. Inter-rater reliability was assessed by two independent evaluators on a sample of 50 patients. All health professionals involved in the research had been previously trained in the use of the DUNDRUM Toolkit.

**Results:** Results revealed that 3.7% of patients in REMS needed more security, while 38.2% could be placed in less restrictive environments. In CRAPs, 56% of patients required a higher level of security.

**Conclusions:** The study demonstrated the validity and reliability of the Italian version of the DUNDRUM Toolkit and highlighted significant discrepancies between current forensic patient placements and needed levels of safety, suggesting the importance of using professional tools in the forensic practice that allow more accurate assessment of safety needs for each patient.

**Keywords:** forensic psychiatric treatment, DUNDRUM Toolkit, structured professional judgment tools, therapeutic use of safety, Italian forensic treatment model

#### Riassunto

**Obiettivi:** Lo scopo è stato dimostrare l'efficacia del DUNDRUM Toolkit in Italia. La ricerca si è concentrata sulla valutazione della collocazione dei pazienti forensi nelle REMS e nelle CRAP, ipotizzando che la corrispondenza tra le valutazioni standardizzate e le decisioni giudiziarie non fosse ottimale.

**Metodi:** È stato tradotto il DUNDRUM Toolkit, dopo averlo adattato alle normative nazionali e testato in diverse regioni. Sono stati valutati 192 pazienti forensi provenienti da 9 REMS e diverse CRAP in Italia, confrontando i livelli di sicurezza terapeutica attuali con quelli calcolati come più appropriati. La valutazione ha utilizzato il DUNDRUM-1 per determinare le necessità di sicurezza prima del trattamento ed i progressi nel trattamento e nel recupero sono stati esaminati attraverso il DUNDRUM-3 ed il DUNDRUM-4. L'affidabilità inter-rater è stata valutata da due valutatori indipendenti su un campione di 50 pazienti. Tutti i professionisti sanitari coinvolti nella ricerca erano stati precedentemente formati all'utilizzo del DUNDRUM Toolkit.

**Risultati:** I risultati hanno rivelato che il 3.7% dei pazienti nelle REMS necessitava di maggiore sicurezza, mentre il 38.2% poteva essere collocato in ambienti meno restrittivi. Nelle CRAP, il 56.3% dei pazienti richiedeva un livello di sicurezza superiore.

**Conclusioni:** Lo studio ha dimostrato la validità e l'affidabilità della versione italiana del DUNDRUM Toolkit e messo in evidenza delle discrepanze significative tra le attuali collocazioni dei pazienti forensi ed i livelli di sicurezza necessari, suggerendo l'importanza di utilizzare nella pratica forense strumenti di giudizio professionale strutturato che consentano una valutazione accurata delle necessità di sicurezza terapeutica per ogni paziente.

**Parole chiave:** trattamento psichiatrico forense, DUNDRUM Toolkit, strumenti di giudizio professionale strutturato, uso terapeutico della sicurezza, modello trattamentale forense italiano

Si ringraziano i dott. Alocci G, Ascolillo C, Barbon I, Bruno M, Castelletti L, Cerabolini M G, Coviello M L, Cuccurullo F, De Donatis T, D'Andrea A, Drosi F, Formichetti M, Franconi F, Guarino, D., Imperadore G, Leozappa M, Liardo R, Liuni F, Montalbò D, Ortenzi R, Padovani F, Paoletti G, Partipilo M, Pascale F, Penta E, Posteraro D E, Prati M, Rizza P, Rossetto I, Spadaro F, Stanga V, Villella C, Visca, F., Zeroli S, Zito A, e tutto il personale sanitario delle REMS del Piemonte, Lombardia, Veneto, Lazio, Toscana, Basilicata, Calabria, Campania, Puglia, Sicilia che hanno partecipato alla ricerca.

Lia Parente, Dottoranda presso l'Università La Sapienza, Roma | Fulvio Carabellese, Dottorando presso l'Università degli Studi di Bari Dipartimento | Mary Davoren, Dottoranda presso l'Università La Sapienza, Roma. DUNDRUM Centre for Forensic Excellence, Trinity College Dublin, Irlanda | Harry G. Kennedy, Trinity College di Dublino, Irlanda | Felice F. Carabellese (Ricercatore principale), Sezione di Psichiatria Forense e Criminologia, Università degli Studi di Bari Aldo Moro, Italia, Policlinico Universitario, p.za G. Cesare, 70124 | Eliseo Secli, Dottore in Psicologia Criminologica e Forense | Donatella La Tegola, Sezione di Criminologia e Psichiatria Forense, Università degli Studi di Bari Aldo Moro, Policlinico Universitario | Viola Ferrante, TRP CRAP Dedicata Puglia | Luigi Buongiorno, Assegnista di Ricerca Università degli Studi di Bari Aldo Moro | Enrico Zanalda, Presidente SIFP | Marco Zuffranieri, Psicologo ASL TO | Gabriele Mandarelli, Prof. Associato Psicopatologia Forense Università degli Studi di Bari Aldo Moro | Giulia Petroni, Dottoranda di Ricerca Università Roma Sapienza | Giuseppe Nicolò, Direttore DSM ASL Roma 5 | Corrado Villella, REMS ASL Roma 5 | Giuseppe Nese, Direttore SMOP Campania | Roberto Catanesi, Prof. Ordinario Psicopatologia Forense Università degli Studi di Bari Aldo Moro

## Structured professional judgment in the new italian forensic treatment model: the validation of the italian version of the DUNDRUM TOOLKIT

### Introduzione

Negli ultimi anni l'Italia ha intrapreso una radicale riorganizzazione del sistema di assistenza psichiatrica forense, abbandonando il modello asilare rappresentato dai sei Ospedali Psichiatrici Giudiziari (OPG), attivati nel '900 e chiusi definitivamente nel 2015. È stato adottato, anche per i pazienti psichiatrici autori di reato, un modello di approccio terapeutico comunitario e riabilitativo, attivo in Italia dal 1978.

Parallelamente, ognuna delle 20 regioni italiane si è organizzata disponendo l'apertura delle Residenze per l'Esecuzione delle Misure di Sicurezza (REMS). Queste strutture, esclusivamente sanitarie e con un massimo di 20 posti letto (Carabellese, 2017; Castelletti et al., 2018; Catanesi et al., 2019), sono coordinate direttamente dai Dipartimenti di Salute Mentale (DSM) e mirano a garantire un trattamento adeguato ai pazienti forensi.

Alcune regioni inoltre hanno sviluppato strutture comunitarie forensi più piccole, dedicate a pazienti forensi in libertà vigilata, strutture che generalmente corrispondono ad un livello di sicurezza inferiore rispetto alle REMS; queste ultime in linea di massima hanno standards di sicurezza terapeutica paragonabili a quelli di medio livello di altri paesi (Kennedy, 2002).

Tuttavia, la diversificazione delle strutture forensi e delle misure di sicurezza psichiatriche ha sollevato la necessità di valutare accuratamente il livello di sicurezza terapeutica richiesto per ogni paziente forense, se inserirlo in una REMS, in una comunità forense a bassa sicurezza (CRAP) o decidere che possa rimanere al proprio domicilio o in qualche altro ambiente con o senza una misura di sicurezza (Carabellese & Carabellese, 2021).

La valutazione accurata del livello di sicurezza terapeutica necessario è un passaggio fondamentale per garantire un trattamento efficace ai pazienti con disturbi mentali che hanno commesso reati. Questo processo deve rispettare i diritti dei pazienti, assicurando che eventuali misure restrittive siano applicate nel contesto più adeguato, in base al livello di rischio che rappresentano. Le misure devono essere temporanee e proporzionate alle necessità terapeutiche e alla prevenzione della recidiva, tenendo conto anche delle esigenze di sicurezza del contesto in cui si trovano. È difatti necessario che i pazienti con disturbi mentali, sottoposti a misure di sicurezza, non mettano a rischio il personale sanitario. Questo significa che il rischio che presentano deve essere gestito in modo tale da non superare la capacità della struttura di garantire un ambiente di lavoro sicuro.

L'assegnazione di un paziente ad un livello di sicurezza

inappropriato può compromettere infatti il successo del trattamento forense ed aumentare i rischi per la sicurezza (Jeandarme et al., 2021).

Il DUNDRUM Toolkit (Kennedy et al., 2016) nasce nel Central Mental Hospital di Dublino in Irlanda, frutto della riflessione dei professionisti sanitari che vi lavoravano, ed ha permesso di assegnare il livello di sicurezza più appropriato ed il trattamento necessario ai pazienti forensi.

È uno strumento finalizzato al cosiddetto *Structured Professional Judgment* (SPJ), giudizio professionale strutturato, utilizzato in altri paesi per pianificare la "legacy", per programmare cioè i bisogni di salute mentale e poi pianificarli (Shaw et al., 1994; Harty et al., 2004; Pierzchniak et al., 1999; O'Neill et al., 2003; Coid e Kahtan, 2000). Gli strumenti di SPJ sono stati utilizzati anche per valutare l'efficacia dei servizi psichiatrici (Coid e Kahtan, 2000; Pillay et al., 2008; O'Dwyer et al., 2011; McCullough et al., 2020).

Il DUNDRUM Toolkit, in particolare, combina indicatori statici di valutazione delle necessità di sicurezza terapeutica (DUNDRUM-1) ed indicatori dinamici di valutazione di efficacia del trattamento in sicurezza terapeutica (completamento del programma terapeutico DUNDRUM-3) e di recovery del trattamento forense (DUNDRUM-4) (Davoren et al., 2012; Eckert et al., 2017; Adams et al., 2018).

Lo strumento è stato utilizzato anche per valutare il livello di sicurezza terapeutica da adottare in carcere (Flynn et al., 2011; O'Neill et al., 2016) in unità di media sicurezza inglesi (Freestone et al., 2015) ed alta sicurezza (Williams et al., 2020), così come in altri paesi (Jeandarme et al., 2019; Habets et al., 2019; Adams et al., 2018; Jones et al., 2019). In particolare, il DUNDRUM Toolkit è stato validato fra le buone pratiche terapeutiche di sicurezza in Irlanda (O'Neill et al., 2016; Davoren et al., 2012; Davoren et al., 2013; Davoren et al., 2015), nel Regno Unito (Freestone et al., 2013; Williams et al., 2020; McCullough et al., 2020), nel Belgio (Jeandarme et al., 2019; Habets et al., 2019; Habets et al., 2020), in Australia (Adams et al., 2018), in Nuova Zelanda (Wharewera-Mika et al., 2020; Jewell et al., 2023), in Canada (Jones et al., 2019; Lam et al., 2022) ed ora anche in Italia (Carabellese et al., 2022; Parente et al., 2023).

Inoltre, nel non valutare correttamente il livello di sicurezza terapeutica più adeguato al gradiente di "pericolosità" che il paziente forense pone, si espone il paziente stesso al rischio di mancato completamento del trattamento (Jeandarme et al., 2021; Lawrence et al., 2018), oltre che mettere in pericolo gli altri pazienti, il personale sanitario e lo stesso paziente.

È stato dimostrato infatti che collocare un paziente forense in un ambiente a maggiore livello di sicurezza terapeutica rispetto a quella di cui ha bisogno può ritardare le risposte al trattamento, ritardare il recupero e portare a periodi di degenza in struttura impropriamente prolungati (Williams et al., 2020; Sharma et al., 2015; Davoren et al., 2013; Eckert et al., 2017); collocare invece un paziente in un ambiente terapeuticamente meno sicuro del necessario può portare al mancato completamento del trattamento, ad una durata della degenza inferiore alle necessità che il paziente pone (Jeandarme et al., 2021) e ad un maggiore rischio di comportamenti dirompenti, di contenzione e di isolamento (Jewell et al., 2023).

Il DUNDRUM Toolkit, nei suoi diversi strumenti di cui si compone (-1, -2, -3 e -4), si ripromette di rispondere efficacemente a tutte queste diverse esigenze e problematiche.

Nel contesto internazionale, l'utilizzo di strumenti strutturati di giudizio professionale, ed in particolare l'utilizzo del DUNDRUM Toolkit, si è dimostrato efficace per guidare le decisioni di collocamento appropriato in sicurezza del malato di mente autore di reato e per valutare l'adeguatezza e l'efficacia dei servizi forensi. Tuttavia, l'applicabilità, la validità e l'efficacia di tali strumenti nel contesto italiano post-riforma non sono state ancora pienamente esplorate.

Il presente studio si propone di valutare l'applicabilità della versione italiana del DUNDRUM Toolkit nel modello trattamentale forense italiano e, contestualmente, valutare l'efficacia dello strumento nell'apprezzare l'appropriatezza della collocazione di autori di reato con disturbi mentali nelle REMS ed in altre strutture forensi italiane a più bassa sicurezza (CRAP). L'obiettivo a lungo termine è fornire una valutazione più obiettiva della corrispondenza tra il livello di sicurezza necessario e quello effettivamente fornito nel modello trattamentale forense italiano, contribuendo così allo sviluppo di politiche e di pratiche più valide ed appropriate.

Abbiamo ipotizzato che, come avviene in altri paesi, la corrispondenza tra una valutazione frutto di SPJ e la collocazione iniziale del paziente in struttura forense, poteva non coincidere.

## Materiali e Metodi

La ricerca è stata approvata dal Comitato Etico dell'Azienda Ospedaliero-Universitaria di Bari (N. 66510/AA. GG del 16.09.2020).

Tutti i pazienti arruolati, informati della finalità della ricerca, hanno fornito il loro consenso alla ricerca in forma scritta. Sono stati intervistati da personale sanitario esperto, con accesso anche ai dati clinici, anamnestici e giudiziari dei pazienti, appositamente formato all'uso degli strumenti di ricerca. Tutti i dati, anonimizzati, sono stati inseriti in file Excel e poi SPSS-28 (IBM Corp., 2021).

## Traduzione in italiano del DUNDRUM Toolkit

La traduzione italiana, redatta per la prima volta da Carabellese & Carabellese, è stata verificata in inglese con l'Autore dello strumento in ogni sessione, arricchita dei suggerimenti dei sanitari coinvolti nella ricerca durante gli anni in cui si è proceduto con la validazione nelle varie regioni, fino a raggiungere una traduzione soddisfacente, utile per il sistema italiano.

In particolare, il DUNDRUM Toolkit è stato tradotto in italiano da dottori bilingue formati per la sua applicazione dal Prof. Kennedy presso il Central Mental Hospital di Dublino negli anni 2018 e 2019.

## Disegno dello studio

Si tratta di uno studio trasversale su 192 pazienti forensi. Sono stati utilizzati: il DUNDRUM-1 per valutare il livello iniziale di sicurezza terapeutica necessario, il DUNDRUM-2 per il triage d'urgenza per l'ammissione di pazienti in lista di attesa ed il DUNDRUM-3 e -4 per valutare i progressi e l'efficacia del trattamento forense in struttura e la recovery forense ottenuta.

L'affidabilità inter-rater è stata valutata su un campione di circa 50 pazienti.

Al momento della ricerca, tra novembre 2020 e maggio 2022, erano attive 31 REMS su tutto il territorio nazionale, che fornivano cure e trattamenti di media sicurezza a 604 malati di mente autori di reato in misura di sicurezza psichiatrica detentiva definitiva o provvisoria. Oltre a questi pazienti forensi, c'era un numero imprecisato di pazienti con misure di sicurezza psichiatrica non detentive collocate in residenze comunitarie di diversa natura. Alcune regioni italiane, come la Puglia, avevano da tempo individuato specificamente residenze comunitarie forensi (CRAP dedicate, a bassa sicurezza). In altre regioni le residenze comunitarie accettavano pazienti forensi e non forensi.

## Partecipanti

L'invito a partecipare a questo progetto di ricerca è stato formalmente rivolto a tutte le REMS italiane. Non tutte le REMS hanno risposto alla richiesta e quindi i criteri di selezione utilizzati per l'arruolamento dei pazienti sono stati (i) geografici: presenza nelle regioni del nord (Lombardia, Piemonte, Veneto), centro (Lazio, Toscana), sud (Basilicata, Calabria, Campania, Puglia) e isole (Sicilia); (ii) disponibilità a partecipare; (iii) disponibilità ad essere formati gratuitamente all'utilizzo del DUNDRUM Toolkit.

In particolare è stato reclutato un campione trasversale composto da 192 pazienti forensi provenienti da diversi contesti: 9 REMS (a media sicurezza), diverse CRAP dedicate (a bassa sicurezza) e da comunità non forensi ma che accoglievano anche pazienti forensi.



Fanno parte del campione pazienti con vizio totale (ex art. 88 c.p.) o parziale di mente (ex art. 89 c.p.), socialmente pericolosi (ex art. 203), che hanno dato il loro consenso scritto a partecipare alla ricerca. Per la valutazione ICC si è provveduto ad esaminare 50 pazienti provenienti da REMS del nord, del centro e del sud Italia, compresi alcuni pazienti provenienti da CRAP dedicate.

### Lo Strumento di SPJ

Il DUNDRUM Toolkit è un insieme di quattro diversi strumenti di SPJ (1 = valutazione del triage di sicurezza; 2 = triage d'urgenza; 3 = completamento del trattamento; 4 = recupero forense), che comprende anche uno strumento di autovalutazione.

Il DUNDRUM-1 nella sua forma a nove item è stato utilizzato per valutare il livello di sicurezza terapeutica più appropriato di cui necessitava il paziente forense al momento dell'invio da parte della competente Autorità Giudiziaria.

Il DUNDRUM-1 fa riferimento ad indicatori *lifetime* di bisogno di sicurezza terapeutica relativamente statici. Ogni elemento è valutato su una scala da 0 a 4, dove un punteggio di "0" indica che non c'è bisogno di sicurezza terapeutica, "1" indica che il paziente potrebbe essere gestito in sicurezza anche ambulatorialmente o soggetto a supervisione, "2" indica la necessità di collocamento in una struttura a bassa sicurezza, "3" indica la necessità di un collocamento in una struttura a media sicurezza (REMS) e "4" indica la necessità di collocamento in un'unità ad alta sicurezza. Questa graduazione nella valutazione è il particolare vantaggio di questo strumento. Ai fini della valutazione della collocazione più appropriata, i nove item sono sommati e divisi per 9 per ottenere un punteggio medio compreso tra 0 e 4, dove un punteggio medio DUNDRUM-1 compreso tra 0 e 0,9 suggerisce al più la necessità di supervisione o prescrizioni più lievi eventualmente in misura di sicurezza non detentiva; da 1,0 a 1,9 indica la necessità di prescrizioni più stringenti anche tramite collocamento in una struttura forense a bassa sicurezza e comunque sia in misura di sicurezza non detentiva; da 2,0 a 2,9 indica la necessità di collocamento in una struttura a media sicurezza come la REMS; da 3,0 a 4 indica la necessità di un posizionamento più sicuro ovvero, nel modello trattamentale forense italiano, in REMS dotate di livelli di sicurezza più stringenti.

Gli item del triage di urgenza del DUNDRUM-2 sono dedicati a fornire un supporto per decidere chi, all'interno delle liste d'attesa per l'ammissione ad un dato livello di sicurezza, ha più urgenza di essere ammesso, al di là dell'ordine cronologico di inserimento nella lista. In generale, un punteggio alto indica un'urgenza maggiore di trattamento forense. Si tratta di uno strumento di SPJ particolarmente utile in paesi come il nostro in cui, in diverse regioni, vi sono liste di attesa in cui i pazienti forensi, prima di poter accedere nelle REMS cui sono stati collocati, attendono mediamente un anno, solitamente seguendo un mero criterio cronologico.

La scala di completamento del programma DUNDRUM-3 è una scala e si compone di sette item che valuta i progressi nelle principali aree di trattamento rilevanti per i pazienti affetti da disturbi mentali che commettono reati, rappresentati da: cura della salute fisica, cura della salute mentale, possibilità di accesso a sostanze di abuso, rischio di condotte violente, cura di sé, coinvolgimento in attività di vita quotidiana, interesse per istruzione, occupazione ed attività creative, relazioni con la famiglia di origine e sentimentali. Sommando i punteggi per i sette item e dividendo per sette si ottiene un punteggio da 0 a 4, dove "4" indica che il paziente non è ancora pronto a spostarsi da un luogo ad alta sicurezza ad un luogo a sicurezza inferiore; "3" indica la possibilità di collocazione in una struttura a media sicurezza; "2" indica la possibilità di collocazione ad un livello di sicurezza più bassa in struttura forense o anche non forense; "1" indica la possibilità di collocazione in strutture o luoghi comunitari a bassa sicurezza o anche di trattamento ambulatoriale con prescrizioni più lievi e "0" indica che il paziente non ha più bisogno di trattamento forense in sicurezza.

La scala di recupero forense DUNDRUM-4 è uno strumento di sette item che valuta soprattutto l'opportunità della dimissione da una struttura forense con o senza condizionalità, ovvero il trasferimento da una struttura a più alta sicurezza ad una a più bassa sicurezza. Per questo, è necessario utilizzarlo sempre insieme al DUNDRUM-3. Gli item sono rappresentati dalla stabilità raggiunta dal paziente, dal suo insight, dal rapporto terapeutico, dall'alleanza di lavoro stabilita con i sanitari, dalla verifica dei permessi, dagli eventuali rischi dinamici legati anche al contesto in cui potrebbe essere collocato il paziente dopo la dimissione, dalle eventuali problematiche legate alla eventualità che la vittima si relazioni ancora con il paziente e dalla progettualità futura del paziente. Ogni elemento è valutato allo stesso modo del DUNDRUM-3 con lo stesso significato di unità significative ed unità di cambiamento.

### Posizionamento più sicuro

Al fine di confrontare il livello di sicurezza terapeutica del paziente al momento dell'arruolamento nel campione con il suo posizionamento più appropriato, abbiamo confrontato la valutazione del DUNDRUM-1 al momento dell'ingresso in struttura forense, con i progressi compiuti nel trattamento (DUNDRUM-3) e nel recupero forense (DUNDRUM-4), utilizzando a tal fine un algoritmo. Ad esempio, se un paziente in REMS otteneva un punteggio medio al DUNDRUM-1 al momento dell'ingresso di 2,5 (nell'intervallo da 2,0 a 2,9), ciò era indicativo della necessità di un livello medio di sicurezza (REMS); se la valutazione attuale dello stesso paziente dava al DUNDRUM-3 un punteggio medio di 1,5 (nell'intervallo da 1,0 a 1,9), cioè possibilità di collocamento ad un livello di sicurezza più basso ed al DUNDRUM-4 un punteggio medio di 0,8 (nell'intervallo da 0 a 0,9), era indicativo della disponibilità di dimissione con prescrizioni

blande, il livello di necessità di sicurezza terapeutica più adeguato sarebbe stato considerato quello indicato dal punteggio più prudente fra i due (1,5 = bassa sicurezza).

### Affidabilità inter-rater ed analisi statistica

Al fine di valutare il livello di affidabilità dello strumento, per ogni scala, sono stati considerati 50 pazienti esaminati da due valutatori indipendenti (A e B). In particolare, è stato utilizzato il Coefficiente di Correlazione Intraclassa (ICC), metodo statistico robusto che misura il grado di concordanza tra valutatori diversi sullo stesso set di soggetti esaminati.

Per ogni scala DUNDRUM sono riportati i Coefficienti di Correlazione Intraclassa sia per le misure singole che per le misure medie. La coerenza interna è stata misurata attraverso l'Alpha di Cronbach, che valuta quanto strettamente un set di item sia correlato come gruppo. Per ogni analisi ICC, sono stati calcolati gli intervalli di confidenza al 95% che danno un'indicazione della precisione della stima dell'affidabilità e per ogni ICC è condotto anche un test F con il relativo valore p, per verificare la significatività statistica. Per ogni scala, sono fornite le statistiche descrittive (media e deviazione standard) dei punteggi assegnati dai due valutatori. L'analisi item-totale è stata condotta per valutare il contributo di ciascun item alla scala complessiva. Per ogni coefficiente di correlazione intraclassa (ICC) è stato riportato anche un test F con il relativo valore p, che risulta, per tutte le scale DUNDRUM Toolkit (DUNDRUM-1, DUNDRUM-2,

DUNDRUM-3, DUNDRUM-4), sempre altamente significativo ( $p < 0,001$ ). La concordanza tra l'attuale livello effettivo di sicurezza terapeutica ed il livello di sicurezza terapeutica più adeguato è stata testata con il coefficiente di correlazione intraclassa (ICC).

Questi calcoli sono stati ripetuti per ciascuna scala del DUNDRUM Toolkit al fine di ottenere una valutazione completa dell'affidabilità inter-rater nella sua versione italiana.

In assenza di uno studio pilota per la popolazione italiana di pazienti forensi si è provveduto a calcolare per la traduzione italiana del DUNDRUM Toolkit l'Indice di Variazione di Affidabilità (RCI) (Evans et al., 1998), che è risultato soddisfacente.

### Risultati

La traduzione italiana del DUNDRUM Toolkit ha mostrato buona coerenza interna, misurata utilizzando l'Alpha di Cronbach: per DUNDRUM-1 Alpha=0,843, per DUNDRUM-3 Alpha=0,746 e per DUNDRUM-4 Alpha=0,842. L'Indice di Variazione di Affidabilità (RCI) per DUNDRUM-1 è stato di 0,77, per DUNDRUM-3 è stato di 0,89 e per DUNDRUM-4 è stato di 0,79. Lo studio ha inoltre dimostrato un Indice di Variazione Affidabile inferiore (RCI) ad un'unità di variazione media per le sotto scale DUNDRUM-1 (RCI=0,77), DUNDRUM-3 (RCI=0,89) e DUNDRUM-4 (RCI=0,79) ed un'eccellente affidabilità inter-rater per tutte le sezioni del DUNDRUM, come di seguito riportato.

**Tabella 1. Dati validazione DUNDRUM-1**

Riepilogo dell'elaborazione dei casi

		N	%
Casi	Valido	50	53.8
	Escluso <sup>a</sup>	43	46.2
	Totale	93	100.0

a. Eliminazione in senso orario in base a tutte le variabili della procedura.

Statistiche degli item

	Media	Deviazione std.	N
DUNDRUM_1_A9	2.4067	.65508	50
DUNDRUM_1_B9	2.5489	.68545	50

Coefficiente di correlazione intraclassa

	Correlazione intraclassa <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			Sig
		Limite inferiore	Limite superiore	Valore	df1	df2	
Misure singole	.824 <sup>a</sup>	.709	.896	10.362	49	49	<.001
Misure medie	.903 <sup>c</sup>	.830	.945	10.362	49	49	<.001

Modello a effetti misti a due vie in cui gli effetti delle persone sono casuali e gli effetti delle misure sono fissi.

a. Lo stimatore è lo stesso, indipendentemente dalla presenza o meno dell'effetto di interazione.

b. Coefficienti di correlazione intraclassa di tipo C utilizzando una definizione di coerenza. La varianza tra le misure è esclusa dalla varianza del denominatore.

c. Questa stima viene calcolata assumendo che l'effetto di interazione sia assente, perché altrimenti non è stimabile.

## Statistiche di affidabilità

Alpha di Cronbach	N di item
.811	9

## Statistiche totali degli item

	Media della scala se l'item viene eliminato	Varianza della scala se l'item viene eliminato	Correzione della correlazione totale degli item	Alpha di Cronbach se l'item viene eliminato
TS1	19.0723	27.312	.329	.813
TS3	19.3253	23.515	.617	.778
TS5	19.5904	22.074	.722	.761
TS6	19.6386	26.234	.411	.804
TS7	19.4096	25.903	.438	.801
TS8	19.3373	25.177	.621	.781
TS9	19.3373	26.177	.531	.791
TS10	20.1325	26.116	.373	.811
TS11	18.6145	25.167	.552	.787

## Coefficiente di correlazione intraclasse

	Correlazione intraclasse <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			
		Limite inferiore	Limite superiore	Valore	df1	df2	Sig
Misure singole	.323 <sup>a</sup>	.244	.420	5.302	82	656	<.001
Misure medie	.811 <sup>c</sup>	.744	.867	5.302	82	656	<.001

Per quanto riguarda il DUNDRUM-1 Triage di sicurezza, su un totale di 93 casi, 50 (53.8%) sono risultati validi per l'analisi, mentre 43 (46.2%) sono stati esclusi. L'esclusione, così come per le successive scale DUNDRUM-2, DUNDRUM-3 e DUNDRUM-4, è stata fatta in base alla procedura di "listwise deletion", che elimina i casi con dati mancanti in qualsiasi variabile dell'analisi.

Il punteggio medio del DUNDRUM-1 per il primo valutatore (A) è stato di 2,4067, mentre per il secondo valutatore (B) è stato leggermente più alto, 2,5489. Questa

piccola differenza (0,1422) suggerisce che lo strumento è abbastanza stabile tra diverse valutazioni.

L'affidabilità inter-rater, misurata attraverso il Coefficiente di Correlazione Intraclasse (ICC), ha mostrato risultati eccellenti. L'ICC per le misure singole è stato di 0,824, con un intervallo di confidenza al 95% tra 0,709 e 0,896. Per le misure medie, l'ICC è stato ancora più alto, pari a 0,903 (con un IC al 95% tra 0,830 e 0,945). Questi valori indicano un'ottima concordanza tra i valutatori nell'utilizzo di questa scala.

**Tabella 2. Dati validazione DUNDRUM-2**

Riepilogo dell'elaborazione dei casi

		N	%
Casi	Valido	50	53.8
	Escluso <sup>a</sup>	43	46.2
	Totale	93	100.0

*a. Eliminazione in senso orario in base a tutte le variabili della procedura.*

## Statistiche degli item

	Media	Deviazione std.	N
DUNDRUM_2_A	12.3600	5.07398	50
DUNDRUM_2_B	12.1400	3.98983	50

## Coefficiente di correlazione intraclasse

	Correlazione intraclasse <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			Sig
		Limite inferiore	Limite superiore	Valore	df1	df2	
Misure singole	.700 <sup>a</sup>	.526	.818	5.665	49	49	<.001
Misure medie	.823 <sup>c</sup>	.689	.900	5.665	49	49	<.001

*Modello a effetti misti a due vie in cui gli effetti delle persone sono casuali e gli effetti delle misure sono fissi.*

*a. Lo stimatore è lo stesso, indipendentemente dalla presenza o meno dell'effetto di interazione.*

*b. Coefficienti di correlazione intraclasse di tipo C utilizzando una definizione di coerenza. La varianza tra le misure è esclusa dalla varianza del denominatore.*

*c. Questa stima viene calcolata assumendo che l'effetto di interazione sia assente, perché altrimenti non è stimabile.*

#### Statistiche di affidabilità

Alpha di Cronbach	N di item
.805	6

#### Statistiche totali degli item

	Media della scala se l'item viene eliminato	Varianza della scala se l'item viene eliminato	Correzione della correlazione totale degli item	Alpha di Cronbach se item viene eliminato
TI1	10.3833	14.647	.658	.752
TU2	10.8333	16.887	.502	.788
TU3	12.7167	17.190	.495	.790
TU4	12.3333	13.887	.564	.778
TU5	10.2833	14.274	.698	.742
TU6	10.1167	14.647	.515	.789

#### Coefficiente di correlazione intraclasse

	Correlazione intraclasse <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			Sig
		Limite inferiore	Limite superiore	Valore	df1	df2	
Misure singole	.407 <sup>a</sup>	.296	.532	5.116	59	295	<.001
Misure medie	.805 <sup>c</sup>	.717	.872	5.116	59	295	<.001

Per il DUNDRUM-2 Triage d'urgenza, su un totale di 93 casi, 50 (53,8%) sono risultati validi per l'analisi, mentre 43 (46,2%) sono stati esclusi (come per il DUNDRUM-1).

Il punteggio medio per il primo valutatore (A) è stato di 12,3600, mentre per il secondo valutatore (B) è stato leggermente inferiore, 12,1400. Anche in questo caso, la differenza è minima (0,22), indicando una buona concordanza.

L'affidabilità inter-rater è risultata affidabile. L'ICC per le misure singole è stato di 0,700 (IC 95%: 0,526-0,818),

mentre per le misure medie è salito a 0,823 (IC 95%: 0,689-0,900). Ciò suggerisce una solida concordanza tra i valutatori, sebbene leggermente inferiore rispetto al DUNDRUM-1. L'Alpha di Cronbach è di 0,805 per i 6 item, valore che indica una buona coerenza interna. Nessun item, se rimosso, aumenterebbe sostanzialmente tale valore. Ciò suggerisce che tutti gli item contribuiscono positivamente alla coerenza interna della scala.

**Tabella 3. Dati validazione DUNDRUM-3**

Riepilogo dell'elaborazione dei casi

		N	%
Casi	Valido	53	57.0
	Escluso <sup>a</sup>	40	43.0
	Totale	93	100.0

*a. Eliminazione in senso orario in base a tutte le variabili della procedura.*

Statistiche degli item

	Media	Deviazione std.	N
DUNDRUM_3_A	2.3747	.79748	53
DUNDRUM_3_B	2.5714	.77236	53

Coefficiente di correlazione intraclassa

	Correlazione intraclassa <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			
		Limite inferiore	Limite superiore	Valore	df1	df2	Sig
Misure singole	.712 <sup>a</sup>	.549	.823	5.944	52	52	<.001
Misure medie	.832 <sup>c</sup>	.709	.903	5.944	52	52	<.001

*Modello a effetti misti a due vie in cui gli effetti delle persone sono casuali e gli effetti delle misure sono fissi.**a. Lo stimatore è lo stesso, indipendentemente dalla presenza o meno dell'effetto di interazione.**b. Coefficienti di correlazione intraclassa di tipo C utilizzando una definizione di coerenza. La varianza tra le misure è esclusa dalla varianza del denominatore.**c. Questa stima viene calcolata assumendo che l'effetto di interazione sia assente, perché altrimenti non è stimabile*

Statistiche di affidabilità

Alpha di Cronbach	N di articoli
.818	7

Statistiche totali degli item

	Media della scala se l'item viene eliminato	Varianza della scala se l'item viene eliminato	Correzione della correlazione totale degli item	Alpha di Cronbach se l'item viene eliminato
P1	14.5132	23.426	.697	.773
P2	14.0658	25.529	.487	.805
P3	15.5789	25.047	.278	.860
P4	14.4079	22.351	.695	.769
P5	14.9342	24.009	.641	.782
P6	14.9211	22.767	.767	.761
P7	14.7632	24.876	.513	.801

Coefficiente di correlazione intraclassa

	Correlazione intraclassa <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			
		Limite inferiore	Limite superiore	Valore	df1	df2	Sig
Misure singole	.391 <sup>a</sup>	.298	.499	5.501	75	450	<.001
Misure medie	.818 <sup>c</sup>	.748	.874	5.501	75	450	<.001

Per il DUNDRUM-3 Completamento del programma di trattamento, su un totale di 93 casi, 53 (57%) sono risultati validi per l'analisi, mentre 40 (43%) sono stati esclusi.

Il punteggio medio del DUNDRUM-3 per il primo valutatore (A) è stato di 2,3747, mentre per il secondo valutatore (B) è stato leggermente più alto, 2,5714. Anche questa differenza, relativamente piccola (0,1967), suggerisce una buona concordanza tra i valutatori.

L'affidabilità inter-rater è risultata nuovamente buona.

L'ICC per le misure singole è stato di 0,712 (IC 95%: 0,549-0,823), per le misure medie di 0,832 (IC 95%: 0,709-0,903). Questi valori indicano una solida concordanza tra i valutatori nell'utilizzo di questa scala.

L'Alpha di Cronbach è 0,818 per i 7 item, indicando ancora una buona coerenza interna.

Anche in questo caso, nessun item, se rimosso, aumenterebbe sostanzialmente l'Alpha di Cronbach oltre 0,818. Ciò suggerisce che tutti gli item contribuiscono positivamente alla coerenza interna della scala e che la struttura



attuale dello strumento è robusta. L'unica eccezione potrebbe essere per l'item P3, la cui rimozione porterebbe

l'Alpha a 0,860, indicando che potrebbe essere l'item meno coerente con il resto della scala.

**Tabella 4. Dati validazione DUNDRUM-4**

Riepilogo dell'elaborazione dei casi

		N	%
Casi	Valido	53	57.0
	Escluso <sup>a</sup>	40	43.0
	Totale	93	100.0

*a. Eliminazione in senso orario in base a tutte le variabili della procedura.*

Statistiche degli item

	Media	Deviazione std.	N
DUNDRUM_4_A	2.2588	.62210	53
DUNDRUM_4_B	2.3962	.61166	53

Coefficiente di correlazione intraclasse

	Correlazione intraclasse <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			
		Limite inferiore	Limite superiore	Valore	df1	df2	Sig
Misure singole	.717 <sup>a</sup>	.556	.827	6.079	52	52	<.001
Misure medie	.836 <sup>c</sup>	.715	.905	6.079	52	52	<.001

*Modello a effetti misti a due vie in cui gli effetti delle persone sono casuali e gli effetti delle misure sono fissi.*

*a. Lo stimatore è lo stesso, indipendentemente dalla presenza o meno dell'effetto di interazione.*

*b. Coefficienti di correlazione intraclasse di tipo C utilizzando una definizione di coerenza. La varianza tra le misure è esclusa dalla varianza del denominatore.*

*c. Questa stima viene calcolata assumendo che l'effetto di interazione sia assente, perché altrimenti non è stimabile*

Statistiche di affidabilità

Alpha di Cronbach	N di item
.777	7

Statistiche totali degli item

	Media della scala se l'item viene eliminato	Varianza della scala se l'item viene eliminato	Correzione della correlazione totale degli item	Alpha di Cronbach se l'item viene eliminato
R1	13.8750	12.618	.635	.719
R2	13.1944	14.807	.616	.727
R3	13.6250	15.562	.560	.740
R4	13.5278	16.872	.401	.767
R5	13.5972	13.850	.678	.710
R6	13.8056	14.356	.473	.757
R7	14.2917	18.097	.172	.801

Coefficiente di correlazione intraclasse

	Correlazione intraclasse <sup>b</sup>	Intervallo di confidenza al 95%		F Test con valore vero 0			
		Limite inferiore	Limite superiore	Valore	df1	df2	Sig
Misure singole	.332 <sup>a</sup>	.239	.442	4.475	71	426	<.001
Misure medie	.777 <sup>c</sup>	.688	.847	4.475	71	426	<.001

Per il DUNDRUM-4 Recupero forense, su un totale di 93 casi, 53 (57%) sono stati considerati validi per l'analisi, mentre 40 (43%) sono stati esclusi (come per il DUNDRUM-3). Il punteggio medio del DUNDRUM-4 per il primo valutatore (A) è stato di 2,2588, mentre per il secondo valutatore (B) è stato leggermente più alto,

2,3962. La differenza di 0,1374 è relativamente piccola e suggerisce una buona concordanza tra i valutatori.

L'ICC per le misure singole è 0,717 (IC 95%: 0,556-0,827), mentre per le misure medie è 0,836 (IC 95%: 0,715-0,905). Questi valori sono buoni e indicano una solida affidabilità inter-rater.

L'Alpha di Cronbach per il DUNDRUM-4 è 0,777, basato su 7 item. Questo valore indica un'accettabile consistenza interna della scala, anche se leggermente inferiore alle altre scale.

In questa scala, l'item R7 ha una correlazione item-totale bassa (0,172) e la sua rimozione aumenterebbe leggermente l'Alpha di Cronbach a 0,801. Tale risultato potrebbe richiedere una revisione o una riconsiderazione dell'item, data la sua bassa correlazione con il totale della scala. Gli altri item mostrano correlazioni item-totali buone e moderate e nessun altro item, se rimosso, aumenterebbe l'Alpha di Cronbach oltre 0,777.

I risultati mostrano un'eccellente affidabilità inter-rater per tutte le sezioni del DUNDRUM. Secondo le linee guida comunemente accettate, valori ICC superiori a 0,75 ed inferiori all'unità sono considerati indicativi di un'eccellente affidabilità. I risultati ottenuti, con valori ICC per misure medie che variano da 0,823 a 0,903, suggeriscono una forte concordanza tra i valutatori nell'applicazione dello strumento DUNDRUM Toolkit 1, 2, 3, 4 nel contesto italiano. L'utilizzo dell'ICC in questo studio è particolarmente appropriato e fornisce una stima dell'affidabilità sia per misure singole che per misure medie, permettendo una valutazione più completa dell'affidabilità dello strumento. Le misure medie mostrano per tutte e 4 le scale DUNDRUM sempre un'affidabilità superiore rispetto alle misure singole.

La coerenza interna, misurata attraverso l'Alpha di Cronbach, è risultata buona o accettabile per tutte le scale, con valori tutti superiori a 0,75, che vanno da 0,777 a 0,818, con tutti i risultati che si sono rilevati statisticamente significativi ( $p < 0.001$ ).

Questi risultati indicano che gli item all'interno di ciascuna scala misurano in modo coerente lo stesso costrutto e suggeriscono una buona coerenza interna per tutte le scale.

È interessante notare che il DUNDRUM-1 (Triage Security) ha mostrato i valori più alti sia per l'ICC (0.903) che per l'Alpha di Cronbach (0.811), suggerendo che que-

sta scala potrebbe essere particolarmente robusta nella versione italiana sia in fase di valutazione della pericolosità sociale, sia in fase di indicazione della misura di sicurezza più appropriata in fase di revisione della pericolosità sociale. Il DUNDRUM-4 (Forensic Recovery), pur mostrando una buona affidabilità inter-rater, ha l'Alpha di Cronbach più basso (0.777), suggerendo che potrebbe beneficiare di ulteriori perfezionamenti per migliorare la sua coerenza interna. È tuttavia possibile che i risultati ottenuti abbiano anche risentito della permanenza inappropriata dei pazienti in REMS o in altra struttura forense.

In particolare il DUNDRUM-1 mostra un'eccellente affidabilità inter-rater ed una buona coerenza interna, il DUNDRUM-2 e il DUNDRUM-3 mostrano una buona affidabilità inter-rater e coerenza interna, il DUNDRUM-4 mostra una buona affidabilità inter-rater e una coerenza interna accettabile, suggerendo che è uno strumento generalmente affidabile, seppur con valori leggermente inferiori alle altre scale, suscettibile di ulteriori miglioramenti.

Inoltre, per tutte le scale del DUNDRUM (-1, -2, -3 e -4), sia per ICC (Coefficiente di Correlazione Intraclassa) che per l'Alpha di Cronbach, i risultati sono statisticamente significativi ( $p < 0.001$ ). Il valore p misura la probabilità che i risultati osservati si verifichino per caso. Il valore ottenuto indica che la probabilità che i risultati osservati si siano verificati per caso è inferiore allo 0.1%. In questo contesto significa che c'è una probabilità estremamente bassa che le correlazioni e le consistenze interne osservate per le 4 scale DUNDRUM siano dovute al caso. Questa forte significatività statistica aumenta la fiducia nella validità e affidabilità della versione italiana del DUNDRUM Toolkit, suggerendo che le correlazioni e le consistenze interne osservate riflettono reali proprietà dello strumento piuttosto che fluttuazioni casuali.

Questi risultati nel loro complesso supportano fortemente la validità e l'affidabilità della versione italiana del DUNDRUM Toolkit, suggerendo che può essere utilizzato con fiducia nella pratica clinica e nella ricerca in ambito forense in Italia.

**Tabella 5. Panoramica delle caratteristiche demografiche e della distribuzione dei pazienti tra le 9 strutture a media sicurezza (REMS) e le 7 residenze comunitarie a bassa sicurezza (CRAP)**

Caratteristica	Totale	Media Sicurezza (REMS)	Bassa Sicurezza (CRAP)
Numero pazienti	192	137	55
Uomini	178 (92.7%)	131 (95.6%)	47 (85.5%)
Donne	14 (7.3%)	6 (4.4%)	8 (14.5%)
Età media (anni)	43,5 (S.D.11,5)	-	-
Durata ricovero (mesi)	14,9 (S.D.10,2)	-	-

Il campione dello studio comprendeva 192 pazienti forensi, suddivisi in due gruppi principali in base al livello di sicurezza della struttura in cui erano inseriti. In particolare, 137 pazienti erano ospitati in 9 REMS (il 22.7% di tutti i residenti in REMS, IC 95% 19.4% - 26.2%), mentre 55 si trovavano in 7 diverse residenze comunitarie (CRAP) a bassa sicurezza. La maggior parte dei partecipanti era di sesso maschile (178 uomini), con solo 14 donne, che rappresentano il 7.3% del totale, in linea con la prevalenza emersa in studio precedente (Catanesi et al., 2019).

L'età media dei pazienti era di circa 43,5 anni (S.D. 11,5, con un intervallo che va dai 18 agli 85 anni, IC 95% 41.9% - 45.2%).

Per quanto riguarda il tempo trascorso dall'ammissione in struttura forense, la durata media era di 14,9 mesi (S.D. 10,2).

Gli anni medi di istruzione erano 9,6 (S.D. 3,1). L'età, gli anni di istruzione ed il tempo trascorso dall'ammissione non differivano in modo significativo tra maschi e femmine o tra REMS e CRAP. La percentuale di donne è più alta nelle strutture a bassa sicurezza (14.5%) rispetto a quelle a media sicurezza (4.4%). Su un campione di 174 pazienti era effettuata comparazione delle diagnosi tra i due tipi di strutture (REMS e CRAP).

**Tabella 6. Distribuzione delle Diagnosi per 174 pazienti tra due tipi di strutture forensi in Italia: le Residenze per l'Esecuzione delle Misure di Sicurezza (REMS) a media sicurezza e le strutture forensi a bassa sicurezza (CRAP)**

Diagnosi	REMS (%)	CRAP (%)
Schizofrenia	48.4	36.5
Solo disturbo di personalità	18.9	11.5
Disturbo schizoaffective	11.5	9.6
Disturbo delirante	5.7	9.6
Disturbo bipolare	5.7	13.5
Disabilità intellettiva	3.3	15.4
Altro	6.5	3.9

**Legenda:**  $X^2 = 14.1$ ,  $df=6$ ,  $p=0.028$

Per quanto riguarda le REMS, sono stati esaminati 122 pazienti. La diagnosi più comune in REMS è stata la schizofrenia, riscontrata nel 48.4% dei casi. Seguono i disturbi di personalità nel 18.9% dei casi e l'11.5% con disturbi schizoaffective. Meno frequenti sono i disturbi deliranti (5.7%), i disturbi bipolari (5.7%) e le disabilità intellettive (3.3%).

Anche nelle strutture forensi a basso livello di sicurezza (CRAP), la schizofrenia è stata la diagnosi più frequente, ma con una percentuale inferiore rispetto alle REMS, attestandosi al 36.5%. Si nota una distribuzione più equilibrata tra le altre diagnosi: i disturbi bipolari e le disabilità intellettive sono rispettivamente al 13.5% ed al 15.4%, seguiti dai disturbi schizoaffective e deliranti, entrambi al 9.6%. I casi con solo disturbo di personalità rappresentano l'11.5%.

È interessante notare che la distribuzione delle diagnosi differisce significativamente tra i due tipi di strutture, come indicato dal test chi-quadrato ( $X^2=14.1$ ,  $df=6$ ,  $p=0.028$ ): per quanto riguarda le diagnosi psichiatriche, la schizofrenia è risultata la più comune; a seguire i più frequenti sono stati i disturbi di personalità.

La disabilità intellettiva come diagnosi primaria era più frequente nelle strutture a bassa sicurezza rispetto a quelle a media sicurezza. Le REMS tendono ad avere una maggiore concentrazione di pazienti schizofrenici, mentre le strutture a basso livello di sicurezza presentano una distribuzione più variegata di diagnosi, con percentuali relativamente più alte di disturbi bipolari e disabilità intellettive.

## Necessità di sicurezza terapeutica

I punteggi medi di sicurezza del triage DUNDRUM-1 non differivano tra unità di sicurezza medie (REMS) ed unità di sicurezza basse (CRAP) (REMS n=137, punteggio DUNDRUM-1 medio di nove item 2,21, SD 0,65, IC 95% da 2,09 a 2,32, intervallo da 0,44 a 3,78; CRAP n=49, 2,07, SD 1,09, IC 95% da 1,86 a 2,29, range da 0,36 a 3,78, ANOVA F=1,4, df=1, p=0,242). Sebbene i pazienti in strutture a bassa sicurezza tendessero ad avere

punteggi DUNDRUM più bassi rispetto a quelli in REMS, l'intervallo era molto ampio e si sovrapponeva.

Le tre successive tabelle rappresentano rispettivamente: il punteggio del DUNDRUM-1 al momento dell'ingresso in struttura, la necessità di sicurezza terapeutica DUNDRUM-1 alla luce dei punteggi del DUNDRUM-3 e la necessità di sicurezza terapeutica DUNDRUM-1 alla luce dei punteggi del DUNDRUM-4.

**Tabella 7. Punteggio di sicurezza del triage DUNDRUM-1 del posizionamento più sicuro al momento dell'affidamento**

Livello di sicurezza	Media sicurezza (n=137)	Bassa sicurezza (n=49)
Elevata	18 (13.1%) IC 95% [7.5-18.8]	7 (14.3%) IC 95% [5.9-27.2]
Media	72 (52.6%) IC 95% [43.9-61.1]	22 (44.9%) IC 95% [30.6-59.8]
Bassa	44 (32.1%) IC 95% [24.4-40.6]	17 (34.7%) IC 95% [21.7-49.6]
Aperto	3 (2.2%) IC 95% [0.4-6.3]	3 (6.1%) IC 95% [1.3-16.9]

**Tabella 8. Necessità di sicurezza terapeutica DUNDRUM-1 compensata dai progressi nel trattamento DUNDRUM-3**

Livello di sicurezza	Media sicurezza (n=136)	Bassa sicurezza (n=48)
Elevata	3 (2.2%) IC 95% [0.5-6.3]	4 (8.3%) IC 95% [2.3-19.9]
Media	73 (53.7%) IC 95% [44.9-62.3]	23 (47.9%) IC 95% [33.3-62.8]
Bassa	56 (41.2%) IC 95% [32.8-49.9]	18 (37.5%) IC 95% [23.9-52.7]
Aperto	4 (2.9%) IC 95% [0.8-7.4]	3 (6.3%) IC 95% [1.3-16.9]

**Tabella 9. Necessità di sicurezza terapeutica DUNDRUM-1 compensata dai progressi nel recupero forense DUNDRUM-4**

Livello di sicurezza	Media sicurezza (n=131)	Bassa sicurezza (n=48)
Elevata	5 (3.8%) IC 95% [1.3-8.7]	4 (8.3%) IC 95% [2.3-19.9]
Media	65 (49.6%) IC 95% [38.9-56.2]	22 (45.8%) IC 95% [31.4-60.8]
Bassa	53 (40.5%) IC 95% [31.9-49.4]	18 (37.5%) IC 95% [23.9-52.7]
Aperto	8 (6.1%) IC 95% [2.7-11.7]	4 (8.3 %) IC 95% [2.3-8.7]

In base ai criteri DUNDRUM-1 di triage al momento dell'ammissione, il 13.1% dei pazienti (18) in REMS a media sicurezza ed il 14.3% (7) in CRAP a bassa sicurezza avrebbero dovuto iniziare il ricovero più appropriatamente in una struttura ad alta sicurezza. Il 52.6% dei pazienti (72) in media sicurezza e il 44.9% (22) in bassa sicurezza avrebbero dovuto iniziare il loro trattamento in una struttura forense a media sicurezza, mentre il 32.1% (44) in media sicurezza e il 34.7% (7) in bassa sicurezza avrebbero dovuto iniziarlo in una a bassa sicurezza. Il 2.2% dei pazienti (3) in media sicurezza e il 6.1% (3) in bassa sicurezza si stima non avessero necessità di alcuna struttura forense ed avrebbero al più avuto necessità di prescrizioni trattamentali più leggere.

In merito ai progressi nel trattamento (DUNDRUM-3) e nel recupero forense (DUNDRUM-4), la corrispondenza tra la collocazione attuale ed il bisogno di sicurezza terapeutica valutato dai due strumenti (DUNDRUM-3 e -4) è risultato meno discordante. Per i pazienti che si trovavano in media sicurezza, solo tra il 2.2% (3 pazienti) e il

3.8% (5 pazienti) necessitava di una struttura a più alta sicurezza, circa la stessa percentuale è assegnata correttamente alla media sicurezza e circa il 38% avrebbe potuto transitare ad un livello di sicurezza inferiore. Per coloro che si trovavano in bassa sicurezza, l'8.3% sembra avere ancora bisogno di alta sicurezza, il 45.8% ha bisogno di media sicurezza e il 37.5% sembra essere correttamente collocato in bassa sicurezza. Sebbene solo un paziente fosse sottoposto a libertà vigilata con affidamento ambulatoriale al DSM, il punteggio del triage DUNDRUM-1 indicava che 6 pazienti (3.1%) collocati in strutture forensi avrebbero potuto iniziare il loro trattamento con medesime modalità e livelli di sicurezza bassi, mentre il punteggio al DUNDRUM-3 indicava che 8 pazienti (4.1%) erano pronti per un collocamento analogo; il recupero forense (DUNDRUM-4) indicava che un numero maggiore di pazienti (13 = 6.7%) erano pronti per un collocamento analogo.

La tabella che segue mostra il valore nominale del posizionamento di corrente più sicuro rispetto al posizionamento di corrente effettivo.

**Tabella 10. Posizionamenti di corrente effettivi rispetto al posizionamento di corrente più sicuro (livello DUNDRUM-1 compensato dal livello attuale DUNDRUM-3 o DUNDRUM-4).**

Collocamento attuale	Collocamento più sicuro	Totale			
	Aperto	Bassa sicurezza	Media sicurezza	Alta sicurezza	
Media sicurezza (REMS)	3 (2.2%) IC95% [0.4-4.7]	49 (36%) IC95% [27.9-44.7]	79 (58.1%) IC95% [49.3-66.5]	5 (3.7%) IC95% [1.2-8.4]	136
Bassa sicurezza (CRAP)	3 (6.3%) IC95% [1.3-17.2]	18 (37.5%) IC95% [23.9-52.7]	23 (47.9%) IC95% [33.3-62.8]	4 (8.3%) IC95% [2.3-19.9]	48
Ambienti in comunità non sicuri (CFS)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	1

**Legenda:**  $X^2=4.9$ ,  $df=6$ ,  $p=0.563$ ; Misura del coefficiente di correlazione intraclassa (ICC)= -0.039,  $F=0.962$ ,  $df=184$ ,  $p=0.914$ . **NOTA:** non c'erano luoghi attuali ad alta sicurezza.

È stata riscontrata una scarsa concordanza tra il posizionamento corrente e quello più appropriato come valutato dal DUNDRUM Toolkit ( $X^2=4.9$ ,  $df=6$ ,  $p=0.563$ ; misura dell'accordo del coefficiente di correlazione intraclassa ICC= -0.039,  $F=0.962$ ,  $df=184$ ,  $p=0.914$ ). Sebbene non fossero disponibili posti letto in strutture ad alta sicurezza, il campione a media sicurezza comprendeva 5 pazienti (il 3.7% di tutti i pazienti a media sicurezza valutati) ed il campione a bassa sicurezza ne comprendeva 4 (l'8.3% del totale dei pazienti a bassa sicurezza valutati), pazienti che avrebbero avuto in realtà necessità di un collocamento a più alta sicurezza. Dei pazienti inseriti in media sicurezza, 79 pazienti (il 58.1% dei pazienti in REMS valutati) sono stati valutati come correttamente collocati, 23 pazienti in-

seriti in strutture a bassa sicurezza sono stati valutati come bisognosi di un livello più alto di sicurezza (il 47.9% dei pazienti a bassa sicurezza valutati), 49 pazienti (il 36% di tutti i pazienti in REMS valutati) sono stati valutati come collocabili in strutture a più bassa sicurezza. Dei 48 pazienti in unità di bassa sicurezza, 18 (il 37.5% di quelli valutati in unità a bassa sicurezza) sono stati valutati come correttamente collocati; 3 pazienti attualmente in unità a media sicurezza e altri 3 in unità a bassa sicurezza avrebbero potuto essere dimessi al più con blande prescrizioni.

In merito all'analisi della collocazione più appropriatezza dei pazienti forensi, il totale dei casi analizzati è stato di 184 pazienti (di cui 136 delle REMS a media sicurezza e 48 delle CRAP a basso livello di sicurezza).



**Tabella 11. Posizionamento attuale, a media sicurezza (REMS) o comunità a bassa sicurezza (CRAP) rispetto all'appropriatezza del posizionamento di corrente più sicuro valutato. Attualmente in media sicurezza Vs attualmente in bassa sicurezza**

Posizionamento attuale	Posizionamento più alto che sicuro	Posizionamento appropriato	Posizionamento più basso che sicuro	Totale
Media sicurezza (REMS)	52 (38.2%) IC 95% [30.0-46.9]	79 (58.1%) IC 95% [49.3-66.5]	5 (3.7%) IC 95% [1.2-8.4]	136
Bassa sicurezza (CRAP)	3 (6.3%) IC 95% [1.3-17.2]	18 (37.5%) [23.9-51.2]	27 (56.3%) IC95% [41.2-70.5]	48

**Legenda:**  $X^2 = 71.4$ ,  $df=2$ ,  $p<0.001$ 

La maggior parte dei pazienti in REMS a media sicurezza (79 pazienti su 136, pari al 58.1% del totale) si trovava collocato nella “posizione appropriata”, mentre 52 pazienti, pari al 38.2%, poteva essere appropriatamente collocato in un ambiente meno restrittivo. 5 pazienti, pari al 3.7% del totale, avevano probabilmente bisogno di una collocazione più sicura.

Per quanto riguarda il basso livello di sicurezza (CRAP), la maggior parte dei pazienti (27 su 48, pari al 56.3% del totale) si trovava sotto posizionato, avrebbero dovuto cioè essere collocati in una struttura forense terapeuticamente più sicura. 18 pazienti (37.5%) si trovavano collocati nella “posizione appropriata”, mentre 3 pazienti (6.3%) avrebbero necessitato al più di prescrizioni più leggere senza alcuna collocazione in strutture forensi. I pazienti collocati in REMS avrebbero beneficiato, più appropriatamente, di una collocazione terapeutica con minore livello di sicurezza, mentre i pazienti attualmente in unità a bassa sicurezza (CRAP) necessitavano di un livello più elevato di sicurezza terapeutica.

L'uso del DUNDRUM Toolkit ha indicato che sia pazienti collocati in REMS che in CRAP avrebbero necessitato di un livello di sicurezza terapeutica alto (inesistente in Italia), ovvero quantomeno in REMS con un più alto profilo di sicurezza, tuttavia statisticamente i pazienti in REMS avevano più probabilità di essere collocati in maniera più appropriata rispetto a quelli collocati in strutture a bassa sicurezza ( $X^2 = 71.4$ ,  $df = 2$ ,  $p < 0.001$ ).

## Discussione

Questo studio fornisce la prima validazione completa della versione italiana del DUNDRUM Toolkit, strumento di SPJ che dimostra di possedere proprietà psicometriche robuste, buona coerenza interna ed un indice di cambiamento affidabile per tutte le scale di cui si compone: DUNDRUM-1, DUNDRUM-2, DUNDRUM-3 e DUNDRUM-4.

I risultati rivelano alcune discrepanze significative tra il livello di sicurezza terapeutica necessario valutato col

DUNDRUM 1, 3 e 4 e quello effettivamente fornito nel sistema trattamentale forense italiano del post-riforma ex Legge 81/2014.

È emerso infatti che il 3.7% dei pazienti collocati nelle REMS avrebbe necessitato di un livello di sicurezza più elevato, mentre ben il 38.2% poteva essere appropriatamente collocato in un ambiente meno restrittivo. Il 56% dei pazienti forensi collocati nelle CRAP (unità a bassa sicurezza), avrebbe necessitato di un livello di sicurezza terapeutica superiore, mentre il 6% poteva essere gestito anche con minori restrizioni.

Clinicamente significativa è risultata poi essere l'individuazione di una piccola ma critica percentuale di pazienti nelle REMS (3.7% di 604 totali, IC 95% 1.2%-8.4%) che, a distanza di tempo dal loro inserimento in struttura forense, avrebbero richiesto un livello di sicurezza più elevato, attualmente non disponibile in Italia; come pure l'8.3% (IC 95% 2.3%-19.9%) dei pazienti forensi collocati in strutture a bassa sicurezza avrebbe avuto necessità di un collocamento a più alta sicurezza.

Il 13.1% (IC 95% dal 7.5% al 18.8%) dei pazienti collocati inizialmente in REMS ed il 14.3% (IC 95% dal 5.9% al 27.2%) dei pazienti collocati nelle CRAP, avrebbe richiesto inizialmente, in maniera più appropriata, una più elevata sicurezza al momento dell'ingresso in struttura forense, ingresso che, come è noto, può avvenire anche a distanza di molto tempo dal giudizio di pericolosità sociale. Pertanto, oltre alla questione del giudizio di pericolosità sociale da parte degli esperti del Giudice, si pone la necessità di una valutazione unica e regionale del livello di pericolosità sociale di tutti i pazienti forensi giudicati socialmente pericolosi, precedentemente al loro ingresso nella struttura forense più appropriata. Da questo punto di vista il DUNDRUM Toolkit si è dimostrato essere particolarmente utile ed efficace. Anche in corso d'opera, nel senso che il DUNDRUM, nelle sue scale 3 e 4, combinate alla 1, fornisce ai sanitari che gestiscono ed assistono il paziente collocato nella struttura forense, una valutazione appropriata, obiettiva e verificabile di eventuali trasferimenti del paziente verso livelli più alti o più bassi di sicurezza terapeutica, ovvero del momento più opportuno di

procedere con la sua dimissione con o senza condizionalità post-dimissione.

Valutazione che, per di più, può essere condivisa da tutti i protagonisti, sanitari o giuristi che siano, del trattamento in senso lato del paziente forense.

L'uso del DUNDRUM-1 per stimare non solo la necessità di sicurezza terapeutica al momento dell'inserimento in struttura forense ma anche per valutare la necessità di trasferimenti in strutture forensi a più elevato o a più basso livello di sicurezza, è stato documentato altrove ed è stato recentemente utilizzato in un modello nazionale di assistenza privo di alta sicurezza, quale quello italiano (Jewell et al., 2023). In tal senso, l'uso del DUNDRUM-1, abbinato al DUNDRUM-3 e DUNDRUM-4 per valutare il giudizio sull'appropriatezza del livello di sicurezza alla luce dei progressi nel trattamento e nella recovery è ben consolidato (O'Dwyer et al., 2011; Eckert et al., 2017; Adams et al., 2018; McCullough et al., 2020; Davoren et al., 2022; Lam et al., 2022) ed ora, applicato in Italia, attraverso questa ricerca che lo ha utilizzato per la prima volta, ha fornito anche nel modello trattamentale forense proiettato verso il recupero riabilitativo del paziente forense, risultati incoraggianti ed utili.

La ricerca dimostra infatti che il 58.1% di coloro che si trovavano in REMS e il 37.5% di quelli collocati in strutture forensi a bassa sicurezza vi si trovavano in modo appropriato. Tuttavia, un numero piccolo ma clinicamente significativo di pazienti forensi era stato collocato in un livello di sicurezza terapeutica inferiore a quello che sembrava dover essere quello più appropriato per le esigenze di sicurezza che la loro condizione esprimeva.

Più precisamente 5 pazienti di quelli in REMS (3.7% del totale) e 4 (8.3%) di quelli collocati in CRAP necessitavano di un'alta sicurezza, percentuale minima di pazienti forensi che tuttavia solleva la questione di un rischio più elevato per gli stessi pazienti e per il personale sanitario.

Si stima pertanto, alla luce dei dati emersi, la necessità in Italia di un certo numero, molto limitato in realtà, di posti letto "ad alta sicurezza".

Per converso, ben il 36% di coloro che erano sistemati in REMS potevano essere collocati, alla luce dei punteggi del DUNDRUM-1, 3 e 4 in strutture forensi a bassa sicurezza, contribuendo in maniera non irrilevante a snellire le procedure di ingresso in REMS.

I risultati di questo studio evidenziano pertanto la necessità di un modello trattamentale forense più dinamico e con una stratificazione più accurata dei livelli di sicurezza, in cui le decisioni in merito al collocamento più appropriato siano basate sulle evidenze e supportate da strumenti di valutazione adeguati, in ossequio ai canoni del SPJ. Il DUNDRUM Toolkit da questo punto di vista si dimostra particolarmente utili nel modello trattamentale forense italiano.

L'implementazione di tale approccio potrebbe migliorare significativamente gli esiti clinici, ridurre i rischi e ottimizzare la allocazione delle risorse.

Il DUNDRUM Toolkit, anche da questo punto di

vista si è dimostrato essere particolarmente efficace per il modello trattamentale italiano.

Il nostro studio presenta alcuni limiti, tuttavia la robustezza metodologica e la rappresentatività geografica del campione conferiscono validità ai risultati e certamente può fornire evidenze cruciali per la pratica clinica forense italiana e per informare adeguatamente le scelte di politica sanitaria, suggerendo la necessità di riconsiderare l'attuale struttura del sistema forense, in particolare la necessità dell'introduzione di piccole unità ad alta sicurezza distribuite sul territorio nazionale ed il miglioramento dei processi decisionali per il collocamento dei pazienti attraverso centri unici, regionali, con sanitari formati all'utilizzo di strumenti appropriati di SPJ. Il DUNDRUM Toolkit, per la sua completezza sarebbe senz'altro strumento utile.

## Conclusioni

La maggior parte dei pazienti è collocata in modo appropriato e sicuro. Tuttavia, una percentuale significativa di pazienti forensi richiederebbe livelli di sicurezza diversi da quelli della loro attuale collocazione.

I risultati suggeriscono la necessità, a livello nazionale, di circa 50 posti letto in più ad alta sicurezza.

La mancanza di concordanza tra la necessità di sicurezza terapeutica valutata e l'effettivo collocamento dei pazienti nelle strutture forensi suggerisce l'importanza di implementare un processo decisionale guidato da valutazioni strutturate del bisogno di sicurezza terapeutica (Carabellese, 2017; Kennedy, 2021).

Migliorare la collocazione del paziente forense in livelli di sicurezza terapeutica appropriati, in un sistema più flessibile e stratificato di livelli di sicurezza che prevedano alta, media e bassa sicurezza potrebbe incrementare significativamente la sicurezza dell'ambiente terapeutico per i pazienti stessi, per gli altri pazienti e per il personale sanitario. La collocazione più appropriata dei pazienti forensi sottoposti a misure di sicurezza psichiatrica è necessaria anche per l'efficacia del trattamento e per questioni etiche, legate alle restrizioni delle libertà imposte dall'applicazione di misure di sicurezza psichiatrica.

L'adozione di strumenti di giudizio professionale strutturati come il DUNDRUM Toolkit in Italia rappresenterebbe un passo importante verso questo obiettivo.

L'utilizzo di tali strumenti decisionali è una parte centrale dell'esercizio delle competenze cliniche e della governance nei servizi forensi di tutto il mondo, per valutazioni più informate e trasparenti (Glancy et al., 2021; Lam et al., 2022). La disponibilità di strumenti di giudizio professionale strutturati in lingua italiana consente tale sviluppo, finora piuttosto trascurato nella pratica forense del nostro Paese.

Per migliorare i modelli di cura forensi appare pertanto cruciale utilizzare sistematicamente strumenti di supporto al SPJ, rafforzare la collaborazione con l'Università e sviluppare percorsi di cura integrati (Kennedy et al., 2019; McLaughlin et al., 2023). L'implementazione di questo

approccio potrebbe portare a migliori esiti clinici, a ridurre il gravoso problema delle liste di attesa, alla riduzione dei rischi ed all'ottimizzazione dell'allocazione delle risorse nel sistema trattamentale forense italiano.

Sono necessarie ulteriori ricerche per modellare le probabili ulteriori esigenze nelle REMS a media sicurezza e nelle residenze comunitarie a bassa sicurezza.

La traduzione italiana del DUNDRUM Toolkit permette di valutare l'attuale funzionamento del modello di cura per la psichiatria forense a seguito delle progressive riforme del 2014.

## Riferimenti Bibliografici

- Adams, J., Thomas, S. D. M., Mackinnon, T., & Eggleton, D. (2018). The risks, needs and stages of recovery of a complete forensic patient cohort in an Australian state. *BMC psychiatry*, 18(1), 35.
- Carabellese, F. (2017). Closing OPG: Socially dangerous mentally ill offenders' diagnostic tools. From forensic-psychiatric evaluation to the treatment. *Rassegna italiana di criminologia*, 11, 173-181.
- Carabellese, Fu., Carabellese, Fe., (2021). Six Years After the Closure of the OPGs and the Establishment of REMS: Considerations and Future Perspectives. *Psychology and Behavioral Science International Journal*, 2021, 16(3).
- Carabellese, Fe., Parente, L., La Tegola, D., Rossetto, I., Franconi, F., Zanalda, E., Mandarelli, G., Catanesi, R., Kennedy, H.G., Carabellese, Fu. (2022). Il DUNDRUM Toolkit, versione italiana e il suo potenziale utilizzo nel modello trattamentale forense italiano. *Rassegna italiana di criminologia* XVI, 4, 271-282.
- Castelletti, L., Scarpa, F. & Carabellese, F. (2018). Treating not guilty by reason of insanity and socially dangerous subjects by community psychiatric services: an Italian perspective. *Rassegna italiana di criminologia*, 12, 182-189.
- Catanesi, R., Mandarelli, G., Ferracuti, S., Valerio, A. & Carabellese, F. (2019). The new Italian residential forensic psychiatric system (REMS). A one-year population study. *Rassegna italiana di criminologia*, numero speciale, 7-23.
- Coid, J. & Kahtan, N. (2000). Uno strumento per misurare le esigenze di sicurezza dei pazienti in media sicurezza. *British Journal of Psychiatry*, 11, 119-134.
- Davoren, M., O'Dwyer, S., Abidin, Z., Naughton, L., Gibbons, O., Doyle, E., McDonnell, K., Monks, S., & Kennedy, H. G. (2012). Prospective in-patient cohort study of moves between levels of therapeutic security: the DUNDRUM-1 triage security, DUNDRUM-3 programme completion and DUNDRUM-4 recovery scales and the HCR-20. *BMC psychiatry*, 12, 80.
- Davoren, M., Abidin, Z., Naughton, L., Gibbons, O., Nulty, A., Wright, B., & Kennedy, H.G. (2013). Prospective study of factors influencing conditional discharge from a forensic hospital: the DUNDRUM-3 programme completion and DUNDRUM-4 recovery structured professional judgement instruments and risk. *BMC Psychiatry*, 13, 185 - 185.
- Davoren, M., Hennessy, S., Conway, C., Marrinan, S., Gill, P. & Kennedy, H. G. (2015). Recovery and Concordance in a Secure Forensic Psychiatry Hospital – the self rated DUNDRUM-3 programme completion and DUNDRUM-4 recovery scales. *BMC Psychiatry*, 15, 61.
- Davoren, M., O'Reilly, K., Mohan, D. & Kennedy, H. G. (2022). Prospective cohort study of the evaluation of patient benefit from the redevelopment of a complete national forensic mental health service: the DUNDRUM Forensic Redevelopment Evaluation Study (D-FOREST) protocol. *BMJ open*, 12, e058581.
- Eckert, M., Schel, S. H. H., Kennedy, H. G., & Bulten, B. H. (Erik). (2017). Patient characteristics related to length of stay in Dutch forensic psychiatric care. *The Journal of Forensic Psychiatry & Psychology*, 28(6), 863-880.
- Evans, C., Margison, F.R., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health*, 1, 70 - 72.
- Flynn, G., O'Neill, C. & Kennedy, (2011a). DUNDRUM-2: Prospective validation of a structured professional judgment instrument assessing priority for admission from the waiting list for a forensic mental health hospital. *BMC Research Notes* 4, 230.
- Flynn, G., O'Neill, C., McInerney, C., & Kennedy, H. G. (2011b). The DUNDRUM-1 structured professional judgment for triage to appropriate levels of therapeutic security: retrospective-cohort validation study. *BMC psychiatry*, 11, 43.
- Freestone, M., Howard, R., Coid, J. W., & Ullrich, S. (2013). Adult antisocial syndrome co-morbid with borderline personality disorder is associated with severe conduct disorder, substance dependence and violent antisociality. *Personality and mental health*, 7(1), 11-21.
- Freestone, M., Bull, D., Brown, R., Boast, N., Blazey, F., & Gilluley, P. (2015). Triage, decision-making and follow-up of patients referred to a UK forensic service: validation of the DUNDRUM toolkit. *BMC psychiatry*, 15, 239.
- Glancy, G., Choptiany, M., Jones, R., & Chatterjee, S. (2021). Measurement-based care in forensic psychiatry. *International journal of law and psychiatry*, 74, 101650.
- Habets, P., Jeandarme, I. & Kennedy, H. G. (2019). Applicability of the DUNDRUM-1 in a forensic Belgium setting. *The Journal of Forensic Practice*, 21, 85-94.
- Habets, P., Jeandarme, I., & Kennedy, H. G. (2020). Determining security level in forensic psychiatry: a tug of war between the DUNDRUM toolkit and the HoNOS-Secure. *Psychology, Crime & Law*, 1-19.
- Harty, M. A., Shaw, J., Thomas, S., Dolan, M., Davies, L., Thornicroft, G., Carlisle, J., Moreno, M., Leese, M., Appleby, L. & Jones, P. (2004). The security, clinical and social needs of patients in high security psychiatric hospitals in England. *Journal of Forensic Psychiatry and Psychology*, 15, 208-221.
- IBM Corporation (2021) IBM SPSS Statistics per Windows. Versione 28. IBM Corp., Armonk.
- Jeandarme, I., Habets, P., & Kennedy, H. (2019). Structured versus unstructured judgment: DUNDRUM-1 compared to court decisions. *International journal of law and psychiatry*, 64, 205-210.
- Jeandarme, I., Habets, P., O'Reilly, K. & Kennedy, H. G. (2021). Is non-completion of treatment related to security need? *Criminal Behaviour and Mental Health*, 31, 321-330.
- Jewell, M., Pillai, K., Cavney, J., Garrett, N., & McKenna, B. (2023). Examining the need for a high level of therapeutic security at a regional forensic mental health service in Aotearoa New Zealand. *Psychiatry, Psychology and Law*, 31, 293 - 310.
- Jones, R. M., Patel, K., & Simpson, A. I. F. (2019). Assessment of need for inpatient treatment for mental disorder among female prisoners: a cross-sectional study of provincially detained women in Ontario. *BMC psychiatry*, 19(1), 98.

- Kennedy, H. G. (2002). Therapeutic uses of security: mapping forensic mental health services by stratifying risk. *Advances in Psychiatric Treatment*, 8(6), 433-443.
- Kennedy, H. G., O'Neill, C., Flynn, G., Gill, P., Davoren, M.. (2016). *Il DUNDRUM Toolkit V1.0.30, 1.0.30, Dublino, TARA*, 1 - 141.
- Kennedy, H. G., O'Reilly, K., Davoren, M., O'Flynn, P. & O'Sullivan, O. P. (2019a). How to Measure Progress in Forensic Care. In: VÖLLM, B. & BRAUN, P. (eds.) *Long-Term Forensic Psychiatric Care: Clinical, Ethical and Legal Challenges*. Cham: Springer International Publishing.
- Kennedy, H. G., Simpson, A. & Haque, Q. (2019b). Perspective on excellence in forensic mental health services: What we can learn from oncology and other medical services. *Frontiers in Psychiatry*, 10, 733.
- Kennedy, H. G. (2021). Models of care in forensic psychiatry. *BJPsych Advances*, 28, 1-14.
- Lam, A. A., Penney, S. R. & Simpson, A. I. F. (2022). Construct Validity and Concordance of Clinician- and Patient-Rated DUNDRUM Programme Completion and Recovery Scales. *International Journal of Forensic Mental Health*, 1-10.
- Lawrence, D., Davies, T. L., Bagshaw, R., Hewlett, P., Taylor, P. & Watt, A. (2018). External validity and anchoring heuristics: application of DUNDRUM-1 to secure service gatekeeping in South Wales. *British Journal of Psychiatry Bulletin*, 42(1), 10-18.
- McCullough, S., Stanley, C., Smith, H., Scott, M., Karia, M., Ndubuisi, B., Ross, C. C., Bates, R., & Davoren, M. (2020). Outcome measures of risk and recovery in Broadmoor High Secure Forensic Hospital: Stratification of care pathways and moves to medium secure hospitals. *BJPsych Open*, 6, e74.
- McLaughlin, P., Brady, P., Carabellese, F., Carabellese, F., Parente, L., Uhrskov Sorensen, L., Jeandarme, I., Habets, P., Simpson, A. I. F., Davoren, M. & Kennedy, H. G. (2023). Excellence in forensic psychiatry services: International survey of qualities and correlates. *BJPsych Open*, 9, e193.
- O'Dwyer, S., Davoren, M., Abidin, Z., Doyle, E., McDonnell, K., & Kennedy, H. G. (2011). The DUNDRUM Quartet: validation of structured professional judgement instruments DUNDRUM-3 assessment of programme completion and DUNDRUM-4 assessment of recovery in forensic mental health services. *BMC research notes*, 4, 229.
- O'Neill, C., Heffernan, P.M., Goggins, R., Corcoran, C., Linehan, S.A., Duffy, D.M., O'Neill, H., Smith, C.K., & Kennedy, H.G. (2003). Long-stay forensic psychiatric inpatients in the Republic of Ireland: aggregated needs assessment. *Irish Journal of Psychological Medicine*, 20, 119 - 125.
- O'Neill, C., Smith, D., Caddow, M., Duffy, F., Hickey, P., Fitzpatrick, M., Caddow, F., Cronin, T., Joynt, M., Azvee, Z., Gallagher, B., Kehoe, C., Maddock, C., O'Keeffe, B., Brennan, L., Davoren, M., Owens, E., Mullaney, R., Keevans, L., . . . Kennedy, H. G. (2016). STRESS-testing clinical activity and outcomes for a combined prison in-reach and court liaison service: A 3-year observational study of 6177 consecutive male remands. *International Journal of Mental Health Systems*, 10, 67.
- Parente L., Carabellese Fu., Felthous A., La Tegola D., Davoren, M., Kennedy H., Carabellese, Fe., (2023). Italian Evaluation and Excellence in REMS (ITAL-EE-REMS): Appropriate Placement of Forensic Patients in REMS Forensic Facilities. *Research Square*.
- Pierzchniak, P., Farnham, F., Taranto, N. D., Bull, D., Gill, H., Bester, P., McCallum, A. & Kennedy, H. (1999). Assessing the needs of patients in secure settings: A multi-disciplinary approach. *Journal of Forensic Psychiatry*, 10(2), 343-354.
- Pillay, S.M., Oliver, B., Butler, L., & Kennedy, H.G. (2008). Risk stratification and the care pathway. *Irish Journal of Psychological Medicine*, 25, 123 - 127.
- Sharma A., Dunn W., O'Toole C., Kennedy H. G. (2015). The virtual institution: Cross-sectional length of stay in general adult and forensic psychiatry beds. *International Journal of Mental Health System*, 9(1), 25.
- Shaw, J., McKenna, J., Snowden, P., Boyd, C., McMahon, D. & Kilshaw, J. (1994a). Clinical features and placement needs of all North West Region patients currently in Special Hospital. *Journal of Forensic Psychiatry*, 5, 93-106.
- Shaw, J., McKenna, J., Snowden, P., Boyd, C., McMahon, D. & Kilshaw, J. (1994b). The North-West Region. I: Clinical features and placement needs of patients detained in Special Hospitals. *Journal of Forensic Psychiatry*, 5, 93-105.
- Wharewera-Mika, J.P., Cooper, E., Wiki, N., Prentice, K., Field, T.R., Cavney, J., Kaire, D., & McKenna, B. (2020). The appropriateness of DUNDRUM-3 and DUNDRUM-4 for Mori in forensic mental health services in New Zealand: participatory action research. *BMC Psychiatry*, 20, 1-9.
- Williams, H.K., Senanayke, M., Ross, C., Bates, R., & Davoren, M. (2020). Security needs among patients referred for high secure care in Broadmoor Hospital England. *BJPsych Open*, 6.