

Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

LE QUATUOR DUNDRUM : DÉFINITION, VALIDITÉ, UTILITÉ ET RECHERCHE FUTURE

Note informative Novembre 2026 Pinel.qc.ca

Forensia
Centre de formation en santé mentale, justice et sécurité

Institut national de psychiatrie légale Philippe-Pinel
Affilié à l'Université de Montréal

Cette note informative a été réalisée sous la responsabilité de l'Unité d'évaluation des technologies et modes d'intervention en santé mentale, justice, sécurité de Santé Québec — Institut national de psychiatrie légale Philippe-Pinel.

Analyse et rédaction

Ann-Pierre Raiche, Ph. D.
Agente de planification, de programmation et de recherche, Unité d'évaluation des technologies et modes d'intervention en santé mentale, justice, sécurité, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Xavier La Rochelle, B. Sc.
Assistant de recherche, étudiant à la maîtrise, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Ashley J. Lemieux, Ph. D.
Professeure adjointe, Département de sexologie, Université du Québec à Montréal, Chercheuse régulière, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Sophie Saint-Louis, M. Sc
Agente de planification, de programmation et de recherche, Forensia, Centre de formation en santé mentale, justice et sécurité, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Yuliya Bodryzlova, M.D., Ph. D.
Agente de planification, de programmation et de recherche, Unité d'évaluation des technologies et modes d'intervention en santé mentale, justice, sécurité, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Responsables de la réalisation du mandat

Anne Crocker, Ph. D.
Directrice, Recherche et enseignement universitaire, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel, Professeure titulaire, Département de psychiatrie et d'addictologie et École de criminologie, Université de Montréal

Jean-Pierre Guay, Ph. D.
Président du comité scientifique, Forensia, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel, Professeur titulaire, École de criminologie, Université de Montréal

Recherche documentaire

Hongyue Pan, M.S.I.
Bibliothécaire, Centre de documentation Jacques-Talbot, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Demandeurs

Mathieu Dufour
Psychiatre légiste, Directeur médical des services professionnels (DMSP) adjoint, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel, Professeur agrégé de clinique, Département de psychiatrie et d'addictologie de la Faculté de médecine, Université de Montréal

Olivier Jackson
Conseiller cadre, volet mission nationale, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Marc-Antoine Leblanc
Directeur des services multidisciplinaires de santé et de services sociaux (DSMSSS), Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Kim Bédard-Charette
Psychiatre légiste, Directrice médicale et des services professionnels (DMSP), Santé Québec — Institut national de psychiatrie légale Philippe-Pinel, Professeure adjointe de clinique, Département de psychiatrie et d'addictologie de la Faculté de médecine, Université de Montréal

Antonella Orsini
Criminologue, Adjointe à la Direction médicale et des services professionnels, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Conseillère statistique

Marie-Christine Stafford, M. Sc.
Statisticienne, Groupe de consultation statistique, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Conseillère scientifique

Marichelle Leclair, Ph. D.
Professeure régulière, Département de psychoéducation et psychologie, Université du Québec en Outaouais, Chercheuse régulière, Santé Québec — Institut national de psychiatrie légale Philippe-Pinel

Production du rapport

Édition
Ann-Pierre Raiche

Correction
Corinne de Vally

Mise en page
Samia Herrera

Pour citer ce document

Raiche, A.-P., Larochelle, X., Saint-Louis, S., Lemieux, A.J., Leclair, M., Pan, H., Bodryzlova, Y., Guay, J.-P. et Crocker, A.G. (2026). *Le Quatuor DUNDRUM: définition, validité, utilité et recherche future.* Montréal, Québec (Canada) : Santé Québec — Institut national de psychiatrie légale Philippe-Pinel.

Pour des renseignements sur l'UETMI-SMJS et pour consulter nos publications : Unité d'évaluation des technologies et modes d'intervention en santé mentale, justice et sécurité (UETMI-SMJS), Institut national de psychiatrie légale Philippe-Pinel, 10 905, boul. Henri-Bourassa Est, Montréal (Québec) H1C 1H1, uetmi.smjs.pinel@ssss.gouv.qc.ca, https://pinel.qc.ca/unite-devaluation-des-technologies-et-modes-dintervention/

Déclaration d'intérêts :
les auteurs et établissements participants ne déclarent aucun conflit d'intérêts dans la réalisation de ce projet.

Dépôt légal : Bibliothèque et Archives nationales du Québec, 2026
ISBN : 978-2-921483-58-2 (PDF)

Tous droits réservés
© UETMI-SMJS — Institut national de psychiatrie légale Philippe-Pinel, 2026

Table des matières

Liste des abréviations et acronymes 3

Mandat 4
Historique 4
Hiérarchisation des services en psychiatrie légale 5
Objectif du rapport 6

Principaux constats 7
Le contexte 7
Le Quatuor DUNDRUM 8
DUNDRUM 1 10
Fidélité 10
Validité 10
DUNDRUM 2 11
Fidélité 11
Validité 11
DUNDRUM 3 12
Fidélité 12
Validité 12
DUNDRUM 4 13
Fidélité 13
Validité 13

Introduction 14
Le Quatuor DUNDRUM 15
Les échelles 15
Question décisionnelle 17

Méthodologie 18
Repérage des documents 18
Sélection des documents 18
Extraction des données, codification, analyse et synthèse 18

Résultats et interprétation 19
3.1. Fidélité 21
Cohérence interne 22
Fidélité interjuge 24

3.2. Validité **26**
Validité critériée 26
Validité de construit 29

Discussion **34**
Limites 34
Conclusion 34

Références **35**

Annexes **40**

LISTE DES ABRÉVIATIONS ET ACRONYMES

ACB	*Anticholinergic Cognitive Burden scale*
APT	Aptitude à subir son procès
CANFOR	*Camberwell Assessment of Need — Forensic Version*
CISSS et CIUSSS	Centre intégré (universitaire) de santé et services sociaux
CH	Centres hospitaliers
COSMIN	*COnsensus-based Standards for the selection of health Measurement Instruments.*
DUNDRUM	*Dangerousness, Understanding, Recovery and Urgency Manual.*
ETMI	Évaluation des technologies et modes d'intervention
GAF	*Global Assessment of Functioning*
HCR-20	*Historical Clinical Risk Management-20*
HoNOS-S	*Health of the Nation Outcome Scale for Users of Secure and Forensic Services*
MSSS	Ministère de la Santé et des Services sociaux
PANSS	*Positive And Negative Syndrome Scale*
RC	Responsabilité criminelle
SAPROF	*Structured Assessment of Protective Factors*
SIM	Suivi intensif dans le milieu
SIV	Suivi d'intensité variable
SM	Santé mentale
SOFAS	*Social and Occupational Functioning Scale*
S-RAMM	*Suicide Risk Assessment and Management Manual*
START	*Short Term Assessment of Risk and Treatability* (évaluation du risque à court terme et traitabilité)

MANDAT

Historique

Après une vaste consultation auprès des parties prenantes du réseau de la santé et des services sociaux au début de l'année 2014, le ministère de la Santé et des Services sociaux (MSSS) dépose le Plan d'action en santé mentale 2015-2020. Ce dernier propose, entre autres, des mesures pour compléter une hiérarchisation des services dans le but d'améliorer l'accès et la continuité des soins, en général et, plus précisément, en psychiatrie légale (MSSS, 2015). En mai 2014, l'Agence de la santé et des services sociaux de Montréal formule la *Proposition d'organisation de services hiérarchisés en psychiatrie légale à Montréal* (Tourigny et al., 2014) dans laquelle le *Dangerousness, UNDerstanding, Recovery and Urgency Manual* (DUNDRUM ; Kennedy, 2016a) est ciblé pour accompagner les intervenants dans la prise de décision et encadrer les patients dans leurs trajectoires de soins. Cet instrument est identifié notamment en raison d'une première recension des écrits non publiée (Perafan et Crocker, 2014). En 2015, le milieu de la santé et des services sociaux est significativement restructuré par la création des Centres intégrés (universitaires) de santé et services sociaux (les CISSS et CIUSSS) et les Agences sont démantelées. Dans la foulée de ces changements, en 2017, l'ancien Institut Philippe-Pinel de Montréal devient l'Institut national de psychiatrie légale Philippe-Pinel (maintenant Santé Québec — Institut national de psychiatrie légale Philippe-Pinel). De par sa mission nationale unique, à l'intersection des milieux de la santé, de la justice et de la sécurité publique, ce dernier joue désormais un rôle-conseil auprès du MSSS pour tout ce qui concerne la psychiatrie légale. Qui plus est, en 2021, le centre de formation intersectoriel en santé mentale, justice et sécurité Forensia de l'Institut est mis sur pied. Le mandat premier de Forensia est d'instrumentler et de soutenir le réseau de la santé et des services sociaux et ses partenaires des réseaux de la justice et de la sécurité publique quant aux meilleures pratiques en santé mentale forensique. L'Institut a donc comme rôle d'identifier et d'implanter des pratiques phares s'appuyant sur les meilleures données du moment dans le milieu de la psychiatrie légale sur l'ensemble du territoire québécois. C'est dans cet esprit que le MSSS, soutenu par l'Institut, met en œuvre un projet national de hiérarchisation des services en psychiatrie légale, projet qui s'inscrit dans les travaux découlant du Plan d'action interministériel en santé mentale 2022-2026 (action 5.6).

Hiérarchisation des services en psychiatrie légale

La hiérarchisation des services en psychiatrie légale a pour objectif d'améliorer la fluidité, l'efficience et la sécurité des services en psychiatrie légale relevant du Code criminel (Gouvernement du Québec, 2022). Pour ce faire, la hiérarchisation demande l'adoption d'une organisation optimale de services hiérarchisés en psychiatrie légale pour l'ensemble du territoire québécois. De ce fait, la hiérarchisation veut tenir compte de la mission des centres hospitaliers qui est de soigner et d'assurer la protection du public, en diminuant le nombre d'hôpitaux désignés pouvant accueillir les patients forensiques. En diminuant le nombre d'hôpitaux désignés, l'objectif est de concentrer l'expertise et de diminuer la variabilité dans l'application des pratiques en lien avec la partie XX.1 du Code Criminel incluant le respect des délais prévus au code criminel pour les évaluations forensiques (c.-à-d. aptitude à subir son procès, responsabilité criminelle et traitement pour rendre apte). Ces objectifs mettent de l'avant l'importance de confier les patients présentant un risque pour la sécurité d'autrui à l'établissement approprié en fonction du niveau de risque qu'ils présentent ainsi que de leur besoin thérapeutique. On parlera, dans ce cas, du niveau de risque et de besoins des personnes en termes de sécurité thérapeutique (De Pau et al., 2021 ; Kennedy, 2002). De manière générale, cette décision doit être prise rapidement par des professionnels variés (policiers, criminologues, agents de la cour, psychiatres, travailleurs sociaux, etc.) au moment de la prise en charge par les services sociaux ou du premier contact avec le système de justice et réévaluée systématiquement au moins une fois par année pour chaque usager confié au responsable d'hôpital désigné par la Commission d'examen de troubles mentaux jusqu'à sa sortie de la trajectoire de psychiatrie légale. Il importe donc, non seulement de choisir un instrument approprié pour prendre ces décisions, mais également d'assurer une formation uniforme des utilisateurs par des connaissances identifiées à travers le réseau de la santé et des services sociaux et le réseau de la justice et de la sécurité publique.

Objectif du rapport

L'objectif du présent rapport est donc de réaliser une recension des écrits des propriétés psychométriques du DUNDRUM afin de documenter les « données » qui soutiennent le choix de cet instrument. Ce rapport soutient la formation actuellement offerte en fournissant un complément d'information visant à éclairer le choix du DUNDRUM comme instrument pour faciliter la prise de décision sur la trajectoire de soins des patients en psychiatrie légale, de même que ses forces et ses limites. Les travaux de ce rapport s'appuient notamment sur le premier rapport (non publié) de 2014 (Perafan et Crocker, 2014), avec une mise à jour des nouvelles données qui ont émergé depuis sa première itération. Ce rapport ne se limite pas à une recension des propriétés psychométriques, il propose également une réflexion sur les différentes implications cliniques de ces propriétés psychométriques.

PRINCIPAUX CONSTATS

Le contexte

À la lumière de la hiérarchisation des soins en psychiatrie légale, il importe d'utiliser un instrument validé pouvant servir à orienter les personnes vers le niveau de soins et de sécurité approprié selon leurs besoins. L'utilisation d'un tel instrument permettra, d'une part, la création d'un langage commun pour une offre de service tout au long de la trajectoire de soins des personnes, en raison notamment de formations dispensées aux utilisateurs de connaissances. D'autre part, le choix d'un instrument validé est essentiel pour garantir des évaluations fiables, objectives et pertinentes afin d'améliorer la qualité des informations sur lesquelles repose la prise de décision. Le Quatuor DUNDRUM est un instrument d'évaluation structurée comprenant quatre échelles destinées à soutenir la prise de décision lors de l'évaluation des besoins en matière de sécurité thérapeutique. Selon les auteurs, le concept de sécurité thérapeutique regroupe les notions de sécurité et de besoins d'encadrement et de traitement. La sécurité thérapeutique, qui inclut la sécurité physique, relationnelle et procédurale, représente une condition essentielle au rétablissement (Kennedy, 2002, 2022). Cet instrument, fortement influencé par les instruments de jugement professionnel structuré en santé mentale forensique, notamment le *Historical Clinical Risk Management-20* (HCR-20 ; Douglas et al., 2013), est conçu pour évaluer et guider les décisions cliniques liées à la sécurité des patients dans le contexte de la psychiatrie légale. Le Quatuor DUNDRUM a été ciblé en raison du nombre limité d'instruments répondant à ce besoin et puisque plusieurs études soutiennent son utilité.

Le Quatuor DUNDRUM

Le Quatuor DUNDRUM est un instrument d'aide à la décision fondé sur une méthode de jugement professionnel structuré, comprenant quatre échelles aux objectifs distincts (DUNDRUM-1, DUNDRUM-2, DUNDRUM-3 et DUNDRUM-4). Le DUNDRUM-1 se concentre sur le triage et vise à déterminer le niveau de sécurité thérapeutique approprié à accorder à un patient en évaluant divers facteurs, notamment la gravité des comportements violents ou auto-agressifs, les besoins spécialisés en psychiatrie légale, le risque d'évasion, et d'autres considérations pertinentes au besoin d'encadrement. Le DUNDRUM-2 évalue l'urgence au triage en fonction des besoins du patient, permettant ainsi de prioriser les admissions sur une liste d'attente en tenant compte de différents critères, y compris la situation juridique du patient. Le DUNDRUM-3 permet d'identifier les cibles d'intervention efficaces en matière de prévention de la récidive criminelle tandis que le DUNDRUM-4 sert à identifier des cibles d'intervention qui sont spécifiques aux personnes qui vivent avec des troubles mentaux et qui ont commis des infractions. Pour les DUNDRUM-3 et 4, il existe également une version auto-évaluée, à remplir par le patient. Chacun des items comporte une échelle d'évaluation en cinq points permettant de déterminer le degré de présence des critères. Il est à noter que, malgré l'utilisation d'items faisant l'objet d'une cotation numérique, il ne s'agit pas d'une mesure actuarielle reposant sur une pondération prédéterminée des différents facteurs. Plus précisément, la cote globale revient à une décision finale de l'évaluateur en soupesant la prépondérance relative des items associés aux besoins prioritaires/risques.

L'utilisation de cet instrument structuré soutient une évaluation plus cohérente, objective et pertinente des besoins en matière de sécurité thérapeutique des patients en psychiatrie légale. Il est conçu pour aider les cliniciens à prendre des décisions éclairées et à garantir un traitement approprié en fonction des caractéristiques spécifiques de chaque patient. Le Quatuor DUNDRUM offre ainsi une approche systématique et professionnelle pour aborder les défis liés à la sécurité et aux besoins d'intervention efficaces dans le domaine de la psychiatrie légale.

PROPRIÉTÉS PSYCHOMÉTRIQUES

La validation psychométrique est un processus scientifique qui consiste à évaluer la fiabilité et la validité des instruments afin de garantir que ceux-ci sont appropriés à l'usage auquel ils sont destinés. Cela permet aux chercheurs, aux cliniciens et aux décideurs de recueillir des données précises et significatives pour prendre des décisions éclairées. Les principales propriétés psychométriques explorées pour valider un instrument sont la fidélité (cohérence interne et fidélité interjuge) ainsi que la validité (validité critériée et validité de construit).

Fidélité : capacité d'un test à donner des résultats constants ou fiables indépendamment du nombre de fois où il est administré, de ses correcteurs et de son contenu particulier.

Cohérence interne : corrélation qui existe entre tous les items qui composent l'échelle (degré d'intercorrélations), indiquant si lesdits items mesurent bien le même construit.

Fidélité interjuge : concordance (degré d'entente) entre les cotes attribuées par différents évaluateurs lors de l'évaluation d'une même personne, au même moment et dans les mêmes conditions, indiquant dans quelle mesure l'instrument est interprété de la même façon par différents évaluateurs.

Validité : capacité d'un instrument à mesurer ce pourquoi il a été conçu (construit prévu).

Validité critériée (aussi appelée validité de critère ou validité liée au critère) : capacité d'un instrument à distinguer les personnes évaluées entre elles et au degré auquel les scores d'un instrument reflètent adéquatement un critère externe, définis comme le gold standard. Lorsque ce critère se définit dans le temps et succède à l'administration de l'instrument, on parlera de validité prédictive (une forme de mesure de la validité critériée).

Validité de construit (aussi appelée validité concourante ou validité convergente) : capacité de l'instrument à mesurer un concept abstrait et le degré avec lequel ce dernier reflète les composantes théoriques sur lequel il repose. Celle-ci est généralement évaluer en examinant le degré auquel l'instrument corrèle avec d'autres instruments mesurant des construits similaires, liés ou non liés.

DUNDRUM 1

FIDÉLITÉ

Cohérence interne

- La cohérence interne du DUNDRUM-1 est variable, quoique généralement bonne, avec de meilleurs résultats observés dans les études disposant de larges échantillons.

- Les éléments liés à la suicidalité (S2 et S4) semblent constituer un aspect distinct, bien que nécessaire dans l'évaluation du requis de sécurité thérapeutique.

Fidélité interjuge

- Le DUNDRUM-1 démontre une bonne fidélité interjuge, bien que le niveau d'entente semble influencé par le degré de formation obtenu par les cliniciens qui font la passation de l'instrument.

VALIDITÉ

Validité critériée

- Le DUNDRUM-1 a une bonne capacité pour distinguer entre les groupes provenant de différents niveaux de sécurité, bien qu'il soit difficile de prédire le niveau de sécurité dans lequel se trouve un patient au moment de l'administration de l'instrument. Plutôt que de mettre en doute la validité de l'instrument, cette limite reflète plutôt l'importance d'utiliser un instrument comme le DUNDRUM pour les décisions de placement des patients. En effet, les différences entre le score total de l'échelle ainsi que l'attribution du niveau d'encadrement/sécurité sans l'utilisation préalable d'un tel instrument est probablement attribuable au fait qu'il existe un plus grand degré de subjectivité dans les évaluations qui ne sont pas guidées par un instrument de jugement professionnel structuré.

- La valeur prédictive du DUNDRUM-1 est significative sur du moyen terme (3-6 mois), mais l'est moins au long terme (12 mois+).

Validité de construit

- Le DUNDRUM-1 corrèle généralement dans le sens attendu avec d'autres instruments conçus pour mesurer des construits similaires tels que les autres échelles du DUNDRUM, le nombre de besoins comblés ou non (CANFOR), les mesures de risques de violence et de comportements problématiques (HCR-20, START, SAPROF) ainsi que les problèmes liés à la sécurité (HoNOS-S).

- Ces résultats soutiennent la validité de construit de l'instrument en tant que mesure du niveau de sécurité thérapeutique requis au moment du triage/évaluation.

DUNDRUM 2

FIDÉLITÉ

Cohérence interne

- Les données actuelles ne permettent pas de confirmer la cohérence interne du DUNDRUM-2, en raison de la faible quantité d'études, aux résultats contradictoires et aux petits échantillons.

Fidélité interjuge

- La fidélité interjuge du DUNDRUM-2 semble variable, avec de meilleurs résultats pour les cliniciens mieux formés.

VALIDITÉ

Validité critériée

- Le DUNDRUM-2 semble pouvoir identifier les patients nécessitant des soins urgents, bien qu'il y ait peu d'études et que les résultats ne soient pas constants à travers celles-ci.

- Tout comme pour le DUNDRUM-1, il est possible que cet enjeu soit plus tributaire du contexte que de l'instrument. Il est possible qu'en l'absence d'un instrument de jugement professionnel structuré, d'autres critères plus subjectifs soient mis de l'avant pour prioriser les demandes d'admissions.

Validité de construit

- Le DUNDRUM-2 corrèle positivement avec le DUNDRUM-1, avec une force variant de faible à élevée selon les études. Il présente également des corrélations positives avec des enjeux liés à la sécurité (sous-échelle Security du HoNOS-S).

- Ces résultats préliminaires appuient la validité de construit de l'instrument en tant que mesure de l'urgence d'un placement sécuritaire au moment du triage.

DUNDRUM 3

FIDÉLITÉ

Cohérence interne

- La cohérence interne du DUNDRUM-3 est excellente, signifiant que les items mesurent de manière cohérente le concept de progression thérapeutique du patient relative à différentes cibles d'intervention identifiées, tel que cela est défini par l'instrument.

Fidélité interjuge

- Une seule étude a exploré la fidélité interjuge du DUNDRUM-3, avec des résultats qui démontrent un niveau d'entente relativement satisfaisant.

VALIDITÉ

Validité critériée

- Le DUNDRUM-3 permet de différencier différents groupes de patients selon leur niveau de sécurité, mais semble évaluer à la hausse le besoin de sécurité thérapeutique des patients.

Validité de construit

- La DUNDRUM-3 corrèle avec le niveau de besoins (CANFOR), les facteurs de risque dynamiques et cliniques (HCR-20), les forces et vulnérabilités (START), la symptomatologie psychiatrique (PANSS), ainsi qu'avec le fonctionnement global et cognitif (SOFAS, ACB, MCCB) et les auto-évaluations des patients.

- Ces résultats soutiennent la validité de construit de la progression thérapeutique, tel que cela est défini par le DUNDRUM-3.

DUNDRUM 4

FIDÉLITÉ

Cohérence interne

- La cohérence interne du DUNDRUM-4 est excellente, signifiant que les items mesurent de manière cohérente le concept de rétablissement, tel que cela est défini par l'instrument.

Fidélité interjuge

- Une seule étude a exploré la fidélité interjuge du DUNDRUM-4, avec des résultats qui démontrent un niveau d'entente relativement satisfaisant.

VALIDITÉ

Validité critériée

- Le DUNDRUM-4 permet de différencier différents groupes de patients selon différentes variables, mais semble évaluer à la hausse le besoin de sécurité thérapeutique des patients.

Validité de construit

- Le DUNDRUM-4 corrèle avec le niveau de besoins (CANFOR), les facteurs de risque (HCR-20), les forces et vulnérabilités (START), la symptomatologie psychiatrique (PANSS), et les auto-évaluations des patients, tout en étant associé négativement au fonctionnement global (GAF) et aux facteurs de protection (SAPROF).

- Ces résultats soutiennent la validité de construit du rétablissement tel que cela est défini par le DUNDRUM-4.

INTRODUCTION

La décision de transférer un patient ayant une problématique psycholégale d'un établissement d'un niveau de sécurité élevée, à modérée, à minimale, et à la communauté (et inversement), est l'une des plus importantes que doivent prendre les professionnels forensiques (Davoren et al., 2012). Par ailleurs, les pratiques d'évaluation du risque de violence ont grandement évolué au cours des dernières années et des approches structurées d'évaluation et de gestion du risque sont maintenant privilégiées (Crocker et al., 2011, 2012, 2023 ; Guay et al., 2022 ; National Institute for Health and Care Excellence [NICE], 2015 ; Wright et Webster, 2011). L'identification des facteurs de risque aide à la planification du plan d'intervention, et par conséquent, l'évaluation du risque est devenue essentielle à la pratique de la santé mentale forensique fondée sur des données probantes (Abidin et al., 2013) et se doit ainsi d'être une partie intégrante de l'organisation des services.

Le Quatuor DUNDRUM

Le Quatuor DUNDRUM (de l'acronyme anglais *Dangerousness, UNDerstanding, Recovery and Urgency Manual*) est un instrument d'aide à la décision fondé sur une méthode de jugement professionnel structuré. Ce dernier, développé au *Central Mental Hospital* de Dundrum en Irlande, vise à appuyer la prise de décision lors de l'évaluation du besoin quant à la sécurité thérapeutique (Kennedy et al., 2016a). Cet instrument, fortement influencé par l'instrument d'évaluation et de gestion de risque HCR-20 (Douglas et al., 2013), comprend quatre échelles distinctes pertinentes à la prise de décision.

LES ÉCHELLES

Les quatre échelles du DUNDRUM sont composées d'une série d'items qui correspondent à des facteurs clés lorsqu'il s'agit de prendre des décisions relatives au besoin de sécurité thérapeutique (c'est-à-dire les besoins d'encadrement et les besoins de traitement et/ou d'intervention efficaces) (Flynn, O'Neill, McInerney, et al., 2011).

Le DUNDRUM-1 *(niveau de sécurité au triage)* porte sur le triage, et plus particulièrement, sur la décision quant à l'assignation d'un patient au niveau approprié de sécurité thérapeutique (Flynn, O'Neill, McInerney, et al., 2011). Il comprend un ensemble de facteurs jugés pertinents dans l'évaluation du besoin de sécurité thérapeutique (Kennedy et al., 2016a). Précisément, le DUNDRUM-1 décrit les 11 éléments suivants: S1) la gravité des gestes violents ; S2) la gravité des comportements auto-agressifs ; S3) l'imminence du risque de violence relié au trouble mental ; S4) l'imminence du risque de suicide ; S5) les besoins spécialisés en psychiatrie légale ; S6) le risque d'évasion et de fugue ; S7) le contrôle des accès ; S8) les considérations envers les victimes et la confiance du public ; S9) les besoins complexes liés au risque de violence ; S10) le comportement en établissement et S11) le processus judiciaire. Bien que la présence de plusieurs troubles mentaux ne soit pas exclue, jusqu'à maintenant les études portent sur des populations présentant au moins un trouble mental (axe 1).

Le DUNDRUM-2 *(niveau d'urgence au triage)* porte sur l'urgence du triage basée sur les besoins du patient, et par conséquent, sur la détermination de sa place sur la liste d'attente par rapport aux autres patients (Flynn, O'Neill et Kennedy, 2011). Le DUNDRUM-2 est composé de 6 items. Le premier évalue le niveau général d'urgence au triage en fonction du milieu de provenance de la personne évaluée. Il existe donc 5 versions différentes de l'item U1, selon si la personne est : a) un patient de psychiatrie légale en suivi externe ; b) en attente d'un procès ou en détention provisoire ; c) un détenu sentencié ; d) en attente d'un transfert vers un niveau de sécurité plus élevé ou e) en attente d'un transfert vers un niveau de sécurité équivalent ou plus faible. Les 5 items subséquents sont les mêmes pour tous et portent sur : U2) la santé mentale ; U3) la prévention du suicide ; U4) les considérations humanitaires ; U5) les considérations systémiques et U6) la priorisation judiciaire. Le DUNDRUM-2 permet d'évaluer qui, sur une liste d'attente de personnes répondant aux critères d'admission, devrait être admis en priorité (Kennedy et al., 2016a).

Le DUNDRUM-3 *(progression dans les programmes)* permet d'identifier les cibles d'intervention efficaces en matière de prévention de la récidive criminelle. Les interventions particulières structurées mises en œuvre devraient correspondre aux facteurs de risque associés aux comportements ayant conduit à l'admission dans un établissement spécialisé (Kennedy et al., 2016a). Ses items sont conceptuellement formulés de manière à situer le patient par rapport à une progression théorique quant aux cibles d'intervention probante en matière de prévention de la récidive criminelle ou en matière de réitération des comportements délinquants. Le DUNDRUM-3 comporte les 7 éléments suivants : P1) santé physique ; P2) santé mentale ; P3) alcool et drogues ; P4) problèmes de comportement ; P5) soins personnels et activités de la vie quotidiennes (AVQ) ; P6) éducation, travail, passe-temps et créativité et P7) famille, réseaux sociaux, amitié et intimité. Il permet donc d'évaluer les domaines à cibler et les objectifs à inclure au plan d'intervention, en lien avec le rétablissement et la gestion du risque (Kennedy et al., 2016a). En plus de l'instrument de jugement structuré, il existe également une version auto-évaluée de cette échelle.

Le DUNDRUM-4 *(rétablissement[1])* quant à lui permet d'identifier des cibles d'intervention qui sont particulières aux personnes qui vivent avec de troubles mentaux et qui ont commis des infractions. Il est important de préciser que l'échelle du rétablissement ne porte pas sur une vision holistique du cheminement individuel de chacun, mais plutôt sur le rétablissement de la maladie. Le DUNDRUM-4 est composé des 6 éléments suivants : R1) la stabilité ; R2) l'introspection ; R3) le lien thérapeutique ; R4) les congés et sorties ; R5) le risque dynamique et R6) la sensibilité face aux victimes (Kennedy et al., 2016a). Un septième élément a également été ajouté dans la version plus récente de cette échelle : l'espoir. Le DUNDRUM-4 complète le DUNDRUM-3 en documentant des cibles d'intervention qui sont particulières aux personnes qui vivent avec des troubles mentaux et qui ont commis des infractions. Le D4 identifie des critères qui portent sur le rétablissement de la maladie. Ainsi l'accent est mis sur la stabilité de la maladie, l'introspection, le lien thérapeutique avec l'équipe traitante et l'octroi de sorties. Selon les auteurs, l'amélioration dans ces domaines contribue au rétablissement de la maladie (Kennedy et al., 2016a). Tout comme pour le DUNDRUM-3, il existe à la fois une version administrée par l'intervenant et une version auto-évaluée, par lequel le patient évalue ses propres besoins en termes de rétablissement (Davoren et al., 2012).

Chacune de ces échelles comprend une évaluation en cinq points allant de 0 (ceux considérés comme pouvant être suivis dans la communauté ou même être libérés) à 4 (ceux considérés comme devant être maintenus dans une institution à sécurité maximale). Ainsi, en ce qui a trait au DUNDRUM-1, ceux recevant une cote de 4 seraient assignés à un établissement à sécurité élevée (tel que Santé Québec — Institut national de psychiatrie légale Philippe-Pinel), une cote de 3 à un établissement à sécurité moyenne (tel que les unités sécurisées des établissements de soins psychiatriques), une cote de 2 à un niveau de sécurité faible (tel que les ailes psychiatriques des centres hospitaliers désignés), ceux recevant une cote de 1 pourraient être suivis en services ambulatoires externes (tel que les équipes SIM et SIV) et ceux avec une cote de 0 pourraient être suivis en services ambulatoires ou être libérés sans conditions (Flynn, O'Neill, McInerney, et al., 2011). En ce qui concerne le DUNDRUM-2, les cotes se réfèrent à un niveau d'urgence, où une cote plus élevée signifie un niveau d'urgence plus important, soit une priorisation lorsqu'il y a une liste d'attente pour l'établissement ciblé à la suite de l'administration du DUNDRUM-1. Enfin, lorsqu'on se réfère au DUNDRUM-3 et au DUNDRUM-4, ceux recevant une cote de 4 présentent des besoins de traitement importants et ne sont pas considérés comme étant prêts à être transférés de l'établissement sécurisé où ils se trouvent (et s'il s'y trouvent) ; ceux avec une cote de 3 sont considérés comme prêts à être transférés d'un établissement à sécurité élevée à un établissement à sécurité moyenne ; ceux avec une cote de 2 seraient prêts à être transférés d'un établissement à sécurité moyenne à un établissement à sécurité faible ; ceux avec une cote de 1 sont prêts à être transférés dans la communauté, et ceux avec une cote de 0 présentent des progrès stables et sont considérés comme étant prêts à recevoir une libération inconditionnelle (O'Dwyer et al., 2011). Toutefois, une cote ne détermine pas automatiquement un transfert. Une telle décision doit être prise à

[1] Il est à noter que, dans le DUNDRUM, le concept de rétablissement renvoie au rétablissement psychopathologique plutôt qu'au rétablissement personnel.

la lumière de l'histoire de l'individu, de son état clinique ainsi que d'une évaluation du risque. Par ailleurs, il importe de préciser que les cotes d'attribution aux différents niveaux de sécurité proposées dans le présent rapport constituent une extrapolation du DUNDRUM visant à l'arrimer au modèle québécois de hiérarchisation des services. Elles doivent donc être interprétées en tenant compte du contexte dans lequel l'instrument est utilisé. À titre d'exemple, en Irlande, les trois niveaux de sécurité sont regroupés au sein d'un même établissement, alors qu'au Québec, ils sont répartis entre différents établissements et organisations de services. Voir la figure 1 pour un exemple des services hiérarchisés en fonction du niveau de besoin tel que cela est déterminé par le DUNDRUM.

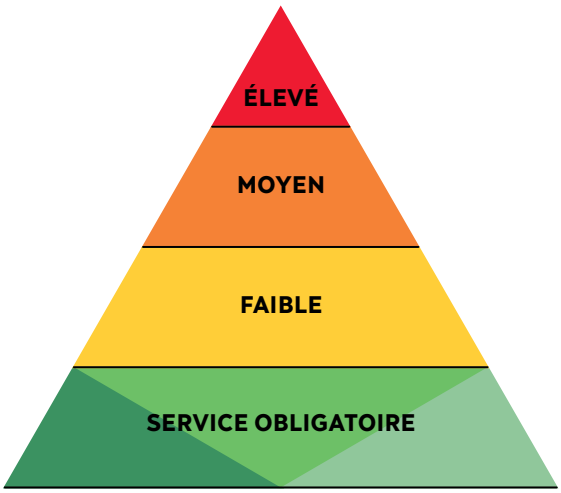

Figure 1. Hiérarchisation des services en psychiatrie légale en fonction du niveau de spécialisation (requis de sécurité et complexité clinique)

QUESTION DÉCISIONNELLE

Le Quatuor DUNDRUM semble être un instrument adéquat pour répondre aux besoins de la hiérarchisation des services en psychiatrie légale en facilitant la prise de décision tout au long de la trajectoire de soin des patients quant au niveau de sécurité thérapeutique nécessaire au moment de la prise en charge (DUNDRUM-1) et tout au long de l'épisode de soin (DUNDRUM-3 et DUNDRUM-4), ainsi qu'au niveau de l'urgence du placement (DUNDRUM-2). Or, pour véritablement soutenir la prise de décision à grande échelle, un instrument se doit d'être, d'une part, standardisé et d'autre part, validé. La standardisation réfère au fait qu'un instrument est développé de manière rigoureuse et mesure un concept donné, de manière objective, à travers une série d'échelles numérotées. Ceci permet l'administration de cet instrument de façon uniforme, ainsi qu'un langage commun pour tous les utilisateurs de l'échelle. La validation, elle, est un processus par lequel on s'assure qu'un instrument mesure réellement le concept escompté : est-ce que cet instrument performe de la même manière s'il est administré par différents professionnels ? Sur différentes populations (selon le genre, la problématique de santé mentale, le point d'entrée dans le système de la santé ou de justice, etc.) ? Dans différentes langues ? Pour répondre à ces questions, plusieurs études auprès de populations variées et menées dans des contextes diversifiés sont nécessaires pour confirmer que l'instrument peut effectivement être utilisé à l'extérieur du contexte dans lequel il a été développé (p. ex. utilisé en anglais par des psychologues dans un hôpital sécurisé d'Irlande avec des patients psychotiques). Lorsqu'il est question de validité d'un instrument, il est habituellement question de ses qualités ou de ses propriétés psychométriques. Ces propriétés nous permettent d'évaluer la qualité d'un instrument. De plus, la compréhension des propriétés psychométriques d'un instrument permet aux utilisateurs d'évaluer plus précisément la qualité des instruments qu'ils emploient. Ainsi, lorsqu'il est question de choisir un instrument pour une utilisation à l'échelle provinciale pour soutenir la hiérarchisation des services en psychiatrie légale, il importe de s'assurer que celui-ci présente des propriétés psychométriques satisfaisantes. Qui plus est, il est essentiel que Santé Québec — Institut national de psychiatrie légale Philippe-Pinel, en tant qu'institut national, promeuve des pratiques phares qui s'appuient sur des données probantes. La question qui se pose est donc la suivante :

Quelles sont les qualités psychométriques du Quatuor DUNDRUM et est-ce que celles-ci nous permettent d'implanter son utilisation à large échelle pour soutenir la hiérarchisation des services en psychiatrie légale à l'échelle provinciale ?

MÉTHODOLOGIE

REPÉRAGE DES DOCUMENTS

Le processus de revue de la littérature a été entamé en 2014 par une agente de recherche. La bibliothèque électronique de recherche de l'Université McGill a été utilisée pour identifier les articles scientifiques traitant du Quatuor DUNDRUM. Les mots-clés « DUNDRUM Quartet », « DUNDRUM Instrument » et « DUNDRUM Toolkit », ont été utilisés lors de la recherche dans les bases de données MEDLINE in-process, EMBASE : Excerpta Medica, PubMed, PsycINFO, SCOPUS, ProQuest Central ainsi que Academic Search Complete (EBSCO). Cette stratégie a été utilisée à nouveau en octobre 2022 lors de la mise à jour du présent document et, par la suite, une veille a été mise en place dans MEDLINE, EMBASE et PsycINFO par une bibliothécaire de Santé Québec — Institut national de psychiatrie légale Philippe-Pinel afin de faciliter l'incorporation de nouveaux articles (2019-2026).

SÉLECTION DES DOCUMENTS

La sélection des documents et l'extraction des données ont été effectuées par un agent de recherche, supervisé par une professionnelle en évaluation des technologies et modes d'intervention. Une première sélection des documents a été faite en consultant les titres et résumés. Ensuite, les articles ont été consultés afin d'identifier ceux qui contenaient des informations quant aux propriétés psychométriques des instruments. La période répertoriée s'étale de 2010, au moment de la création de la première version du Quatuor DUNDRUM, à octobre 2026, au moment de la préparation de ce rapport. Un article a été identifié par le biais de recherches manuelles. En somme, 31 articles ont été répertoriés. Les données descriptives relatives à chaque article inclus dans la revue systématique sont disponibles en annexe (voir tableau supplémentaire 1, en annexe).

EXTRACTION DES DONNÉES, CODIFICATION, ANALYSE ET SYNTHÈSE

Afin de faciliter l'analyse systématique des articles, une grille d'extraction a été développée selon les recommandations COSMIN (Prinsen et al., 2018)[2]. Ces dernières ont comme objectif d'aider à la sélection des meilleurs instruments pour mesurer l'état de santé dans la pratique clinique et la recherche. Développé d'abord pour les instruments auto-rapportés par les patients, ce guide peut également servir dans la sélection d'instruments de cotation par les cliniciens ou d'instruments de diagnostic, en adaptant la méthodologie. Nous avons suivi une version adaptée du protocole COSMIN pour les revues systématiques. Il importe cependant de noter que dans le cadre de la présente recension, les recommandations COSMIN ont seulement servi de feuille de route pour guider le travail, plutôt que de suivre une application rigide de ses principes. La grille d'extraction incluait des informations au sujet de l'échantillon (n, âge, sexe), le contexte d'administration de l'instrument (établissement, description des patients, pays, échelle du DUNDRUM utilisée), ainsi que les indices psychométriques. Les différents indices répertoriés ont été regroupés selon leur habileté à rendre compte 1) de la fidélité de l'instrument, soit la capacité d'un test de donner des résultats constants ou fiables indépendamment du nombre de fois où il est administré, de ses correcteurs et de son contenu particulier (Hogan, 2017) ou 2) de sa validité, soit la capacité d'un test à mesurer ce qu'il a été conçu pour mesurer (Hogan, 2017).

[2] COnsensus-based Standards for the selection of health Measurement Instruments

RÉSULTATS ET INTERPRÉTATION

Au total, 31 articles ont été inclus dans cette recension, toutes échelles et toutes propriétés psychométriques confondues. Pour connaître les données descriptives des études retenues, consultez le tableau supplémentaire 1 présenté en annexe. Pour une synthèse des résultats, selon chaque échelle pour chaque propriété psychométrique évaluée, consulter le tableau 1. À l'aide de la grille d'extraction, chaque article a été révisé selon sa capacité à rendre compte des propriétés psychométriques des échelles du DUNDRUM. Pour la cohérence interne, la fidélité interjuge et la validité critériée, les articles étaient côtés d'un '+' si les résultats étaient jugés satisfaisant, d'un '–' si les résultats étaient jugés insatisfaisant et d'un '?' si les analyses effectuées ne permettaient pas de se prononcer sur la propriété psychométrique en question. Pour la validité de construit, les résultats sont fournis de manière descriptive: les instruments ayant démontré une corrélation significative et supérieure à 0,3 avec les échelles DUNDRUM sont présentés, et les corrélations supérieures à 0,5 sont identifiées en gras. Les critères utilisés pour cette évaluation sont ceux énumérés dans la section appropriée du rapport.

Tableau 1. Synthèse des propriétés psychométriques du Quatuor DUNDRUM

		Fidélité		Validité	
Instrument	**Références**	**Cohérence interne**	**Fidélité interjuge**	**Validité critériée**	**Validité de construit**
DUNDRUM–1	Flynn, O'Neill, McInerney, et al., (2011)*	+	+, †	+, +	‡
	Flynn, O'Neill et Kennedy (2011)*			+	**DUNDRUM-2**
	O'Dwyer et al., (2011)*				DUNDRUM-3, DUNDRUM-4
	Davoren et al., (2012)*			–	
	Abidin et al., (2013)*	–	+, +		START-S, START-V
	Davoren et al., (2013)*			–	
	Davoren, Byrne, et al., (2015)*			–, –	
	Freestone et al., (2015)	+		–, ?, –	
	O'Neill et al., (2016)			+	
	Eckert et al., (2017)			+	
	Adams et al., (2018)	+			**HCR-20, DUNDRUM-3, DUNDRUM-4**
	Lawrence et al., (2018)			+	
	De Page et al., (2019)			+	‡
	Gulati et al., (2019)			?	
	Habets et al., (2019)	+			
	Jeandarme et al., (2019)			–	
	Jones et al., (2019)		+	+	

		Fidélité		Validité	
Instrument	**Références**	**Cohérence interne**	**Fidélité interjuge**	**Validité critériée**	**Validité de construit**
	Habets et al., (2020)	–	+, †	–, –	HoNOS-Secure
	Kennedy et al., (2020)*	–			
	Williams et al., (2020)			+, +, +	
	Jeandarme et al., (2021)			+	
	Parente et al., (2024)	+			
	Chaffe et al., (2026)			+	
	Hudon et al., (2026)			+	
DUNDRUM–2	Flynn, O'Neill et Kennedy (2011)*	+	+, †	+	**DUNDRUM-1**
	Davoren, Byrne, et al., (2015)*			–, –	
	Freestone et al., (2015)			–	
	O'Neill et al., (2016)			+	
	De Page et al., (2019)			+	‡
	Gulati et al., (2019)			?	
	Jones et al., (2019)		+	+	
	Habets et al., (2020)	–	+, †	–	HoNOS-Secure
	Jeandarme et al., (2026)	–	+, †		
DUNDRUM–3	O'Dwyer et al., (2011)*	+	+, †	+, +	DUNDRUM-1, **DUNDRUM-4, PANSS, GAF, HCR-20, S-RAMM,** CANFOR
	Davoren et al., (2012)*			–, +, +	
	Abidin et al., (2013)*	+		+, +	**START-S, START-V, SAPROF**
	Davoren et al., (2013)*			+	
	Davoren, Hennessy, et al., (2015)*	+		+, +, +, +	**DUNDRUM-3 -SR**
	O'Reilly et al., (2016)*				**PANSS,** ACB, MCCB, **SOFAS**
	Eckert et al., (2017)			+	
	Adams et al., (2018)	+		–, +	**DUNDRUM-1, DUNDRUM-4, HCR-20,** CANFOR
	McCullough et al., (2020)			+, +, ?	
	Lam et al., (2023)			+	
	Parente et al., (2024)	+			
	Pouls et al., (2026)	+		+	
DUNDRUM–4	O'Dwyer et al., (2011)*	+	+, †	+, –	DUNDRUM-1, **DUNDRUM-3, PANSS, GAF, HCR-20, S-RAMM**
	Davoren et al., (2012)*			–, +, +	
	Abidin et al., (2013)*	+	+, +		**START-S, START-V, SAPROF**
	Davoren et al., (2013)*	+			

		Fidélité		Validité	
Instrument	**Références**	**Cohérence interne**	**Fidélité interjuge**	**Validité critériée**	**Validité de construit**
	Davoren, Hennessy, et al., (2015)*		+	+, +, +, +	**DUNDRUM-4 -SR**
	Eckert et al., (2017)			+	
	Adams et al., (2018)		+	–, +	**DUNDRUM-1, DUNDRUM-3, HCR-20**
	McCullough et al., (2020)			+, +, ?	
	Lam et al., (2023)			+	
	Smothers et al., (2023)			+	HCR-20
	Parente et al., (2024)	+			
	Pouls et al., (2026)	+		+	

Note: *L'étude a été menée au Central Mental Hospital où l'instrument a été développé. Notation des résultats (adaptée des directives COSMIN), par étude. Pour la cohérence interne, la fidélité interjuge et la validité critériée: satisfaisant (+) / insatisfaisant (–) / indéterminé (?) selon les critères propres à chaque type d'analyse (détaillés dans les sections appropriées); Pour la fidélité interjuge: † Une analyse de fidélité interjuge a été effectuée au niveau des items. Pour la validité critériée: la multitude de symboles représente la quantité d'analyses effectuées. Pour la validité de construit: le nom des instruments qui corrèlent à 0,3 ou plus sont indiqués; les instruments qui corrèlent à 0,5 ou plus sont indiqués en gras; ‡ = une analyse a été effectuée, sans résultats significatifs.

Vingt-quatre (24) études rapportaient des informations pertinentes quant aux propriétés psychométriques du DUNDRUM-1, incluant des échantillons variant de 13 à 643 participants[3]. La majorité de ces études ont été menées en Irlande (n = 10), mais le DUNDRUM-1 a également été testé en Australie, en Nouvelle-Zélande, en Belgique, en Italie, au Canada, aux Pays-Bas et au Royaume-Uni (Angleterre et Pays de Galles). Neuf (9) études ont été identifiées comme détenant des informations pertinentes quant aux propriétés psychométriques du DUNDRUM-2, incluant des échantillons variant de 13 à 643 participants. Les pays où ces études ont été effectuées incluent la Belgique, le Canada, l'Irlande et le Royaume-Uni (Angleterre). Douze (12) études détenaient des données pertinentes quant aux propriétés psychométriques du DUNDRUM-3 et douze (12) d'entre elles traitaient également du DUNDRUM-4, incluant des échantillons variant de 56 à 327 participants. Ces échelles ont été testées en Australie, en Italie, en Belgique, en Irlande, au Canada, aux Pays-Bas et au Royaume-Uni (Angleterre et Pays de Galles). Puisqu'aucune étude n'utilisait uniquement le DUNDRUM-4, ces deux échelles seront décrites conjointement. La prochaine section de ce rapport se veut une synthèse des informations recueillies dans la littérature consultée. Les sections qui suivent comprennent les informations se rapportent aux propriétés psychométriques investiguées, soit la fidélité, regroupant la cohérence interne et fidélité interjuge, ainsi que la validité, regroupant la validité critériée et la validité de construit.

3.1. Fidélité

La fidélité réfère à la précision et à la constance des mesures obtenues à l'aide d'un instrument. Elle se rapporte à sa capacité à donner des résultats constants ou fiables indépendamment du nombre de fois où il est administré, de ses correcteurs et de son contenu particulier (notion de reproductibilité) (Hogan, 2017). Bien qu'il existe plusieurs méthodes pour évaluer la fidélité d'un instrument, les travaux portant sur la validation des échelles du DUNDRUM se sont penchés sur deux facettes de la fidélité, c'est-à-dire la cohérence interne et la fidélité interjuge.

[3] À noter que puisque dix (10) des 30 études ont été conduites au Central Mental Hospital en Irlande, il pourrait y avoir du chevauchement entre les participants.

COHÉRENCE INTERNE

La cohérence interne d'un instrument réfère à la corrélation qui existe entre tous les items qui le composent (degré d'intercorrélations) et examine à quel point tous les items d'une même échelle mesurent le même concept ou construit (Hogan, 2017). En théorie, plusieurs items mesurant le même construit (p. ex. le « rétablissement » pour le DUNDRUM-4) devraient être fortement et positivement corrélés entre eux. La cohérence interne d'un instrument est traditionnellement mesurée à l'aide de l'alpha de Cronbach (α). Nous baserons notre interprétation des différents indices de fiabilité répertoriés selon les lignes directrices suivantes : $\alpha \geq 0{,}90$ est excellent, $\alpha \geq 0{,}80$ est bon, $\alpha \geq 0{,}70$ est satisfaisant, $\alpha \geq 0{,}50$ est acceptable et $\alpha < 0{,}50$ est inacceptable (Cicchetti, 1994 ; Peterson et al., 1985 ; Taber, 2018).

Les résultats répertoriés à travers les études et leur interprétation respective sont présentés dans le tableau 2. De manière générale, on constate que les échelles portant sur la progression et le rétablissement (DUNDRUM-3 et DUNDRUM-4) performent mieux en termes de cohérence interne que celles relatives au triage (DUNDRUM-1 et -2).

Tableau 2. Cohérence interne des échelles du DUNDRUM

Articles	N	α	Interprétation[a]
DUNDRUM-1[b]			
Flynn, O'Neill, C., McInerney et al., 2011* (Irlande)	316	0,95	Excellent
Abidin et al., 2013* (Irlande)	92	0,60	Acceptable
Freestone et al., 2015 (Angleterre)	195	0,77	Satisfaisant
Adams et al., 2018 (Australie)	327	0,84	Bon
Habets et al., 2019* (Belgique)	50	0,79	Satisfaisant
Habets et al., 2020* (Belgique)	38	0,61	Acceptable
Kennedy et al., 2020 (Irlande)	28	0,67	Acceptable
Parente et al., 2024 (Italie)	193	0,84	Bon
DUNDRUM-2			
Flynn, O'Neill et Kennedy, 2011* (Irlande)	65	0,72	Satisfaisant
Habets et al., 2020 (Belgique)	38	0,26	Inacceptable
Jeandarme et al., 2026 (Belgique)	120	0,40	Inacceptable
DUNDRUM-3			
O'Dwyer et al., 2011* (Irlande)	95	0,91	Excellent
Abidin et al., 2013* (Irlande)	98	0,91	Excellent
Davoren, Hennessy, et al., 2015* (Irlande)	97	0,90	Excellent
Adams et al., 2018 (Australie)	327	0,90	Excellent
Parente et al., 2024 (Italie)	193	0,75	Satisfaisant
Pouls et al., 2026 (Belgique)	82	0.82	Bon
DUNDRUM-4[c]			
O'Dwyer et al., 2011* (Irlande)	95	0,89	Bon
Abidin et al., 2013* (Irlande)	98	0,89	Bon
Davoren, Hennessy, et al., 2015* (Irlande)	97	0,88	Bon
Adams et al., 2018 (Australie)	327	0,89	Bon
Parente et al., 2024 (Italie)	193	0,84	Bon
Pouls et al., 2026 (Belgique)	27	0,83	Bon

Note: *L'article a été mené au Central Mental Hospital. [a]L'interprétation est conforme aux lignes directrices proposées par Cicchetti (1994) et Peterson, Albaum et Beltramini (1985). [b]Pour les articles dont le nom des auteurs est inscrit en italique, le DUNDRUM-1 a été utilisé sans les items liés à la suicidalité (S2 et S4). [c]À l'exception de Parente et ses collègues (2024), aucun autre article relatif à la cohérence interne du DUNDRUM-4 n'a utilisé la nouvelle version incluant l'item lié à l'espoir (R7).

Pour le DUNDRUM-1, bien que la cohérence interne semble plus variable, on note que les études effectuées auprès d'échantillons plus grands obtiennent généralement de meilleurs résultats. On note également que certains items, si supprimés, mènent à une augmentation de l'alpha de Cronbach, particulièrement les items S2 (gravité des comportements auto-agressifs) et S4 (imminence du risque de suicide) (Flynn, O'Neill, McInerney, et al., 2011 ; Freestone et al., 2015 ; Habets et al., 2019). En effet, selon ces études, les éléments relatifs à la suicidalité (S2 et S4) refléteraient un construit distinct de ceux évaluant le risque d'agression ou d'évasion (Freestone et al., 2015). Selon ce point de vue, les considérations liées aux tendances suicidaires d'un individu seraient associées à une facette différente, mais tout de même pertinente au besoin de sécurité thérapeutique.

En ce qui concerne la cohérence interne du DUNDRUM-2, les trois études rapportent des résultats divergents. Tout d'abord, l'étude menée par Flynn, O'Neill et Kennedy (2011) au *Central Mental Hospital* en Irlande a permis de conclure que cet instrument avait une cohérence interne satisfaisante (α = 0,72). Les auteurs rapportent également que le retrait de l'item U6 (priorisation judiciaire) permet d'augmenter l'alpha de Cronbach à 0,81, indiquant que cet item corrèle moins aux autres. Toutefois, ces résultats n'ont pas pu être répliqués dans deux études subséquentes menées par Habets, Jeandarme et Kennedy (2020) et Jeandarme et ses collègues (2026) en Belgique, qui ont obtenu une cohérence interne inacceptable (α = 0,26 et α = 0,40, respectivement). Ces résultats sont nuancés en partie par la faible taille de l'échantillon utilisé dans l'étude de Habets, Jeandarme et Kennedy (2020) (N = 38). Ces derniers soulignent également la possibilité que les items constituant le DUNDRUM-2 s'appliquent difficilement au contexte belge. En somme, les données disponibles ne nous permettent pas de nous prononcer avec certitude sur la cohérence interne du DUNDRUM-2. De plus, les items pourraient donc être conceptualisés comme des construits unitaires représentant chacun une dimension du construit d' « urgence ». À la lumière de cette réalité clinique, il est donc opportun de se questionner sur la pertinence de la cohérence interne pour évaluer la qualité de cet instrument (Habets et al., 2020). En somme, les données disponibles ne nous permettent pas de nous prononcer avec certitude sur la cohérence interne du DUNDRUM-2.

Les échelles DUNDRUM-3 et DUNDRUM-4, quant à elles, présentent de manière générale, des scores très élevés de cohérences internes. Les études ont mis en évidence des alphas allant de 0,75 à 0,91 pour le DUNDRUM-3 et de 0,83 à 0,89 pour le DUNDRUM-4 (Abidin et al., 2013 ; Adams et al., 2018 ; Davoren, Hennessy, et al., 2015 ; O'Dwyer et al., 2011 ; Parente et al., 2024 ; Pouls et al., 2026). Ces données suggèrent que les items qui composent le DUNDRUM-3 et DUNDRUM-4 covarient fortement. Il importe de noter que la majorité des études ont investigué la cohérence interne du DUNDRUM-4 sans l'ajout de l'item R7 (espoir). Soulignons également que la cohérence interne des versions auto-évaluées des échelles DUNDRUM-3 et DUNDRUM-4 n'a pas été investiguée dans les études recensées.

La cohérence interne repose sur le principe qu'une échelle est unidimensionnelle, c'est-à-dire qu'elle mesure un seul construit (Hogan, 2017). Selon cette conceptualisation, la présence du trait causerait la présence ou l'absence des caractéristiques individuelles mesurées par chaque item et la présence du trait devrait donc toujours corréler avec la présence d'une majorité d'items (et vice-versa). Un indice d'intercorrélation élevé indique donc que tous les items mesurent bel et bien un seul et même construit. Bien qu'il puisse sembler a priori que chacune des échelles DUNDRUM mesure techniquement un construit unidimensionnel (la sécurité au triage, l'urgence au triage, la progression dans les programmes et le rétablissement), il pourrait être préférable de considérer ces concepts comme multidimensionnels puisque plusieurs facteurs indépendants les uns des autres peuvent en être la cause. En effet, les auteurs du DUNDRUM **indiquent que les décideurs ne sont pas liés au score final de l'évaluation, car « un item avec une cote élevée pourrait suffire à justifier une admission dans un milieu à sécurité thérapeutique de niveau maximal si indiqué par le contexte du cas précis »** (Kennedy et al., 2016a, p. 16). Ce type d'interprétation est inhérent à l'approche du jugement professionnel structuré. Ainsi, les échelles du DUNDRUM seraient mieux décrites comme

des instruments de mesure formatifs, soit des instruments composés d'un ensemble hétérogène d'items visant chacun à évaluer un facteur qui causerait le construit visé. Par exemple, pour le DUNDRUM-1, il est plausible que les items relatifs à l'historique d'agression envers autrui (S1 et S3) influencent la décision de sécurité au triage indépendamment de ceux mesurant l'historique d'autoagression (S2 et S4). Par conséquent, on ne s'attend pas nécessairement à ce que tous les items corrèlent entre eux. Ces résultats mettent de l'avant l'importance d'interpréter cette propriété psychométrique en relation avec d'autres indices de fiabilité, tels que la fidélité interjuge (Cicchetti, 1994).

FIDÉLITÉ INTERJUGE

La fidélité interjuge nous informe sur le degré d'entente entre différents évaluateurs lors de l'évaluation auprès d'une même personne, au même moment et dans les mêmes conditions, indiquant dans quelle mesure l'instrument est interprété de la même façon par différents évaluateurs (Hogan, 2017). Les scores obtenus à la suite de l'administration des échelles du DUNDRUM par deux évaluateurs indépendants pour un même patient devraient être corrélés fortement et positivement. Il existe plusieurs façons d'évaluer la fidélité interjuge et, d'ailleurs, le choix des indices statistiques utilisés varie considérablement d'une étude à l'autre (p. ex. Kappa simple/pondéré, corrélations intra-classe, corrélation de Spearman, corrélation de Pearson, etc.). De plus, il ne semble pas y avoir de consensus quant à une méthode à privilégier pour évaluer la fidélité interjuge des échelles du DUNDRUM. Alors que certaines études rapportent le degré d'accord entre les scores globaux, d'autres s'intéressent à celui-ci pour chaque item individuellement. Toutefois, selon Fleiss (1999), les deux indices à privilégier sont le Kappa pondéré et la corrélation intra-classe (ICC). Cicchetti (1994) propose des critères d'interprétation identiques pour ces deux mesures : ≥ 0,75 est excellent, ≥ 0,60 est bon, ≥ 0,40 est acceptable et < 0,40 est inacceptable. L'hétérogénéité des méthodes utilisées limite toute tentative de comparaison systématique des résultats à travers les études. Toutefois, il est possible de brosser un portrait global à partir des études de validation de chaque échelle.

Les premières études de validation, réalisées par les auteurs de l'instrument, évaluant la fidélité interjuge des échelles du DUNDRUM ont été effectuées au *Central Mental Hospital*. Pour chacun des onze items du DUNDRUM-1, Flynn, O'Neill, McInerney et al. (2011) ont observé une corrélation élevée ($r > 0,75$) entre les scores attribués par deux évaluateurs indépendants sur un échantillon de 18 participants. Ne soulevant aucune différence significative entre la performance individuelle de chaque item, ces auteurs en concluent que le DUNDRUM-1 démontre une bonne fidélité interjuge (Flynn, O'Neill, McInerney, et al., 2011).

Deux études subséquentes évaluant cette fois-ci le DUNDRUM-2 décrit dans quelle mesure les scores fournis par deux cliniciens variaient lors de l'évaluation d'un échantillon de 24 participants (Flynn, O'Neill et Kennedy, 2011) et de 30 participants (Jeandarme et al., 2026) pour chacun des six items de cette échelle. Dans l'étude de Flynn et ses collègues (2011), les items U1 (situation initiale) et U4 (considérations humanitaires) ont présenté les plus grandes divergences entre les cliniciens, les évaluations différant de plus d'un point dans 25 % des cas. À l'inverse, l'item U6 (priorisation judiciaire) a présenté la meilleure concordance, avec une seule évaluation différant de plus d'un point. De manière similaire, dans l'étude de Jeandarme et al. (2026), U1 présentait le coefficient de concordance le plus faible, tandis que U6 présentait une concordance parfaite. Les auteurs s'appuient sur ces résultats pour conclure que le DUNDRUM-2 démontre une bonne fidélité interjuge (Flynn, O'Neill et Kennedy, 2011 ; Jeandarme et al., 2026).

Une étude comparative menée par Habets et al. (2020) en Belgique, évalue conjointement la fidélité interjuge du DUNDRUM-1 et du DUNDRUM-2 (voir tableau 3). L'objectif était de comparer les propriétés psychométriques de ces échelles avec celles d'un autre instrument mesurant le besoin de sécurité thérapeutique : le HoNOS-secure (de l'anglais *Health of the Nation Outcome Scale for Users of Secure and Forensic Services*). D'abord, les chercheurs ont observé que le niveau d'entente était généralement plus élevé pour les items du DUNDRUM-1 que pour ceux du DUNDRUM-2 et du HoNOS-secure. Le seul item problématique pour le DUNDRUM-1 était S11 (processus judiciaire),

qui a obtenu un niveau d'entente moyen. Le DUNDRUM-2 présentait quant à lui trois items ayant obtenu de très faible niveau d'entente : U2 (santé mentale), U5 (considérations systémiques) et U6 (priorisation judiciaire). Finalement, ces auteurs soulignent que le niveau d'entente est meilleur entre les cliniciens lorsque ceux-ci ont suivi une formation portant sur l'utilisation des échelles du DUNDRUM. En somme, ces résultats soulignent, d'une part, la complexité de mesurer le niveau d'urgence au triage, surtout lorsque cet instrument est construit pour s'adapter à des populations de patients issus de divers milieux, en tenant compte non seulement de l'état de la personne, mais également des différents contextes juridiques (présence d'une ordonnance de la cour) et systémique (disponibilité des lits) qui peuvent jouer sur la décision de prioriser un placement. D'autre part, ces résultats démontrent également l'importance d'une formation de qualité pour soutenir l'implantation du DUNDRUM dans le réseau de la psychiatrie légale pour assurer des évaluations justes.

Tableau 3. Résultats de Habets et al. (2020) relatifs à la fidélité interjuge des items des DUNDRUM-1 et -2

Item	N	ICC[a]	Interprétation[b]
DUNDRUM-1			
S1 : Gravité des gestes violents	37	0.86	Excellent
S2 : Gravité des comportements auto-agressifs	38	0.88	Excellent
S3 : Imminence du risque de violence relié au trouble mental	37	0.74	Bon
S4 : Imminence du risque de suicide	38	0.69	Bon
S5 : Besoins spécialisés en psychiatrie légale	36	0.66	Bon
S6 : Évasion et fugue	38	0.80	Excellent
S7 : Contrôle des accès	36	0.77	Excellent
S8 : Considérations envers les victimes et confiance du public	38	0.70	Bon
S9 : Besoins complexes reliés au risque de violence	38	0.72	Bon
S10 : Comportement en établissement	36	0.85	Excellent
S11 : Processus judiciaire	38	0.52	Acceptable
DUNDRUM-2			
U1 : Location	38	0.61	Bon
U2 : Santé mentale	38	0.18	Inacceptable
U3 : Prévention du suicide	37	0.74	Bon
U4 : Considérations humanitaires	36	0.61	Bon
U5 : Considérations systémiques	23	0.27	Inacceptable
U6 : Priorisations judiciaires	37	0.39	Inacceptable

Note: [a]Corrélation intra-classe de type A,1 (absolute agreement, single measurement). [b]L'interprétation est conforme aux lignes directrices de Cicchetti (1994).

À notre connaissance, l'étude de O'Dwyer et al. (2011) constitue la seule démonstration empirique de la fidélité interjuge du DUNDRUM-3 et du DUNDRUM-4. Ces auteurs rapportent que, pour chaque item du DUNDRUM-3, la proportion des participants (N = 21) qui se sont vu attribuer exactement le même score allait de 62 % pour l'item P4 (Problèmes de comportement) à 90 % pour le P1 (Santé physique). Pour le DUNDRUM-4, cette proportion variait de 62 % pour l'item R4 (Congés et sorties) à 86 % pour le R5 (Risque dynamique). Ces résultats semblent démontrer un niveau d'entente relativement satisfaisant pour ces échelles (O'Dwyer et al., 2011). Bien que ces résultats soient prometteurs, il ne faut pas passer sous silence le fait qu'il existe des biais inhérents associés au processus de validation par les auteurs de l'instrument. Il sera donc important que de futures études évaluent la fidélité interjuge du DUNDRUM-3 et du DUNDRUM-4 en dehors du Central Mental Hospital, où l'expérience des évaluateurs assure un degré d'entente plus élevé.

Pour conclure cette section sur la fidélité interjuge, nous rapportons qualitativement les résultats de deux études concernant la concordance entre les évaluations DUNDRUM-3 et DUNDRUM-4 complétées par des cliniciens et celles complétées par des patients (Davoren et al., 2015, Lam et al., 2023). Le constat principal de ces études est que les patients semblent s'évaluer de manière plus optimiste que les cliniciens, affichant des scores plus élevés pour leur processus de rétablissement (c.-à-d. des scores plus faibles de dangerosité). Les évaluations DUNDRUM-3 et DUNDRUM-4 auto-évaluées étaient en corrélation avec les évaluations des cliniciens, mais étaient systématiquement inférieures. Notons également que la différence entre les scores du clinicien et ceux auto-évalués semblait diminuer à mesure que les patients approchaient de leur sortie. Ce dernier constat résulte vraisemblablement d'une variabilité réduite des scores auto- évalués en général, ceux-ci étant plus concentrés vers des scores faibles. En effet, les scores auto- évalués étaient plus similaires à travers les différents niveaux de sécurités thérapeutiques comparativement aux scores des cliniciens. Par conséquent, les évaluations auto-évaluées étaient moins précises en ce qui a trait à la prédiction des décisions de placements.

3.2. Validité

La validité d'un instrument réfère à sa capacité à mesurer ce pourquoi il a été conçu pour mesurer (Hogan, 2017). En ce qui concerne le DUNDRUM, à l'exception du DUNDRUM-2 qui évalue l'urgence du triage, les construits mesurés par le Quatuor DUNDRUM correspondent globalement au besoin de sécurité thérapeutique (tel que représenté par le système de cotation allant de 0 à 4 selon le niveau de sécurité recommandé). Il existe différentes manières de vérifier le degré de validité d'un instrument, mais celles-ci sont généralement regroupées en trois grands types de validité (validité de contenu, validité critériée et validité de construit). Dans le cadre de cette revue, nous nous attarderons aux évidences identifiées dans la littérature recensée, c'est-à-dire la validité critériée ainsi que la validité de construit.

VALIDITÉ CRITÉRIÉE

La validité critériée réfère à la capacité d'un instrument à distinguer les personnes évaluées entre elles et au degré auquel les scores d'un instrument reflètent adéquatement un critère externe (Hogan, 2017). Dans le cadre du DUNDRUM, le niveau de sécurité thérapeutique réel d'un patient pourrait constituer ce critère externe pour les échelles DUNDRUM-1, -3 et -4. Ainsi, il serait possible d'émettre l'hypothèse qu'un patient est placé au niveau de sécurité approprié et de valider si l'instrument arrive aux mêmes conclusions. Lorsque ce critère se définit dans le temps et succède à l'administration de l'instrument, on parlera de validité prédictive, qui est une forme de validité critériée (p. ex. le score de l'instrument administré à l'admission prédit les comportements agressifs dans les mois subséquents). En ce sens, une démonstration de validité prédictive est également une démonstration de validité critériée. Nous parlerons donc de validité critériée dans ce rapport, en englobant la notion de validité prédictive.

Pour évaluer systématiquement la validité critériée, des critères externes (ainsi que les analyses correspondantes) ont été compilés et organisés de façon à regrouper les données issues des articles selon chaque échelle du DUNDRUM. Les critères de signification sont ceux proposés par le guide COSMIN pour chaque type d'analyse (Prinsen et al., 2018). Le tableau 4 résume les coefficients et les critères de signification pour chaque échelle du DUNDRUM.

Il existe plusieurs techniques statistiques qui permettent d'estimer la validité critériée d'un test, incluant la corrélation et, précisément pour la validité prédictive, l'aire sous la courbe (Area Under the Curve, AUC). Prinsen et al. (2018) décrivent comme adéquates une corrélation ≥ à 0,70 avec le critère externe, ou encore une aire sous la courbe

≥ à 0,70 selon le critère en question (p. ex. décision d'admission de type « oui/non »). Le test ANOVA et le Test-t (incluant leurs équivalents non paramétriques) comparant différents niveaux de sécurité thérapeutique ont également été considérés comme une démonstration de validité critériée, dans la mesure où les scores moyens des groupes de patients se distribuent de manière cohérente avec l'attribution du niveau de sécurité (c.-à-d. l'ordre des moyennes correspond à l'ordre des groupes en termes de sécurité thérapeutique).

Tableau 4. Critère externe pour vérifier la validité critériée des échelles du DUNDRUM

Coefficient	Critères externe	Exemples
ANOVA/Kruskal-Wallis	Niveau de sécurité	Décision de placement dans un hôpital psychiatrique Décision de placement dans un système de santé Décision de libération dans un hôpital psychiatrique Décision de transfert dans un hôpital psychiatrique Recommandation de placement dans un système de santé (comité) Recommandation d'hospitalisation dans un institut sécurisé (psychiatre) Recommandation d'hospitalisation dans un institut sécurisé (travailleur social)
	Priorité d'admission	Décision d'admission dans un hôpital psychiatrique
	Autre	Niveau d'accès à des privilèges de sortie
Test-t/Mann-Withney U	Niveau de sécurité	Décision de placement dans un système de santé
	Autre	Achèvement du traitement
AUC	Niveau de sécurité	Décision de libération dans un hôpital psychiatrique Décision de transfert dans un hôpital psychiatrique
	Violence	Violence envers autrui Violence envers soi-même
	Autre	Isolement Infraction dans l'année suivant la libération Infraction dans les 6 mois suivant la libération

DUNDRUM-1

Dix-neuf (19) des études recensées ont fourni des données pertinentes pour évaluer la validité critériée du DUNDRUM-1. Certaines de ces études ont comparé les résultats du DUNDRUM-1 à la décision finale d'admission et d'attribution du niveau d'encadrement/sécurité des patients dans différentes unités psychiatriques hiérarchisées selon leur niveau de sécurité thérapeutique (Flynn, O'Neill, McInerney, et al., 2011 ; Freestone et al., 2015 ; Jeandarme et al., 2019 ; Lawrence et al., 2018). Dans ces études, les groupes de patients obtiennent des scores significativement différents, ce qui suggère une certaine capacité de l'instrument à distinguer les patients selon leur niveau de sécurité. Par exemple, dans l'étude de Flynn, O'Neill, McInerney, et al. (2011), les scores moyens du DUNDRUM-1 étaient cohérents avec la décision de placement pour chacun des groupes de patients.

En revanche, les résultats n'ont pas été concluants dans d'autres études. Freestone et al. (2015) ont observé que le groupe de patients non-admis obtenait une moyenne légèrement trop élevée, correspondant davantage à un besoin de sécurité thérapeutique de faible à moyen. En Belgique, Habets et al. (2019) n'ont pas observé de variations des scores DUNDRUM-1 en fonction du niveau de sécurité. De manière similaire, Jeandarme et al. (2019) ont constaté que, même si les scores du DUNDRUM-1 concordaient lorsque les tribunaux recommandaient un placement dans un établissement à sécurité élevé, il existait une très faible correspondance entre le score à l'échelle DUNDRUM-1 et le niveau de sécurité recommandé. Les auteurs suggèrent que, considérant l'absence de critères objectifs justifiant les décisions de placement émis par les tribunaux, il est possible de remettre en question la légitimité du jugement de la cour à titre de standard

de comparaison pour la validité du DUNDRUM-1 (Jeandarme et al., 2019). Enfin, l'étude de Lawrence et al. (2018) obtient des différences significatives entre les différents groupes de patients. Par contre, l'information rapportée dans l'article ne permettait pas de se prononcer quant à la correspondance entre leur score et le niveau de sécurité thérapeutique auquel ils étaient assujettis.

Dans une étude d'Abidin et al. (2013), plusieurs instruments de jugement professionnel structuré, dont le DUNDRUM-1, ont été comparés dans leur capacité à prédire le risque de violence envers soi-même et autrui sur une période de 6 mois. Les scores au DUNDRUM-1 ont permis de prédire les comportements violents envers autrui (AUC = 0,74 [0,60 – 0,88]) et envers soi-même (AUC = 0,71 [0,50 – 0,93]). L'étude de Freestone et al. (2015) quant à elle a démontré que les scores au DUNDRUM-1 ne permettaient pas de prédire le risque de récidive de violence sur une période de 12 mois chez les patients référés et non-admis. D'autres études ont vérifié l'association des scores du DUNDRUM-1 avec divers critères externes définis comme des indicateurs cliniques pertinents. Notons l'étude de Flynn, O'Neill, McInerney, et al. (2011), qui a démontré que, parmi les patients n'ayant pas été admis au Central Mental Hospital, ceux ayant connu un suivi psychiatrique dans les trois mois suivant ce refus d'admission obtenaient en moyenne des scores significativement plus élevés que ceux n'ayant pas eu un tel suivi. Chaffe et ses collègues (2026), quant à eux, soulignent que le DUNDRUM-1, et particulièrement les items liés à la gravité et l'imminence des actes de violence (S1 et S3), permettaient de prédire l'admission dans un *extended forensic care and rehabilitation pathway* (EFCR) en Nouvelle-Zélande. Williams et al. (2020), quant à eux, rapportent que les patients dont l'admission était recommandée par un travailleur social forensique avaient des scores plus élevés au DUNDRUM-1 que ceux non recommandés. Pour conclure, dans une étude s'intéressant à stabilité longitudinale des recommandations produites par le DUNDRUM, Hudon et ses collègues (2026) soulignent que sur un suivi d'au moins 12 mois, la majorité des participants sont restés dans les limites de leur niveau de soins initialement recommandé à l'aide du DUNDRUM-1 tandis que 15 % ont eu besoin d'un changement de niveau de sécurité thérapeutique. Cette étude mentionne que le score DUNDRUM moyen lié à la violence est le seul facteur prédictif significatif des variations du niveau de soins.

En somme, les résultats recensés concernant la validité critériée du DUNDRUM-1 permettent de conclure que l'échelle du DUNDRUM-1 a une **bonne capacité pour distinguer entre les groupes provenant de différents niveaux de sécurité, bien qu'il soit difficile de prédire le niveau de sécurité dans lequel se trouve un patient au moment de l'administration de l'instrument. Plutôt que de mettre en doute la validité de l'instrument, cette limite reflète plutôt l'importance d'utiliser un instrument comme le DUNDRUM pour les décisions de placement des patients.** En effet, les différences entre le score total de l'échelle ainsi que le niveau de placement sans l'utilisation préalable d'un tel instrument est probablement attribuable au fait qu'il existe un plus grand degré de subjectivité dans les évaluations qui ne sont pas guidées par un instrument de jugement structuré (Hart et al., 2016).

DUNDRUM-2

Huit (8) des études recensées ont investigué la validité critériée du DUNDRUM-2. Parmi ces études, Flynn, O'Neill et Kennedy (2011) ont entre autres démontré que, dans un contexte où le nombre de demandes d'admission excède la capacité d'un hôpital psychiatrique, les scores du DUNDRUM-2 permettaient de prédire quels patients seraient priorisés à l'admission. En effet, parmi les 65 patients sur la liste d'attente du Central Mental Hospital, les 38 patients ultimement admis avait en moyenne des scores plus élevés que ceux à qui n'ont pas été offerte l'admission. Ces auteurs notent également que la distinction entre les groupes était plus marquée au moment du retrait de la liste d'attente, comparativement au moment de l'ajout sur la liste d'attente. Les études menées par Freestone et al. (2015) et Habets et al. (2020) cherchaient toutes deux à démontrer l'existence d'un lien entre les scores du DUNDRUM-2 et la vitesse d'admission, posant l'hypothèse qu'un niveau d'urgence plus élevé devrait être associé à un temps d'attente moindre au moment de l'admission. Freestone et al. (2015) ne sont pas parvenus à démontrer une corrélation significative entre le temps d'admission et les scores du DUNDRUM-2 (corrélation de

Pearson : r = −0,33 ; NS ; N = 175). Habets et al. (2020) sont arrivés à une conclusion similaire, n'observant pas de corrélation significative entre le DUNDRUM-2 et la durée du séjour en prison pour un échantillon de 100 détenus référés en établissement psychiatrique (corrélation de Spearman : r = −0,02 ; NS). L'étude de O'Neill et al. (2016), quant à elle, s'intéressait plutôt à l'association entre le DUNDRUM-2 et le niveau de sécurité thérapeutique dans lequel se trouvait la personne, interprété comme un autre indicateur clinique pertinent. Ces auteurs rapportent des différences significatives entre les scores des patients référés à différents niveaux de sécurité thérapeutique. En somme, ces résultats préliminaires indiquent que le DUNDRUM-2 arrive partiellement à distinguer les patients nécessitant des soins urgents et d'autres études seraient nécessaires. De plus la gestion des places disponibles est intimement liée à un écosystème de services souvent segmentés en fonction de plusieurs critères qui sont extérieurs à celui d'urgence (p. ex. région géographique, juridiction, ordonnance émise, etc.)

DUNDRUM-3 et DUNDRUM-4

Dix (10) des études recensées ont rapporté des résultats pertinents concernant la validité critériée du DUNDRUM-3, tandis que onze (11) ont rapporté des résultats pertinents pour le DUNDRUM-4. Parmi celles-ci, Davoren et al. (2013) ont démontré que les scores du DUNDRUM-3 et DUNDRUM-4 permettent de prédire la décision de libération conditionnelle. Toutefois, les résultats révèlent que le groupe de patients ayant bénéficié d'une libération conditionnelle obtenait en moyenne des scores sensiblement trop élevés (selon les recommandations DUNDRUM) pour bénéficier d'une telle décision (μ_{D3} = 0,56 [0,18 – 0,94] ; μ_{D4} = 0,73 [0,40 – 1,07]). Une étude similaire menée dans un hôpital psychiatrique hautement sécurisé a démontré que la décision de transférer un patient vers un établissement moyennement sécurisé se reflétaient dans les scores des DUNDRUM-3 et DUNDRUM-4 (McCullough et al., 2020). Par contre, tout comme pour l'étude de Davoren et al. (2013), les patients transférés obtenaient en moyenne des scores correspondant à un niveau élevé de sécurité thérapeutique (μ_{D3} = 3,14 [3,00 – 3,29] ; μ_{D4} = 3,24 [3,07 – 3,41] et ce, autant pour le DUNDRUM-3 que pour le DUNDRUM-4.

Enfin, six études ont démontré que le DUNDRUM-3 et DUNDRUM-4 permettent de distinguer des patients localisés à différents niveaux de sécurité (Adams et al., 2018 ; Davoren et al., 2012 ; Davoren, Hennessy, et al., 2015 ; Lam et al., 2023 ; O'Dwyer et al., 2011, Pouls et al., 2026). Cependant, pour chacune de ces études, l'analyse de correspondance entre la cote DUNDRUM et le placement se révèle non concluante. On observe un écart entre la cote au DUNDRUM et l'encadrement sécuritaire où se trouvent les patients au moment de la collecte de données. En effet les patients évoluaient dans des unités dont l'encadrement était plus faible que ce qu'indiquait le DUNDRUM. En somme, les DUNDRUM-3 et DUNDRUM-4 ont été utilisés avec succès dans plusieurs études et permettent de différencier différents groupes de patients selon leur niveau de sécurité. Toutefois, il semblerait que ces échelles surestiment généralement le besoin de sécurité thérapeutique des patients. Ce résultat nous rappelle que l'attribution d'un niveau d'encadrement ne repose pas uniquement sur le DUNDRUM-3 et DUNDRUM-4. Le DUNDRUM-1 contient des informations spécifiques à cette décision. Les résultats doivent donc être considérés comme intéressants, mais incomplets. Il serait pertinent que la recherche se penche sur des variables permettant de tester l'apport unique des échelles DUNDRUM-3 et DUNDRUM-4.

VALIDITÉ DE CONSTRUIT

Une autre forme de validité utilisée est la validité de construit (ou validité concourante ou validité convergente/divergente). Dans ce cas, on se base sur le construit que l'instrument prétend mesurer pour émettre diverses hypothèses sur le degré auquel l'instrument corrèle avec d'autres instruments mesurant des construits similaires, liés ou non liés. Par exemple, un chercheur pourrait émettre l'hypothèse qu'un instrument qui mesure un construit similaire (p. ex. le risque de violence) devrait être corrélé avec le construit mesuré (p. ex. sécurité thérapeutique). Dans ce cas, la validation de l'instrument examinerait le niveau de corrélation entre les scores des deux instruments pour une même personne. La validité de construit (aussi nommée validité convergente ou divergente) se mesure en comparant un instrument à des construits qui sont jugés similaires, liés ou inversement liés.

Chaque auteur émet ses propres hypothèses quant aux construits qui peuvent être utiles pour valider une échelle. Selon le guide COSMIN, pour évaluer systématiquement la validité de construit d'un instrument et assurer la réplicabilité du processus d'extraction, des tests d'hypothèses génériques doivent être définis au préalable (Prinsen et al., 2018). Prinsen et al. (2018) recommandent que les hypothèses soient groupées selon si elles mesurent une corrélation avec un construit similaire (degré de signification proposé : $r \geq ,50$ ou $r \leq -,50$), lié mais non similaire (degré de signification proposé : $,50 > r \geq ,30$ ou $-,50 < r \leq -,30$) ou non lié (degré de signification proposé : $-,30 < r < ,30$). La validité de construit de l'instrument évalué est jugée satisfaisante si plus de 50 % des hypothèses identifiées sont confirmées par les résultats obtenus dans l'étude.

À l'issue de la phase d'extraction, 136 corrélations ont été identifiées comme pertinentes à l'examen de la validité de construit. En incluant le Quatuor DUNDRUM, les corrélations recensées portent sur 11 instruments regroupant chacun de une à huit mesure(s) (et/ou sous-échelles), pour un total de 41 mesures. Durant la rédaction de ce manuscrit, il est devenu apparent que l'approche de vérification proposée par les directives COSMIN dépasse le cadre de la présente revue. En effet, cette méthode nécessite de porter un jugement sur la similarité des échelles DUNDRUM avec chacune des mesures recensées. Les directives COSMIN recommandent également d'ajuster les seuils d'interprétation de corrélation (± ,30 et ± ,50) selon les tailles d'effet habituelle du domaine d'étude. En l'absence de critères objectifs pour déterminer les niveaux de similarité et les seuils d'interprétation de corrélation, il a été jugé préférable d'adopter une approche descriptive. Ainsi, le tableau 4 offre une synthèse qualitative de l'ensemble des corrélations recensées pour chaque échelle DUNDRUM. Les corrélations positives et négatives sont identifiées par « + » et d'un « - » respectivement et la force des corrélations est indiquée à l'aide de l'astérisque (pas d'astérisque : $-,30 < r < ,30$, « * » : $,50 > r \geq ,30$ ou $-,50 < r \leq -,30$, « ** » : $r \geq ,50$ ou $r \leq -,50$). Les corrélations non significatives ($p > ,05$) sont identifiées par « NS ». Enfin, les corrélations mesurées dans un échantillon constitué exclusivement d'hommes ou de femmes sont identifiées en bleu et en rouge respectivement.

Les construits

Divers instruments ont été recensés pour examiner la validité de construit des différentes échelles du DUNDRUM. Les sous-échelles du DUNDRUM ont notamment été corrélées entre elles, puisqu'il est postulé que le niveau de sécurité requis (D-1), l'urgence au triage (D-2), la progression dans les programmes (D-3) et le rétablissement (D-4) représentent des concepts similaires ou étroitement liés. D'autres instruments de jugement clinique structuré ont également été utilisés comme étalons de comparaison afin de soutenir la validité de construit, en reflétant soit le niveau de besoins, le risque (violence, suicide, victimisation, etc.), soit les facteurs de protection des individus. Par exemple, le *Camberwell Assessment of Need — Forensic Version* (CANFOR) vise à identifier les besoins particuliers des patients en psychiatrie légale (25 domaines couvrant les besoins de base, fonctionnels, de santé et propres au contexte médico-légal). Il distingue les besoins rencontrés (« no needs ») et non rencontrés (« unmet needs »), produisant ainsi un profil de besoins plutôt qu'un score de sévérité. Le *Short-Term Assessment of Risk and Treatability* (START) a pour objectif d'évaluer le risque de comportements problématiques à court terme (violence, automutilation, suicide, négligence, victimisation) à partir de 20 items considérés à la fois comme vulnérabilités et comme forces. Le HCR-20 est un autre instrument de jugement clinique structuré, développé pour évaluer le risque de violence sur la base de 20 items regroupés en facteurs historiques (p. ex. antécédents de violence, troubles de la personnalité), cliniques (p. ex. symptômes, impulsivité) et de gestion (p. ex. soutien, conformité au traitement). De son côté, le *Structured Assessment of Protective Factors for Violence Risk* (SAPROF) est complémentaire au HCR-20 et centré sur l'évaluation des facteurs de protection. Il comprend 17 items couvrant les domaines internes (p. ex. intelligence, empathie), motivationnels (p. ex. attitudes prosociales, engagement dans le traitement) et externes (p. ex. réseau social, soutien). Enfin, le *Suicide Risk Assessment and Management Manual* (S-RAMM) est destiné à évaluer et à gérer le risque suicidaire. Il comporte 21 items regroupés en trois sections (facteurs historiques, cliniques et de gestion), chacun coté de 0 (absent) à 2 (présent), l'évaluation finale reposant sur le jugement clinique global.

D'autres instruments mobilisés mesurent plutôt des construits associés au fonctionnement global ou cognitif de la personne. L'*Anticholinergic Cognitive Burden scale* (ACB) quantifie l'exposition cumulative des patients à des médicaments anticholinergiques (p. ex. antipsychotiques, antidépresseurs tricycliques), chaque molécule étant cotée de 1 (effet possible) à 3 (effet marqué). Un score total de 3 ou plus indique déjà un risque accru de déclin cognitif, de confusion et de chute. Le *Global Assessment of Functioning* (GAF), introduit dans le DSM-III, évalue de façon globale le fonctionnement psychologique, social et professionnel d'un individu sur une échelle allant de 0 à 100. Le *MATRICS Consensus Cognitive Battery* (MCCB) est un instrument standardisé conçu pour évaluer le fonctionnement cognitif des personnes atteintes de schizophrénie ou d'autres troubles psychotiques. Il couvre sept domaines cognitifs (vitesse de traitement, attention, mémoire de travail, apprentissage verbal et visuel, raisonnement et cognition sociale), avec des scores composites normalisés. La *Positive and Negative Syndrome Scale* (PANSS) est utilisée pour évaluer la sévérité des symptômes de schizophrénie. Elle comprend 30 items répartis en trois sous-échelles (symptômes positifs, symptômes négatifs et psychopathologie générale), chacun coté de 1 (absent) à 7 (extrême), pour un total de 30 à 210 points. L'échelle *Social and Occupational Functioning Assessment Scale* (SOFAS) mesure le niveau de fonctionnement social et professionnel indépendamment de la sévérité symptomatique, sur une échelle de 0 à 100. Contrairement au GAF, le SOFAS se concentre exclusivement sur le fonctionnement, sans intégrer directement les symptômes psychiatriques. Enfin, le *Health of the Nation Outcome Scales — Secure* (HoNOS-S) est une version adaptée du HoNOS pour les milieux sécuritaires. Il évalue la sévérité des problèmes de santé mentale et de fonctionnement social à travers 12 items cotés de 0 (pas de problème) à 4 (problème sévère), pour un score total allant de 0 à 48.

Tableau 5. Synthèse qualitative des corrélations recensées

	Résultats			
Instrument et mesure	**D1**	**D2**	**D3**	**D4**
ACB				
Score total			+*	
CANFOR				
Total needs	NS, +**		+*, +*, +**	+*, +*, +**
Met needs			NS	NS
Unmet needs	+*, NS		+**, +**, +*	+**, +**, NS
Total needs (auto-évalué)			NS	NS
Met needs (auto-évalué)			NS	NS
Unmet needs (auto-évalué)	+*	+		
DUNDRUM				
D1		+**, +	+*, +**, +**	+*, +**, +**
D2	+**, +			
D3	+*, +**, +**		+**	+**, +**, +**
D4	+*, +**, +**		+**, +**, +**	
D3 (auto-évalué)			+**	
D4 (auto-évalué)				+**
GAF				
Score total			–**	–**
HCR-20				
Score total	+**, +**		+**, +**, +**	+**, +**, +**
Score dynamique			+**	+**
Historical	NS, +*, +**		+*, +*, +**	+*, +*, +**, NS
Clinical	NS, +**, +**		+**, +**, +**	+**, +**, +**,+**
Risk Management	+*, +**		+**, +*, +**	+**, +*, +**, +*

	Résultats			
Instrument et mesure	**D1**	**D2**	**D3**	**D4**
HoNOS-S				
Score total	+*, NS	+*, NS		
Security	+*	+**		
Behavior	NS	NS		
Impairment	NS	NS		
Symptoms	NS	NS		
Social	NS	NS		
General12	NS	NS		
General13	NS	NS		
MCCB				
Score total			-*	
PANSS				
Score total			+**, +**	+**
Score de symptômes positifs			+**	+**
Score de symptômes négatifs			+**	+*
Score de symptômes généraux			+**	+*
SAPROF				
Score total	–		–**	–**
SOFAS				
Score total			–**	
S-RAMM				
Score total			+**	+**
Score dynamique			+**	+**
Background			+	NS
Current			+**	+**
Futur			+*	+**
START				
Strengths	–*		–**	–**
Vulnerabilities	+*		+**	+**

Note: NS = Corrélation non-significative (p > ,05) ; +/– = Corrélation (significative) positive ou négative respectivement ; * Taille d'effet modérée (,50 > *r* ≥ ,30 ou –,50 < *r* ≤ –,30) ; ** Taille d'effet élevée (*r* ≥ ,50 ou *r* ≤ – ,50) ; Indices en bleu ou en rouge : corrélation mesurée dans un échantillon composé exclusivement d'hommes ou de femmes respectivement ; D1, D2, D3, et D4 réfèrent aux échelles DUNDRUM ; les références relatives aux corrélations rapportées ci-dessus sont disponibles sur demande.

DUNDRUM-1

Sept (7) études ont rapporté des corrélations pertinentes à l'évaluation de la validité de construit du DUNDRUM-1. Notons tout d'abord que le DUNDRUM-1 semble corréler positivement avec les autres échelles DUNDRUM, et ce, autant chez les hommes que chez les femmes (Adams et al., 2018 ; Flynn, O'Neill et Kennedy, 2011 ; Habets et al., 2020). Chez les femmes, le nombre total de besoins identifiés par le CANFOR (Total needs) corrèle positivement et fortement avec le DUNDRUM-1, tandis que chez les hommes, on trouve plutôt une corrélation positive et modérée avec le nombre total de besoins non satisfaits (Unmet needs) (Adams et al., 2018). La majorité des résultats recensés indiquent également que le DUNDRUM-1 corrèle positivement avec toutes les mesures et sous-échelles du HCR-20 sans effet important associé au genre (Adams et al., 2018 ; Flynn, O'Neill et Kennedy, 2011). Pour le HoNOS-S, des corrélations positives et modérées ont été observées avec le score total (dans un échantillon d'hommes) et la sous-échelle Security (De Page et al., 2019 ; Habets et al., 2020). Enfin, l'étude d'Abidin et al. (2013) rapporte une corrélation faible avec le SAPROF et des corrélations modérées avec les deux sous-échelles du START.

Dans l'ensemble, ces résultats suggèrent que le DUNDRUM-1 présente une validité de construit acceptable, puisqu'il tend à corréler avec des mesures établies de besoins, de risques de violence et de problèmes liés à la sécurité, ainsi qu'avec les autres échelles du DUNDRUM, ce qui appuie sa pertinence pour évaluer le niveau de sécurité thérapeutique requis lors du triage.

DUNDRUM-2

Trois (3) articles ont rapporté des corrélations pertinentes à l'évaluation de la validité de construit du DUNDRUM-2. Ce dernier semble corréler positivement avec le DUNDRUM-1, avec une force variant de faible à élevée d'une étude à l'autre (Flynn, O'Neill et Kennedy, 2011 ; Habets et al., 2020). Les résultats recensés indiquent également que le DUNDRUM-2 corrèle positivement avec le score total et la sous-échelle Security du HoNOS-S (De Page et al., 2019 ; Habets et al., 2020). Dans l'ensemble, ces résultats suggèrent que le DUNDRUM-2 possède une validité de construit acceptable, puisqu'il tend à corréler avec le DUNDRUM-1 et avec des mesures établies de problèmes liés à la sécurité, appuyant son utilité pour évaluer l'urgence au triage.

DUNDRUM-3 et DUNDRUM-4

Cinq (5) articles ont rapporté des corrélations pertinentes à l'évaluation de la validité de construit des DUNDRUM-3 et DUNDRUM-4. En premier lieu, soulignons que ces deux échelles semblent fortement corrélées (Adams et al., 2018 ; O'Dwyer et al., 2011). De plus, les résultats comprennent des corrélations positives de force modérée à élevée avec le DUNDRUM-1 pour ces deux échelles (Adams et al., 2018 ; O'Dwyer et al., 2011). Le nombre total de besoins identifiés à l'aide du CANFOR (Total needs) et le nombre de besoins non satisfaits (Unmet needs) corrèlent positivement avec les DUNDRUM-3 et DUNDRUM-4 (Adams et al., 2018 ; O'Dwyer et al., 2011). Ces deux échelles démontrent des corrélations positives élevées pour le score total, le score dynamique et la sous-échelle Clinical du HCR-20, ainsi que des corrélations positives modérées à élevées avec les sous-échelles Historical et Risk Management (Adams et al., 2018 ; O'Dwyer et al., 2011). Selon Smothers et al. (2023), le DUNDRUM-4 présente également des corrélations allant de modérées à élevées avec les sous-échelles Clinical et Risk Management du HCR-20.

Les scores totaux du GAF et du SAPROF sont corrélés fortement et négativement avec le DUNDRUM-3 et DUNDRUM-4 (Abidin et al., 2013 ; O'Dwyer et al., 2011). Ces deux échelles démontrent aussi de fortes corrélations positives et négatives avec les sous-échelles Vulnerabilities et Strengths du START respectivement (Abidin et al., 2013). Les DUNDRUM-3 et DUNDRUM-4 corrèlent positivement avec l'ensemble des mesures du PANSS, mais notons que les relations semblent généralement plus fortes avec le DUNDRUM-3 (O'Dwyer et al., 2011 ; O'Reilly et al., 2016). On constate également que les deux échelles démontrent des corrélations positives modérées à élevées avec l'ensemble des mesures du S-RAMM à l'exception de la sous-échelle Background (O'Dwyer et al., 2011). Le DUNDRUM-3 semble corréler fortement avec le SOFAS et modérément avec les échelles ACB et MCCB (O'Reilly et al., 2016). Notons également que les DUNDRUM-3 et DUNDRUM-4 corrèlent positivement et fortement avec leur version auto-évaluée respective (Davoren, Hennessy, et al., 2015). Enfin, l'étude de O'Reilly et al. (2016) rapporte une corrélation élevée entre les scores du DUNDRUM-3 mesurés à trois ans d'intervalle.

Dans l'ensemble, les résultats suggèrent que les DUNDRUM-3 et DUNDRUM-4 présentent une validité de construit robuste, corrélant avec une large gamme d'instruments mesurant les besoins, le risque, le fonctionnement et la symptomatologie. Ces corrélations soutiennent leur utilité clinique pour évaluer la progression dans les programmes et le rétablissement.

DISCUSSION

Ce rapport visait à analyser les qualités psychométriques du Quatuor DUNDRUM afin d'appuyer les démarches d'implantation et son utilisation dans la hiérarchisation des services en psychiatrie légale au niveau provincial. De façon générale, les résultats indiquent que les échelles du Quatuor DUNDRUM présentent des qualités psychométriques globalement satisfaisantes. Le DUNDRUM-1 et le DUNDRUM-2 montrent une fidélité interjuge adéquate et une validité de construit appuyée par des corrélations avec des instruments de référence portant sur les besoins, les risques et la sécurité. Leurs validités critériées suggèrent une capacité à différencier les niveaux de sécurité, malgré certaines limites liées au contexte d'application. Les DUNDRUM-3 et DUNDRUM-4 présentent une excellente cohérence interne et une validité de construit robuste, appuyées par des corrélations avec des mesures de besoins, de risques, de fonctionnement et de symptomatologie. Ces constats soutiennent l'idée que le Quatuor DUNDRUM mesure de façon cohérente les construits qu'il vise à opérationnaliser (sécurité thérapeutique, urgence, progression et rétablissement) et offre un potentiel intéressant pour soutenir la hiérarchisation des services.

LIMITES

Un certain nombre de limites doivent toutefois être considérées. D'abord, les études recensées reposent sur des échantillons de taille variable et parfois restreinte, ce qui limite la généralisation des résultats. La fidélité interjuge, bien que globalement satisfaisante, demeure tributaire de la formation des cliniciens, ce qui souligne la nécessité d'un encadrement rigoureux pour assurer une utilisation fiable de l'instrument. Certaines propriétés psychométriques, comme la cohérence interne du DUNDRUM-2, restent par ailleurs à documenter davantage. Sur un plan conceptuel, il importe de rappeler que les instruments de jugement clinique structuré comme le DUNDRUM ne visent pas à fournir un score « numérique », mais plutôt à soutenir un jugement raisonné. Leur validité et leur fidélité doivent donc être comprises dans cette perspective. Concernant la validité critériée, il demeure difficile d'identifier des critères externes exempts de biais : le niveau de placement (D-1, D-3, D-4) et la rapidité de placement (D-2) sont les principaux indicateurs utilisés, mais ceux-ci sont influencés par divers facteurs contextuels, tels que les contraintes de ressources et d'espaces, les préoccupations sociopolitiques et la subjectivité inhérente aux évaluations réalisées sans instrument de jugement structuré. Enfin, la quasi-totalité des études recensées ont été menées en contexte anglophone et à l'extérieur du Québec. Bien qu'une version traduite ait été développée localement (Kennedy et al., 2016b), elle n'a pas encore fait l'objet d'une validation empirique. De plus, au-delà de la fidélité interjuge, de la cohérence interne, de la validité critériée et de la validité de construit, la validité de contenu n'a pas été explorée, alors qu'elle requiert habituellement un jugement d'experts. Une telle démarche de validation de la version francophone québécoise apparaît donc essentielle en vue d'une implantation à grande échelle.

CONCLUSION

Cette réponse rapide visait à recenser les propriétés psychométriques du DUNDRUM afin de documenter les données soutenant le choix de cet instrument. La compréhension de ces propriétés est essentielle, puisqu'elle permet aux utilisateurs d'évaluer avec davantage de précision la qualité des instruments qu'ils utilisent. Compte tenu de l'importance des décisions relatives aux transferts de patients présentant une problématique psycholégale, il est essentiel de s'assurer que l'instrument utilisé présente des propriétés psychométriques satisfaisantes. Cette considération est d'autant plus importante lorsqu'il s'agit de sélectionner un instrument destiné à une utilisation à l'échelle provinciale pour soutenir la hiérarchisation des services en psychiatrie légale.

Ainsi, ce rapport permet une meilleure compréhension des propriétés psychométriques du DUNDRUM, tout en suscitant une réflexion sur ses implications cliniques. Il soutient également la formation actuellement offerte en proposant un complément d'information destiné à éclairer le choix du DUNDRUM comme instrument d'aide à la décision concernant la trajectoire de soins en psychiatrie légale, ainsi que ses forces et ses limites.

RÉFÉRENCES

Abidin, Z., Davoren, M., Naughton, L., Gibbons, O., Nulty, A. et Kennedy, H. G. (2013). Susceptibility (risk and protective) factors for in-patient violence and self-harm: prospective study of structured professional judgement instruments START and SAPROF, DUNDRUM-3 and DUNDRUM-4 in forensic mental health services. *BMC Psychiatry*, 13(1), 197. https://doi.org/10.1186/1471-244X-13-197

Adams, J., Thomas, S. D. M., Mackinnon, T. et 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. https://doi.org/10.1186/s12888-017-1584-8

Chaffe, J., Pillai, K., Gupta, R., McKenna, B., & Garrett, N. (2026). To admit or not admit : the utility of structured decision-making tools in forensic mental health services in Aotearoa New Zealand. *Psychiatry, Psychology and Law*, 0(0), 1 15. https://doi.org/10.1080/13218719.2026.2673496

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. https://doi.org/10.1037/1040.3590.6.4.284

Crocker, A. G., Braithwaite, E. et Côté, G. (2012). L'évaluation et la gestion du risque de violence. Dans *Intervention de pointe en réadaptation psychiatrique*. (p. 275-299). Presses de l'Université de Québec.

Crocker, A. G., Côté, G. et Braithwaite, E. (2011). Perspectives in violence risk assessment and management in mental health services. Dans *Applied Research and Evaluation in Community Mental Health Services: A current update and overview* (p. 94-114). McGill-Queen's University Press.

Crocker, A. G., Kahlaoui, K., Luigi, M., et Leclair, M. (2023). Troubles mentaux et comportements violents. Dans Lecomte, T. (Ed), Manuel de réadaptation psychiatrique, 3e édition. Presses de l'Université du Québec.

Davoren, M., Abidin, Z., Naughton, L., Gibbons, O., Nulty, A., Wright, B. et 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. https://doi.org/10.1186/1471-244X-13-185

Davoren, M., Byrne, O., O'Connell, P., O'Neill, H., O'Reilly, K. et Kennedy, H. G. (2015). Factors affecting length of stay in forensic hospital setting: need for therapeutic security and course of admission. *BMC Psychiatry*, 15, 301. https://doi.org/10.1186/s12888-015-0686-4

Davoren, M., Hennessy, S., Conway, C., Marrinan, S., Gill, P. et 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(1), 61. https://doi.org/10.1186/s12888-015-0433-x

Davoren, M., O'Dwyer, S., Abidin, Z., Naughton, L., Gibbons, O., Doyle, E., McDonnell, K., Monks, S. et 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. https://doi.org/10.1186/1471-244X-12-80

De Page, L., Boulanger, M., Bréart, M., Depuydt, A., Palmers, P., Salsac, C., Vandenbussche, S., Zaarour, Z. et Titeca, P. (2019). HoNOS-Secure in a Belgian Community and Forensic Patients with Psychotic Disorders. *Journal of Forensic Psychology Research and Practice*, 19(5), 341-358. https://doi.org/10.1080/24732850.2019.1660938

De Pau, M., Nicaise, P., Bourmorck, D., Vanderplasschen, W. et Vander Laenen, F. (2021). Organizing Forensic Mental Health Care Delivery: Putting the Trinitarian Model of Therapeutic Security to the Test. *International Journal of Forensic Mental Health*, 20(3), 291-302. https://doi.org/10.1080/14999013.2021.1876795

Douglas, K. S., Hart, S. D., Webster, C. D. et Belfrage, H. (2013). HCR-20V3: *Assessing risk of violence — User guide*. Mental Health, Law, and Policy Institute, Simon Fraser University. http://hcr-20.com/about/

Eckert, M., Schel, S. H. H., Kennedy, H. G. et 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. https://doi.org/10.1080/14789949.2017.1332771

Fleiss, J. L. (1999). Reliability of Measurement. Dans *The Design and Analysis of Clinical Experiments* (p. 1-32). John Wiley & Sons, Ltd. https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118032923.ch1

Flynn, G., O'Neill, C., McInerney, C. et Kennedy, H. G. (2011). The DUNDRUM-1 structured professional judgment for triage to appropriate levels of therapeutic security: retrospective-cohort validation study. *BMC Psychiatry*, 11(1), 43. https://doi.org/10.1186/1471-244X-11-43

Flynn, G., O'Neill, C. et Kennedy, H. G. (2011). 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. https://doi.org/10.1186/1756-0500-4-230

Freestone, M., Bull, D., Brown, R., Boast, N., Blazey, F. et 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. https://doi.org/10.1186/s12888-015-0620-9

Gouvernement du Québec. (2022). Plan d'action interministériel en santé mentale 2022-2026. S'unir pour un mieux-être collectif. https://publications.msss.gouv.qc.ca/msss/document-003301/

Guay, J.-P., Da Silva Guerreiro, J. et Crocker, A. G. (2022). Les méthodes et enjeux relatifs à l'évaluation du risque de la violence hétérodirigée. *Santé mentale au Québec*, 47(1), 63–85. https://doi.org/10.7202/1094145ar

Gulati, G., Otuokpaikhian, K., Crowley, M., Pradeep, V., Meagher, D. et Dunne, C. P. (2019). Mental healthcare interfaces in a regional Irish prison. *International Journal of Prisoner Health*, 15(1), 14-23. https://doi.org/10.1108/IJPH-06-2017-0029

Habets, P., Jeandarme, I. et Kennedy, H. G. (2019). Applicability of the DUNDRUM-1 in a forensic Belgium setting. J*ournal of Forensic Practice*, 21(1), 85-94. https://doi.org/10.1108/JFP-11-2018-0043

Habets, P., Jeandarme, I. et 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*, 26(9), 868-886. https://doi.org/10.1080/1068316X.2020.1742338

Hart, S. D., Douglas, K. S. et Guy, L. S. (2016). The structured professional judgment approach to violence risk assessment : Origins, nature, and advances. Dans D. P. Boer (Éd.), The Wiley *Handbook on the Theories, Assessment and Treatment of Sexual Offending* (pp. 643-666). Wiley.

Hogan, T. P. (2017). *Introduction à la psychométrie* (2e éd.). Chenelière Éducation

Hudon, A., Chainey, F., Degré, C., Courtemanche. A.-R., Meunier, M.-A., Fortier, A., Bourque, C., Boulé, L., Latreille, V., Mongeon, S., Allard, J.-M., Borduas Pagé, S. (2026). Longitudinal validity of the DUNDRUM-1 triage security scale in a Canadian not criminally responsible cohort: A retrospective study of 85 Patients. *The Canadian Journal of Psychiatry*, 0(0). doi:10.1177/07067437261497471

Jeandarme, I., Habets, P. et Kennedy, H. (2019). Structured versus unstructured judgment: DUNDRUM-1 compared to court decisions. *International Journal of Law and Psychiatry*, 64, 205-210. https://doi.org/10.1016/j.ijlp.2019.04.006

Jeandarme, I., Habets, P., O'Reilly, K. et Kennedy, H. G. (2021). Is non-completion of treatment related to security need? *Criminal Behaviour and Mental Health : CBMH*, 31(5), 321-330. https://doi.org/10.1002/cbm.2213

Jones, R. M., Patel, K. et 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. https://doi.org/10.1186/s12888-019-2083-x

Kennedy, H. G. (2002). Therapeutic uses of security: mapping forensic mental health services by stratifying risk. *Advances in Psychiatric Treatment*, 8(6), 433-443. https://doi.org/10.1192/apt.8.6.433

Kennedy, H. G. (2022). Models of care in forensic psychiatry. BJPsych Advances, 28(1), 46-59. doi:10.1192/bja.2021.34

Kennedy, H. G., Mullaney, R., McKenna, P., Thompson, J., Timmons, D., Gill, P., O'Sullivan, O. P., Braham, P., Duffy, D., Kearns, A., Linehan, S., Mohan, D., Monks, S., McLoughlin, L., O'Connell, P., O'Neill, C., Wright, B., O'Reilly, K. et Davoren, M. (2020). A tool to evaluate proportionality and necessity in the use of restrictive practices in forensic mental health settings: the DRILL tool (Dundrum restriction, intrusion and liberty ladders). *BMC Psychiatry*, 20(1), 515. https://doi.org/10.1186/s12888-020-02912-6

Kennedy, H. G., O'Neill, C., Flynn, G., Gill, P. et Davoren, M. (2016a). The DUNDRUM Toolkit V1.0.30. http://www.tara.tcd.ie/handle/2262/76545

Kennedy, H. G., O'Neill, C., Flynn, G., Gill, P. et Davoren, M. (2016b). DUNDRUM v1.0 : Quatre instruments de jugement professionnel structuré pour l'évaluation des niveaux de sécurité et d'urgence, la progression dans les programmes et le rétablissement (traduit par J.-P. Gagnon et A. Crocker ; version 1.0). Institut national de psychiatrie légale Philippe-Pinel. (Ouvrage original publié en 2016). http://www.santecom.qc.ca/bibliothequevirtuelle/pinel/127316.pdf

Lam, A. A., Penney, S. R., et Simpson, A. I. F. (2023). Construct validity and concordance of clinician- and patient-rated DUNDRUM Programme Completion and Recovery Scales. *International Journal of Forensic Mental Health*, 22(3), 252–261. https://doi.org/10.1080/14999013.2022.2151671

Lawrence, D., Davies, T.-L., Bagshaw, R., Hewlett, P., Taylor, P. et Watt, A. (2018). External validity and anchoring heuristics: application of DUNDRUM-1 to secure service gatekeeping in South Wales. *BJPsych Bulletin*, 42(1), 10-18. https://doi.org/10.1192/bjb.2017.6

McCullough, S., Stanley, C., Smith, H., Scott, M., Karia, M., Ndubuisi, B., Ross, C. C., Bates, R. et 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(4), e74. https://doi.org/10.1192/bjo.2020.61

Ministère de la Santé et des Services sociaux du Québec. (2015). *Faire ensemble et autrement. Plan d'action en santé mentale 2015-2020*. Ministère de la Santé et des Services sociaux. https://publications.msss.gouv.qc.ca/msss/fichiers/2017/17-914-17W.pdf

O'Dwyer, S., Davoren, M., Abidin, Z., Doyle, E., McDonnell, K. et 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(1), 229. https://doi.org/10.1186/1756-0500-4-229

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. https://doi.org/10.1186/s13033-016-0097-z

O'Reilly, K., O'Connell, P., Donohoe, G., Coyle, C., O'Sullivan, D., Azvee, Z., Maddock, C., Sharma, K., Sadi, H., McMahon, M. et Kennedy, H. G. (2016). Anticholinergic burden in schizophrenia and ability to benefit from psychosocial treatment programmes: a 3-year prospective cohort study. *Psychological Medicine*, 46(15), 3199-3211. https://doi.org/10.1017/S0033291716002154

Perafan, V., et Crocker, A. G. (2014). The Dundrum Quartet: Definition, Validity, Usefulness and Need for Future Research. Court rapport préparé pour le groupe de travail sur « la hiérarchisation des services de psychiatrie légale à Montréal » — Agence de santé et des services sociaux de Montréal (p. 13). Montréal, Canada: Institut universitaire en santé mentale Douglas

Parente, L., Carabellese, F., Felthous, A., La Tegola, D., Davoren, M., Kennedy, H. G., et Carabellese, F. F. (2024). Italian Evaluation and Excellence in REMS (ITAL-EE-REMS): Appropriate placement of forensic patients in REMS forensic facilities. *International journal of mental health systems*, 18(1), 33. https://doi.org/10.1186/s13033-024-00647-5

Peterson, R. A., Albaum, G. et Beltramini, R. F. (1985). A Meta-Analysis of Effect Sizes in Consumer Behavior Experiments. *Journal of Consumer Research*, 12(1), 97-103. https://doi.org/10.1086/209039

Pouls, C., Habets, P. et Jeandarme, I. (2026). Taking the DUNDRUM toolkit to the real world: field validity after implementation in Belgium. *Journal of Forensic Practice*, 1-17. https://doi.org/10.1108/JFP-02-2026-0019

Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W. et Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, 27(5), 1147-1157. https://doi.org/10.1007/s11136-018-1798-3

Smothers, M., Hill, C., Lawrence, D., Bagshaw, R., et Watt, A. (2023). Predictors of recovery in a medium secure service: Influence of the Welsh Government's Mental Health (2010) Measure. *International journal of law and psychiatry*, 91, 101935. https://doi.org/10.1016/j.ijlp.2023.101935

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, 48(6), 1273-1296. https://doi.org/10.1007/s11165-016-9602-2

Tourigny, M., St-Germain, A., Lanoue, A. et Trépanier, J. (2014). *Psychiatrie légale : Proposition d'organisation des services hiérarchisés de psychiatrie légale à Montréal*. Agence de la santé et des services sociaux de Montréal. http://www.santecom.qc.ca/bibliothequevirtuelle/Montreal/9782895108177.pdf

Williams, H. K., Senanayke, M., Ross, C. C., Bates, R. et Davoren, M. (2020). Security needs among patients referred for high secure care in Broadmoor Hospital England. *BJPsych Open*, 6(4), e55. https://doi.org/10.1192/bjo.2020.35

Wright, P. et Webster, C. D. (2011). Implementing structured professional judgment risk assessment schemes: An example of institutional change. *The International Journal of Forensic Mental Health*, 10(1), 1-6. https://doi.org/10.1080/14999013.2011.552365

ANNEXES

Tableau supplémentaire 1. Données descriptives des articles inclus dans la recension

Population				Administration de l'instrument			Instruments (et version) utilisés			
Références	**N**	**Âge μ (É-T)**	**Sexe (%♂)**	**Établissement**	**Participants**	**Pays**	**D1**	**D2**	**D3**	**D4**
Abidin et al., (2013)	100	40,5 (12,8)	94,0	Central Mental Hospital Dundrum	Tous les patients au Central Mental Hospital entre mars et avril 2010	Irlande	✓		✓	6 items
Adams et al., (2018)	327	45,2 (–)	89,6	Districts de santé locaux de la Nouvelle-Galles du Sud	Tous les patients médico-légaux adultes (âgés de plus de 18 ans) avec un diagnostic primaire de maladie mentale majeure	Australie	9 items		✓	✓
Chaffe et al., (2026)	141	36,3 (–)	81,5	Auckland Regional Forensic Service (ARFPS)	Tous les patients présentés au comité d'admission entre le 1er avril 2016 et le 31 mars 2022 en vue d'une admission dans un parcours (EFCR)	Nouvelle-Zélande	✓			
Davoren et al., (2012)	86	40,6 (12,8)	100,0	Central Mental Hospital Dundrum	Patients évalués entre février et mars 2010	Irlande	✓		✓	6 items
Davoren et al., (2013)	56	43,7 (12,8)	100,0	Central Mental Hospital Dundrum	Tous les patients de l'hôpital ont été évalués entre février et mars 2011	Irlande	✓		✓	6 items
Davoren, Byrne, et al., (2015)	230	34,9 (10,5)	83,3	Central Mental Hospital Dundrum	Tous les patients admis au Central Mental Hospital Dundrum entre le 1er janvier 2010 et le 30 juin 2014	Irlande	✓	✓		
Davoren, Hennessy, et al., (2015)	64	41 (12,3)	90,6	Central Mental Hospital Dundrum	Tous les patients de l'hôpital psychiatrique central en février 2012	Irlande	✓		v. clinicien; v. patient	v. clinicien; v. patient
De Page et al., (2019)	189	41,5 (11,9)	—	Centre hospitalier Jean Titeca et Centre Neuro-Psychiatrique Saint-Martin	Les patients traités dans l'un ou l'autre des établissements participants	Belgique	✓	✓		
Eckert et al., (2017)	139	—	100,0	Deux établissements du centre psychiatrique médico-légal néerlandais de haute sécurité de la Pompe Fondation	Tous les patients hospitalisés séjournant en LFPC ou en RFPC	Pays-Bas	9 items		✓	✓
Flynn, O'Neill, McInerney, et al., (2011)	316	—	100,0	Central Mental Hospital Dundrum	Hommes nouvellement incarcérés à la prison de Cloverhill entre janvier 2008 et décembre 2009, et nouvelles incarcérations (réceptions) dépistés entre avril et juin 2009	Irlande	✓			
Flynn, O'Neill et Kennedy (2011)	65	—	84,6	Central Mental Hospital Dundrum	Toutes les personnes placées sur la liste d'attente de janvier à juin 2010	Irlande	✓	✓		
Freestone et al., (2015)	188	34,6 (10,7)	89,2	Le East London NHS Foundation Trust Forensic Service	Les patients référés aux services médico-légaux entre juin 2011 et juin 2013	Royaume-Uni : Angleterre	✓	✓		
Gulati et al., (2019)	13	40 (–)	84,6	Les établissements de soins en santé mentale en Irlande	Patients référés du système de justice vers un établissement de soins de santé mentale à travers le pays	Irlande	✓	✓		

Population				Administration de l'instrument			Instruments (et version) utilisés			
Références	**N**	**Âge µ (É-T)**	**Sexe (%♂)**	**Établissement**	**Participants**	**Pays**	**D1**	**D2**	**D3**	**D4**
Habets et al., (2019)	50	38,6 (10,2)	100,0	Unité de sécurité moyenne à l'institut OPZC Rekem	Une sélection aléatoire de patients masculins référés pour admission dans une unité de sécurité moyenne par le tribunal	Belgique	9 items			
Habets et al., (2020)	100	43,6 (11,6)	100,0	Différents niveaux de sécurité thérapeutique	Une sélection aléatoire de 100 patients psychiatriques en milieu correctionnel	Belgique	9 items	✓		
Hudon et al., (2026)	85	39.0 (11.9)	91.8	Centre de psychiatrie médico-légale universitaire tertiaire au Québec	Adultes jugés non criminellement responsables en raison d'un trouble mental qui ont fait l'objet d'une évaluation entre le 1er janvier 2019 et le 31 décembre 2024.	Canada	9 items			
Jeandarme et al., (2019)	150	41,5 (11,3)	100,0	Unité de sécurité moyenne à l'institut OPZC Rekem ; Différents niveaux de sécurité thérapeutique	Une sélection aléatoire de 50 patients masculins référés pour admission dans une unité de sécurité moyenne par le tribunal et une sélection aléatoire de 100 patients masculins référés à différents niveaux de sécurité thérapeutique par le tribunal	Belgique	9 items			
Jeandarme et al., (2021)	50	36,1 (10,7)	100,0	Département de santé mentale médico-légale d'un hôpital psychiatrique général (sécurité moyenne)	Un échantillon aléatoire d'une cohorte de 160 patients admis entre 2001 et 2010	Belgique	9 items			
Jeandarme et al., (2026)	120	34,1 (11,7)	95,0	Centres psychiatriques médico-légaux à sécurité élevée (Ghent et Antwerp).	Patients ayant fait l'objet d'une admission (prioritaire et régulière) entre le 17 novembre 2014 et le 31 janvier 2023	Belgique		✓		
Jones et al., (2019)	643	34,2 (12,2)	0,0	Toutes les prisons provinciales détenant des femmes en Ontario	Toutes les femmes en détention provisoire ou condamnées dans une prison provinciale de l'Ontario, le 18 juillet 2017	Canada	✓	✓		
Kennedy et al., (2020)	28	34,6 (10,2)	100,0	Central Mental Hospital Dundrum	Tous les patients de sexe masculin admis dans une unité de 12 lits avec un ratio personnel/patients élevé (deux pour un en équivalent temps plein)	Irlande	✓			
Lam et al., (2023)	118	35.29 (11,5)	86.8	Trois niveaux de sécurité dans un hôpital psychiatrique au Canada	Tous les patients hospitalisés dans un hôpital psychiatrique entre 2020 et 2022	Canada			v. clinicien ; v. patient	v. clinicien ; v. patient
Lawrence et al., (2018)	121	—	—	Service de santé mentale médico-légale du sud du Pays de Galles	Tous les patients référés pour une premières fois aux Wales Secure Services entre janvier 2010 et juin 2013	Royaume-Uni : Pays de Galles	✓			
McCullough et al., (2020)	142	39,5 (10,1)	100,0	Broadmoor High Secure Hospital	Tous les patients hospitalisés ayant terminé l'ensemble des évaluations	Royaume-Uni : Angleterre			✓	✓
O'Dwyer et al., (2011)	95	40,9 (–)	91,6	Central Mental Hospital Dundrum	Patients à l'hôpital lors de la collecte des données en février et mars 2010	Irlande	✓		✓	6 items
O'Neill et al., (2016)	349	—	100,0	Cloverhill (prison pour hommes)	Toutes les admissions dans les hôpitaux médico-légaux, les admissions en psychiatrie générale et les déviations en consultation externe de 2012 à 2014	Irlande	✓	✓		

Population				Administration de l'instrument			Instruments (et version) utilisés			
Références	**N**	**Âge μ (É-T)**	**Sexe (%♂)**	**Établissement**	**Participants**	**Pays**	**D1**	**D2**	**D3**	**D4**
O'Reilly et al., (2016)	70	39,5 (11,1)	94,3	Central Mental Hospital Dundrum	Patients hospitalisés ayant reçu un diagnostic de schizophrénie ou de trouble schizo-affectif en 2012	Irlande			✓	
Parente et al.,. (2024)	193	43.5 (11,5)	92.8	Différents niveaux de sécurité thérapeutique (sécurité faible et moyenne)	Tous les patients admis dans l'un de ces services entre 2020 et 2021.	Italie	9 items		✓	✓
Pouls et al., (2026)	87	—	92,0	Eight organizations provided assessments for forensic psychiatric patients across different security levels.	Forensic psychiatric patients who were deemed not guilty by reason of insanity (NGRI). Between June 2022 and December 2024	Belgique			✓	✓
Smothers et al., (2023)	198	—	80,8	Unité de sécurité moyenne dans le sud du Pays de Galles	Tous les patients admis dans une unité de sécurité moyenne entre 2007 et 2017	Royaume-Uni : Pays de Galles		✓		6 items
Williams et al., (2020)	204	34,2 (9,5)	100,0	Broadmoor High Secure Hospital	Tous les patients référés d'hôpitaux ou de prisons à sécurité moyenne vers l'hôpital de Broadmoor, pour une éventuelle admission, entre 2015 et 2017	Royaume-Uni : Angleterre	9 items			