

**Évaluation de l'utilité des outils  
d'évaluation du risque et des  
mesures de la personnalité pour  
la prédiction de la récidive avec  
violence chez les délinquants adultes**

**2007-04**

par  
Mary Ann Campbell, Ph.D.  
Sheila French, M.A.  
et  
Paul Gendreau, Ph.D.  
Centre for Criminal Justice Studies  
Université du Nouveau-Brunswick à Saint John

Le présent rapport a été rédigé dans le cadre d'un contrat conclu avec le ministère de la Sécurité publique et de la Protection civile. Les opinions exprimées n'engagent que les auteurs et ne sont pas nécessairement celles du Ministère.

N° de cat. : PS3-1/2007-4F-PDF  
N° ISBN : 978-0-662-09392-3

## **Note des auteurs**

Mary Ann Campbell, Ph.D, est chargée d'enseignement et directrice du Centre for Criminal Justice Studies, Département de psychologie, Université du Nouveau-Brunswick (campus de Saint John). Sheila French est étudiante au doctorat en psychologie sous la direction de M<sup>me</sup> Campbell. Paul Gendreau est chercheur invité à la Division of Criminal Justice de l'Université de Cincinnati et professeur émérite au Département de psychologie de l'Université du Nouveau-Brunswick (campus de Saint John). Les demandes de renseignements au sujet du manuscrit doivent être envoyées à Mary Ann Campbell au Centre for Criminal Justice Studies, Université du Nouveau-Brunswick, C. P. 5050, Saint John (Nouveau-Brunswick), Canada, E2L 4L5; adresse électronique : [mcampbel@unbsj.ca](mailto:mcampbel@unbsj.ca).

Les auteurs tiennent à rendre hommage à Laurie Green, qui a calculé le coefficient d'objectivité du présent projet. Nous remercions Karl Hanson pour ses suggestions au sujet du taux de base et Paula Smith pour son aide concernant l'établissement de la base de données. Enfin, nous remercions Delphine Gossner, Karl Hanson, Peter Raynor, Steve Van Dine, Glenn Walters et Steve Wormith, qui n'ont ménagé ni leur temps ni leurs efforts pour nous aider et qui nous offrent gracieusement des documents de référence et/ou des données non publiées.

Une version antérieure du présent rapport a paru en 2007, mais elle a été révisée depuis en raison d'erreurs statistiques. Cela n'a à peu près pas entraîné de changements dans l'interprétation, la conclusion ou les recommandations du rapport, mais pour plus de clarté et d'exactitude, les citations devraient renvoyer au document actuel.



## Table des matières

Sommaire .....	i
Introduction.....	1
Prédicteurs du risque et principes procéduraux importants pour l'évaluation du risque.....	1
Variations des outils d'évaluation du risque de violence .....	2
Résumé.....	7
Méthodes et procédure .....	9
Échantillon d'études.....	9
Codage des études .....	9
Calcul de la valeur de l'effet .....	10
Hétérogénéité de la valeur de l'effet .....	10
Estimation du nombre garantissant la certitude des résultats.....	11
Résultats.....	13
Description de la base de données.....	13
Mesures du risque : validité prédictive de la violence en établissement.....	14
Mesures du risque : validités prédictives de la récidive avec violence .....	15
Comparaison des valeurs de l'effet selon la génération de l'outil d'évaluation du risque.....	16
Comparaisons fondées sur le contenu de l'outil : statique par rapport à dynamique .....	17
Comparaisons fondées sur la méthode d'administration des mesures .....	18
Comparaisons fondées sur la pertinence des outils d'évaluation pour les services correctionnels .....	19
Discussion .....	21
Outils actuariels et structurés et échelle de psychopathie .....	21
Recommandations pour guider la sélection des outils d'évaluation de la violence future.....	25
Autres questions qui pourraient faire l'objet d'études à l'avenir .....	30
Conclusion .....	31
Bibliographie.....	33
Annexe A .....	47



## Sommaire

Malgré l'existence d'outils d'évaluation du risque de violence [p. ex. le guide d'évaluation du risque de violence (VRAG)], d'outils généraux d'évaluation du risque lié à la criminalité et à la violence permanentes [p. ex. inventaire du niveau de service – révisé (INS-R)] et de mesures de la personnalité associée à l'agression [échelle de psychopathie – révisée (PCL-R)], relativement peu de méta-analyses comparatives de ces méthodes d'évaluation ont été réalisées concernant leur validité prédictive et leur valeur limite pour l'établissement d'objectifs en matière d'évaluation du risque de violence. Ces objectifs comprennent la prédiction du risque, la définition d'objectifs de réduction du risque et la conception d'un moyen de surveiller les variations du niveau de risque. Par conséquent, la présente étude visait à procéder à une évaluation méta-analytique de l'utilité relative des outils d'évaluation du risque et d'autres mesures psychologiques comme moyen de contribuer à l'établissement de normes de pratique pour la réalisation d'évaluation du risque de violence.

La présente méta-analyse a permis de comparer divers outils (p. ex. auto-évaluation, protocoles d'évaluation du risque clinique actuariels/structurés) qui ont servi à l'estimation du risque de violence. Afin de figurer dans la méta-analyse, une étude devait avoir été effectuée après 1980, être de nature prospective et comprendre un échantillon de cas de délinquants adultes bénéficiant de services de psychiatrie légale. De plus, il était essentiel que l'étude présente une estimation statistique de la relation entre un outil d'évaluation particulier et un résultat violent (non sexuel) ayant trait à la violence en établissement ou à la récidive avec violence qui pourrait être transformé en valeur de l'effet. Les études sur la récidive devaient aussi comprendre une période de suivi après la mise en liberté d'au moins six mois dont il fallait tenir compte dans l'analyse alors qu'aucune période de suivi minimale n'était nécessaire pour les études portant sur la violence en établissement si elle était de nature prospective. On n'a ménagé aucun effort pour recueillir des données publiées et non publiées.

Un guide de codage a été conçu afin d'obtenir des renseignements sur les caractéristiques de l'étude et de l'échantillon ainsi que le genre et la présentation des méthodes d'évaluation du risque utilisées et de recueillir les données statistiques pour le calcul des estimations de la valeur de l'effet. L'interprétation des résultats était fondée sur la statistique  $Z^+$ , qui est une estimation de la valeur de l'effet ajustée en fonction de la taille de l'échantillon. La statistique  $Z^+$  a été calculée au moyen de  $r'$ , qui est la corrélation entre un outil d'évaluation et une variable des résultats après ajustement pour tenir compte du taux de violence de base de l'échantillon. Lorsque les intervalles de confiance de 95 % entre deux valeurs de l'effet moyennes se chevauchaient dans une proportion d'au plus un quart de la longueur moyenne des deux intervalles, les deux valeurs de l'effet moyennes ont été interprétées comme représentant deux paramètres de population différents et étaient statistiquement différentes les unes des autres [c.-à-d. que  $p \leq 0,05$ ; Cummings & Finch (2005)]. Le chevauchement des intervalles de confiance concernant la valeur de l'effet qui excédait ce critère indiquait que les valeurs résultaient du même paramètre de population de l'échantillon et qu'elles n'étaient donc pas statistiquement différentes les unes des autres.

La méta-analyse a permis de faire un certain nombre de constatations significatives. En particulier, il y avait très peu de variation dans l'ampleur de la validité prédictive de la récidive violente pour les outils d'évaluation du risque actuariels/structurés utilisés et ayant fait l'objet de recherche le plus souvent (c.-à-d. HCR-20, INS-R, échelle d'ISR et PCL-R). Le nombre le plus élevé de valeurs de l'effet a été obtenu pour l'échelle PCL-R, mais les cinq outils ont produit des intervalles de confiance relativement précis pour leurs estimations du risque de violence dans la collectivité. Les valeurs de l'effet variaient de 0,22 (pour l'échelle d'ISR et l'échelle HCR-R) à 0,32 (pour le guide VRAG), ce qui témoigne de la validité prédictive généralement modérée de ces outils pour le résultat de la récidive violente. Les intervalles de confiance pour le guide VRAG, l'INS-R et l'échelle PCL-R se chevauchaient tous, ce qui

indique que ces mesures étaient des prédicteurs égaux de la récidive violente. Les intervalles de confiance concernant la valeur de l'effet de l'échelle HCR-20 chevauchaient également l'INS-R, l'échelle PCL-R et l'échelle d'ISR tandis que le guide VRAG a donné de meilleurs résultats que l'échelle HCR-20 et l'échelle d'ISR.

On a obtenu beaucoup moins de valeurs de l'effet pour la prédiction de la violence en établissement que pour la récidive violente. Seule l'échelle HCR-20 excédait les critères minimums de 10 valeurs de l'effet nécessaires pour une interprétation significative. Une forte variabilité était également évidente dans les intervalles de confiance pour les valeurs de l'effet de chaque outil examiné (y compris l'échelle HCR-20), ce qui a limité encore plus l'interprétation de leur valeur prédictive. Compte tenu de ces limites, les données préliminaires donnent à penser que les estimatives prédictives les plus solides provenaient de l'échelle HCR-20 ( $Z^+ = 0,28$ ), de l'INS-R ( $Z^+ = 0,24$ ) et de l'échelle PCL-VD ( $Z^+ = 0,22$ ) et d'une variable sur les antécédents criminels ( $Z^+ = 0,26$ ). Même si seulement l'échelle HCR-20 avait assez de valeurs de l'effet pour permettre d'avoir une confiance relative dans l'interprétation de sa vraie validité prédictive, ces valeurs étaient surtout fondées sur des échantillons de délinquants bénéficiant de services de psychiatrie légale. Par conséquent, même l'utilisation de l'échelle HCR-20 comme prédicteur de la violence en établissement exige des recherches supplémentaires concernant son applicabilité aux milieux carcéraux généraux.

D'autres variables ont été examinées comme modérateurs possibles de la validité prédictive des outils utilisés pour évaluer le risque de violence. D'après ces analyses, en tant que groupe, les outils d'évaluation du risque de deuxième génération (conçus à partir de procédures statistiques, qui contenaient des éléments principalement statiques et étaient de nature non théorique) étaient les prédicteurs les plus solides de la violence en établissement. Par contre, les outils d'évaluation du risque de troisième génération (axés sur la théorie, qui contenaient des éléments dynamiques et qui consistaient à mesurer les variations du risque) ont fourni ensemble les valeurs de l'effet les plus fortes pour la prédiction de la récidive avec violence dans la collectivité. La variation de la validité prédictive des outils de deuxième et de troisième génération d'évaluation des résultats de la récidive avec violence et de la violence en établissement peut être attribuable à la longueur des périodes de suivi. Les études sur la violence en établissement ont généralement une durée plus brève que les études sur la récidive. De plus, il y avait un lien plus étroit entre les outils qui avaient un contenu concernant la théorie criminologique et/ou ceux qui étaient spécialement conçus comme outils d'évaluation du risque et la prédiction de la violence en établissement et la récidive avec violence qu'avec ceux qui n'avaient pas été conçus à l'origine pour prédire le risque et/ou ceux qui mesuraient les concepts non pertinents ou faiblement associés au comportement criminel (p. ex. l'estime de soi).

En conclusion, la plupart des outils d'évaluation du risque disponibles actuellement sont des prédicteurs modérés de la violence future. Les outils fondés sur des facteurs historiques peuvent fournir des estimations plus fiables du risque de violence pendant l'incarcération (c.-à-d. prédictions à court terme), mais l'inclusion de facteurs dynamiques du risque et des besoins est importante pour la prédiction de la violence après la mise en liberté (c.-à-d. prédictions à long terme). Comme la plupart de ces outils prédisent la violence à des degrés semblables de précision, le choix de l'outil d'évaluation du risque le plus approprié devrait être fondé sur : a) le but de l'évaluation du risque, b) la capacité de l'outil de définir adéquatement les facteurs criminogènes du risque de violence d'un délinquant; c) la valeur informative de l'outil pour les programmes de traitement afin de réduire le risque de violence; d) la capacité de l'outil de mesurer les variations du niveau de risque. Chacun de ces facteurs facilite les pratiques de gestion des cas efficaces dans le cas des délinquants incarcérés ou sous surveillance dans la collectivité.



## **Évaluation de l'utilité des outils d'évaluation du risque et des mesures de la personnalité pour la prédiction de la récidive avec violence chez les délinquants adultes**

L'évaluation du risque de violence future d'un délinquant joue un rôle prépondérant dans les décisions prises concernant la peine à imposer à cet individu, la gestion du cas, la mise en liberté et les préoccupations en matière de sécurité publique [Andrews et Bonta (2003); Hoge et Andrews (1996)]. Ces évaluations devraient aussi aider à guider le choix des objectifs et des stratégies d'intervention qui permettront de réduire le risque [Heilbrun (1997)]. La plus grande partie des connaissances actuelles concernant la prédiction du risque de violence ont été acquises par suite de la recherche de réponses à des préoccupations (surtout dans les années 50 et 60) concernant la validité des critères utilisés pour prendre des décisions en matière de risque [Andrews (1989); Heilbrun (1997); Litwack et Schlesinger (1999); Monahan et Steadman (1994); Rice (1997)]. Plus précisément, les évaluations du risque de première génération au milieu du XX<sup>e</sup> siècle [Bonta (2002)] étaient fondées sur des jugements cliniques subjectifs du risque, qui étaient généralement formulés au moyen de méthodes d'évaluation non structurées et non systématiques [Hoge et Andrews (1996)]. La plupart des études ultérieures ont montré que la précision des évaluations du risque non structurées est inférieure aux estimations du risque provenant de méthodes de prédiction objectives, structurées et fondées sur des données (actuarielles) [Bonta, Law et Hanson (1998); Grove, Zald, Lebow, Snitz et Nelson (2000)]. L'essentiel, c'est que la prédiction exacte de la dangerosité future s'est révélée difficile pour les spécialistes [Hanson (2005); Quinsey, Harris, Rice et Cormier (1998)], mais elle peut être facilitée par l'utilisation d'outils d'évaluation du risque structurés. Pour faciliter davantage la pratique de l'évaluation du risque de violence, l'étude actuelle présentera un examen méta-analytique des divers outils d'évaluation utilisés pour guider ce processus. Afin d'établir la base des résultats et de la discussion, la première section de l'introduction décrira brièvement les facteurs importants pour la prédiction exacte du risque ainsi que les principes qui devraient guider le processus d'évaluation et de réduction du risque. Nous examinerons ensuite les outils d'évaluation contemporains du risque de violence et leurs variations sur le plan du résultat prévu, du contenu et de l'administration.

### *Prédicteurs du risque et principes procéduraux importants pour l'évaluation du risque*

Une caractéristique nécessaire d'une évaluation du risque efficace est la définition des variables qui contribuent au comportement criminel d'un individu [Bonta (2002)]. Beaucoup d'études ont été consacrées à cette tâche et ont fait ressortir un certain nombre de facteurs historiques et psychosociaux importants qui sont pertinents pour la prédiction de la dangerosité et la criminalité persistante [voir Andrews et Bonta (2003); Borum (1996)]. Pour résumer brièvement deux méta-analyses exhaustives sur les prédicteurs de la récidive, Bonta et coll. (1998) et Gendreau, Little et Goggin (1996) ont défini les attitudes antisociales, les pairs antisociaux, la toxicomanie, le dysfonctionnement familial, les conflits interpersonnels et les modes de vie négatifs ou instables ainsi que certaines variables démographiques (sexe masculin, célibataire et jeunes) comme prédicteurs utiles de la récidive générale et violente. En ce qui concerne la récidive avec violence, il existe d'autres prédicteurs, notamment le diagnostic de trouble de la personnalité antisociale ou de psychopathie, les antécédents de comportement violent et des problèmes d'emploi [Bonta et coll. (1998)]. De plus, ces prédicteurs importants du risque sont courants chez les populations carcérales générales et les délinquants souffrant de troubles mentaux [Bonta et coll. (1998); Phillips et coll., (2005)]. Il vaut également la peine de mentionner que bon nombre des facteurs cliniques traditionnels utilisés pour l'évaluation du risque (p. ex. intelligence, troubles de l'humeur, psychose, estime de soi) ont produit les validités prédictives les plus faibles pour la récidive avec violence et générale [Bonta et coll. (1998); Gendreau et coll. (1996)].

L'évaluation exacte du risque constitue une étape essentielle de la réduction efficace de celui-ci. Les principes directeurs à la base de l'évaluation efficiente du risque ou de la réadaptation sont les principes du risque, du besoin et de la réceptivité décrits par Andrews et Bonta (2003). Selon leur modèle, le *principe du risque* est fondé sur l'hypothèse qu'il est possible de prédire le comportement criminel et que l'intensité de l'intervention visant à réduire ce risque doit correspondre au niveau de risque du délinquant. Deuxièmement, le *principe du besoin* reconnaît que certains facteurs de risque peuvent être changés de manière à réduire le risque. Ces « facteurs criminogènes » ont trait au mode de vie, aux cognitions et au comportement du délinquant (p. ex. attitudes antisociales, toxicomanie) et sont liés empiriquement au risque de violence et/ou à la criminalité générale. Les interventions visant à réduire le risque doivent permettre de faire face aux facteurs criminogènes plutôt qu'aux facteurs qui ont de faibles liens avec la récidive (p. ex. estime de soi, dépression). Enfin, le *principe de la réceptivité* concerne le mode et la méthode d'intervention utilisés pour cibler les facteurs criminogènes. Le choix du traitement doit essentiellement être fondé sur des programmes qui s'appuient sur des données empiriques pour la réduction du comportement criminel comme les approches cognitivo-comportementales et d'apprentissage social [c.-à-d. réceptivité générale; Andrews et Bonta (2003)]. L'intervention doit également tenir compte du mode d'apprentissage du délinquant et d'autres facteurs qui peuvent nuire à sa capacité de répondre à l'intervention, comme un trouble mental, la motivation de changer ou les déficiences physiques [c.-à-d. réceptivité particulière; Andrews et Bonta (2003)]. On a montré que l'application des principes du risque, du besoin et de la réceptivité contribue plus à la réduction du risque que les interventions qui ne tiennent pas compte de ces principes ou qui en tiennent peu compte [Andrews et Bonta (2003); Dowden et Andrews (2000); French et Gendreau (2006); Gendreau, Goggin, French et Smith (2006)]. Les études poussées sur la prédiction du risque et les principes du risque, du besoin et de la réceptivité ont fourni des balises utiles permettant de concevoir des outils valides pour l'évaluation du risque de violence.

#### *Variations des outils d'évaluation du risque de violence*

Dans le but d'améliorer la qualité, la validité et l'efficacité des décisions prises en matière de risque de violence, on a accordé beaucoup d'attention à la conception d'outils d'évaluation du risque uniformisés, et un certain nombre d'outils d'évaluation du risque prometteurs fondés sur des données empiriques ont été élaborés [Borum (1996); Webster, Douglas, Eaves et Hart (1997a)]. Certains de ces outils sont conçus spécialement pour prédire la dangerosité comme le mécanisme de prédiction de la violence [VPS; Webster, Harris, Rice, Cormier et Quinsey (1994)], l'échelle d'évaluation du risque de violence [EERV; Wong et Gordon (2006)] et l'échelle des variables historiques, cliniques et de gestion du risque [HCR-20; Webster, Douglas, Eaves et Hart (1997b)]. Dans quelques cas, les outils visent à prédire une forme particulière de violence comme la violence à l'endroit d'un partenaire intime [guide d'évaluation du risque de violence conjugale – ERVC; Kropp, Hart, Webster et Eaves (1995)] ou la récidive sexuelle [risque de violence sexuelle-20 – SVR-20; Boer, Hart, Kropp et Webster (1997)]. Par conséquent, les spécialistes ont le choix des outils d'évaluation pour la prédiction de la dangerosité générale ainsi que certains types de comportement violent.

Même s'ils ne sont pas considérés comme une mesure du risque, les traits de personnalité psychopathique évalués par l'échelle de psychopathie – révisée [PCL-R; Hare (1991); 2003] se sont révélés utiles pour la prédiction de la violence future [Gendreau, Goggin et Smith (2002); Hemphill, Hare et Wong (1998)]. De plus, on a montré que les mesures du risque de récidive générale [p. ex. inventaire du niveau de supervision – révisé (INS-R); Andrews et Bonta (1995)] prédisent assez bien la violence future [Gendreau et coll. (2002); Harris, Rice et Quinsey (1993)]. L'utilité des outils d'évaluation du risque généraux pour

la prédiction de la violence est probablement attribuable au chevauchement des prédicteurs de risque pour la récidive avec violence et générale [Bonta et coll. (1998)]. Par conséquent, en plus des outils qui s'appliquent à la violence, les mesures conçues à d'autres fins peuvent aider à prédire la violence.

Comme l'a indiqué Bonta (2002), les méthodes d'évaluation du risque comprennent les méthodes papier-crayon [p. ex. l'échelle de sentiments criminels – ESC, Andrews et Wormith (1984); le questionnaire d'auto-évaluation – QAE; Loza, Dhaliwal, Kroner et Loza-Fanous (2000)], les méthodes d'examen des dossiers [p. ex. guide d'évaluation du risque de violence – VRAG; Harris et coll. (1993)], et les approches fondées sur des entrevues qui sont combinées aux examens des dossiers (p. ex. INS-R et HCR-20). Certaines de ces approches mesurent un seul concept relatif au risque [p. ex. les attitudes antisociales mesurées par l'ESC modifiée, Simourd (1997)], tandis que d'autres puisent à de multiples domaines associés à la récidive (p. ex. l'INS-R évalue 10 domaines de risque et de besoin). Même si l'administration et le contenu des outils d'évaluation du risque sont diversifiés, la pratique consistant à combiner les outils d'évaluation du risque pour produire une estimation consensuelle du risque peut présenter des problèmes. Mills et Kroner (2006) ont utilisé l'échelle PCL-R, l'INS-R, le VRAG et l'information statistique générale sur la récidive [ISGR, Bonta, Harman, Hann et Cormier (1996); Nuffield (1982)] pour prédire la récidive avec violence et générale après la mise en liberté. Pour la plupart des délinquants, ils s'entendaient sur les scores de risque uniformisés produits pour chacun de ces outils. Malheureusement, l'exactitude prédictive était considérablement réduite dans les cas où il y avait un degré élevé de désaccord entre les scores de risque uniformisés de ces outils. Les défis liés aux évaluations du risque fondées sur l'utilisation de plusieurs outils d'évaluation du risque font ressortir la nécessité d'effectuer des études qui indiquent l'outil d'évaluation du risque le plus approprié pour une population carcérale, une évaluation et un cadre médico-légal donnés.

*Production d'outils d'évaluation du risque.* L'utilité des renseignements que contiennent les outils d'évaluation du risque pour l'atteinte des objectifs de réadaptation et la surveillance des variations du risque au fil du temps varie en fonction de leur nature. Comme il est mentionné plus haut, les outils d'évaluation du risque de première génération étaient fondés sur des évaluations cliniques non structurées, non systématiques et subjectives du risque et étaient sujets à des erreurs et à des biais [p. ex. Grove et coll. (2000)]. Compte tenu des limites des méthodes d'évaluation de première génération, les outils d'évaluation du risque de deuxième génération ont été conçus pour permettre d'obtenir des prédictions du risque efficaces et uniformisées. Toutefois, le choix des éléments des méthodes de deuxième génération était basé sur des données purement statistiques [c.-à-d. actuarielles; Bonta (2002)]. Seuls les éléments qui permettaient une prédiction maximale de la récidive étaient inclus (p. ex. sexe masculin, groupe ethnique), sans égard à leur valeur théorique ou en matière de réadaptation. Le VRAG [Harris et coll. (1993)], l'échelle des facteurs prépondérants [SFS; Hoffman (1983)] et l'ISGR [Bonta et coll. (1996); Nuffield (1982)]. Malgré le fait que certains outils d'évaluation du risque actuariels affichent une validité prédictive assez bonne [p. ex. de 0,30 à 0,35; Bonta et Yessine (2005); Glover, Nicholson, Hemmati, Berfeld et Quinsey (2002); Loza et Green (2003); Polvi (2001)], ils se composent surtout d'éléments statiques [Andrews et Bonta (2003)]. Les facteurs de risque statiques sont de nature historique et/ou sont inchangeables (p. ex. sexe, âge, antécédents criminels). Parmi les facteurs statiques, les antécédents criminels constituent l'un des prédicteurs les plus solides de la violence future et de la récidive générale [Bonta et coll. (1998); Gendreau et coll. (1996); Webster et coll. (1997a)]. Cependant, les prédicteurs fondés sur les infractions antérieures ont été critiqués parce qu'ils ne permettent pas de saisir la complexité des facteurs qui contribuent à la récidive et de mesurer la variation du niveau de risque au fil du temps, et tous deux sont essentiels à la gestion des cas pour la réduction du risque

[Andrews, Bonta et Hoge (1990); Hoge et Andrews (1996)]. Les critiques des outils de deuxième génération soutiennent que le véritable objectif de l'évaluation du risque devrait être de contribuer à la réduction du risque plutôt que de prédire seulement le risque [Wong et Gordon (2006)].

En réponse à ces critiques, les outils d'évaluation du risque de troisième génération [Andrews, Bonta et Wormith (2006); Bonta (2002)] ont mis l'accent sur la valeur informative des modèles de prédiction pour les décideurs de la gestion des cas (p. ex. agents de libération conditionnelle et de probation, commissions des libérations conditionnelles, psychologues médico-légaux). Tout comme pour les mesures de deuxième génération, ces outils comprenaient des facteurs de risque fondés sur des données empiriques; toutefois, le choix des éléments était basé sur la connaissance théorique de la criminalité et de la violence persistantes (c.-à-d. théories de l'apprentissage social et de la cognition sociale, principes du risque, du besoin et de la réceptivité [Andrews et Bonta (2003) Gendreau et coll. (2006)]). De plus, les évaluations du risque de troisième génération comprenait des facteurs de risque dynamiques, qui de par leur nature varient et peuvent changer avec le temps ou sous l'influence de facteurs sociaux, psychologiques, biologiques ou contextuels comme le traitement, par exemple les facteurs malléables comme la toxicomanie, les conflits interpersonnels et les attitudes antisociales [voir Douglas et Skeem (2005)]. Même si les facteurs de risque statiques et dynamiques se sont révélés aussi utiles pour la prédiction du risque [Gendreau et coll.(1996); Wong et Gordon (2006)], selon la théorie actuelle, les facteurs dynamiques sont plus pertinents lorsque l'accent est mis sur la réduction du risque [Andrews (1989); Douglas et Skeem (2005); Heilbrun (1997)]. Dans le contexte de l'évaluation du risque, les facteurs dynamiques sont souvent appelés « facteurs criminogènes » en raison de leurs liens empiriques avec le comportement criminel. Par conséquent, l'utilisation d'outils d'évaluation des facteurs de risque dynamiques présente l'avantage suivant : ils sont sensibles aux variations du niveau de risque qui pourraient se produire au fil du temps et/ou par suite des efforts de réadaptation [Andrews et Bonta (2003); Heilbrun (1997)].

La génération la plus récente d'outils d'évaluation du risque [c.-à-d. la quatrième génération; Andrews et Bonta (2003); Andrews et coll. (2006)] a été conçue expressément pour être intégrée : a) au processus de gestion du risque; b) à la sélection de modes d'intervention et de cibles de traitement; c) à l'évaluation des progrès du traitement. Ces outils sont administrés à de multiples occasions (c.-à-d. lors des réévaluations) et sont informatifs, parce qu'ils documentent les variations de certains facteurs criminogènes et, en ce qui concerne le potentiel de risque général, qui pourraient se produire depuis le contact initial d'un délinquant avec le système de justice pénale jusqu'à sa sortie du système. Les outils de quatrième génération peuvent servir à définir les domaines de succès dans un plan de gestion des cas visant à réduire le risque et les domaines où des stratégies devraient être modifiées pour maximiser leur capacité de réduire le risque. Les formes dynamiques d'outils d'évaluation du risque ne sont pas très courantes en ce moment [Bonta (2002)], mais le niveau de service-inventaire de gestion des cas (NS-IGC) [Andrews, Bonta et Wormith (2004)] et l'échelle de risque de violence (ERV) [Wong et Gordon, (2006)] constituent deux exemples prometteurs.

De toute évidence, la pratique consistant à prédire la dangerosité et la criminalité futures s'est beaucoup améliorée depuis l'utilisation des évaluations du risque cliniques non structurées. Toutefois, il y a encore un manque d'uniformité quant à savoir si et dans quelle mesure il faut introduire une évaluation clinique dans la procédure d'évaluation du risque. Même si l'objectif des outils de deuxième génération était de ne pas recourir, sinon très peu, à l'évaluation clinique subjective, certaines mesures du risque encouragent une certaine souplesse clinique dans l'établissement d'estimations du risque (p. ex. NS-IGC, HCR-20, ERVC). Par exemple, plusieurs outils (p. ex. NS-IGC) permettent à l'évaluateur de recourir au concept de la « primauté des interventions cliniques », ce qui signifie que l'estimation actuarielle du risque peut être

ajustée d'après les jugements subjectifs de l'évaluateur concernant le rôle des facteurs de protection, les circonstances atténuantes ou d'autres facteurs particuliers à un cas. La souplesse clinique est également prise en compte dans les outils qui dissuadent d'utiliser les calculs purement actuariels des estimations du risque lorsqu'ils sont utilisés à des fins cliniques. Par exemple, il n'y a pas de scores-seuil cliniques ou de probabilités numériques du risque pour l'échelle HCR-20 [Webster et coll. (1997b)]. L'estimation du risque est plutôt fondée sur le jugement subjectif de l'évaluateur qui détermine si le délinquant se classe dans la catégorie de risque « faible », « modéré » ou « élevé ». Ce jugement est basé sur un examen systématique et attentif des facteurs de risque pertinents sur le plan théorique et empirique définis dans l'échelle HCR-20. Pour tenir compte de la structure prévue dans l'examen de chacun des facteurs de risque évalués, cette approche de l'évaluation du risque est désignée sous le nom de « jugement prédictif structuré ». Cependant, les outils fondés sur cette approche ont été critiqués parce qu'ils sont trop subjectifs et qu'ils souffrent des mêmes limites que les évaluations du risque cliniques non structurées [p. ex. Hilton, Harris et Rice (2006)].

*Méthode d'administration et pertinence du contenu des mesures d'évaluation du risque.* Pour le choix de l'outil d'évaluation du risque, il faut aussi se demander s'il faut utiliser une mesure évaluée par l'observateur et/ou une mesure d'auto-évaluation. Comme l'a indiqué Bonta (2002), les méthodes possibles d'évaluation du risque comprennent les méthodes papier-crayon [p. ex. l'échelle de sentiments criminels – ESC, Andrews et Wormith (1984); le questionnaire d'auto-évaluation – QAE, Loza, Dhaliwal, Kroner et Loza-Fanous (2000)], les méthodes d'examen des dossiers [p. ex. VRAG; Harris et coll. (1993)] et les approches fondées sur des entrevues qui sont combinées aux examens des dossiers (p. ex. INS-R et HCR-20). La majorité des outils de prédiction du risque sont fondés sur les évaluations par des spécialistes ayant reçu une formation des facteurs de risque statiques et/ou dynamiques individuels. Ces évaluations sont effectuées après un examen approfondi des renseignements auxiliaires et des renseignements provenant de dossiers correctionnels, qui peuvent aussi comprendre une entrevue semi-structurée avec le délinquant. Il s'agit d'une approche qui exige beaucoup de temps, mais qui permet de procéder à une évaluation exhaustive et valide du risque. Les mesures d'auto-évaluation [p. ex. MMPI-2 Pd scale, Hathaway et McKinley (1967); le système de classification MMPI de Megargee, Megargee et Bohn (1979) et les caractéristiques antisociales et les sous-échelles d'agression de l'inventaire d'évaluation de la personnalité, Morey (1991)] ont également été utilisés comme source d'information pour l'évaluation du risque que pose un délinquant (p. ex. Douglas, Hart et Kropp (2001); Magargee et Carbonell (1995); Morey et Quigley (2002); Osberg et Poland (2001)]. Même si l'administration des mesures d'auto-évaluation peuvent exiger peu de temps et être peu coûteuses, elles prêtent le flanc à une critique importante : elles n'ont généralement pas été conçues pour servir à une évaluation du risque. Par conséquent, elles ne sont pas nécessairement représentatives des facteurs de risque et de besoins définis de façon empirique utiles pour la prédiction et la gestion du risque [Bonta (2002); Walters (2006)].

En réponse à cette critique, certains chercheurs ont conçu des outils d'auto-évaluation qui visent expressément à évaluer les facteurs pertinents aux résultats du risque criminel. L'échelle de sentiments criminels-modifiée [ESC-M; Simourd (1997); Simourd et Van de Ven (1999)], le Psychological Inventory of Criminal Thinking Styles [PICTS, Walters (1995 et 1996)], et le questionnaire d'auto-évaluation [QAE; Loza et coll. (2000)] sont des exemples d'outils d'auto-évaluation du risque. Afin de déterminer la valeur des outils d'auto-évaluation pour la prédiction du risque, Walters (2006) a effectué une méta-analyse qui a servi à comparer certains outils d'évaluation du risque structurés/actuariels [c.-à-d. HCR-20, INS-R, PCL-R, VRAG et le Lifestyle Criminality Screening Form – LCSF créé par Walters, White et

Denney (1991)] à un certain nombre de mesures d'auto-évaluation qui ont été utilisées pour effectuer des évaluations du risque dans les cas d'inconduite en milieu carcéral, de récidive générale et de violence. Certaines des mesures d'auto-évaluation s'appliquaient expressément à la prédiction du risque (p. ex. PICTS, QAE), tandis que d'autres correspondaient à des concepts cliniques généraux considérés comme applicables au risque ou au moins à la personnalité générale et au fonctionnement psychologique d'un individu (p. ex. NEO Personality Inventory-Revised, répertoire multidimensionnel de la colère, échelle de désespoir de Beck; MMPI-Pd scale). Les conclusions de Walters corroboraient la validité prédictive des mesures d'auto-évaluation du risque, mais *seulement si* ces outils étaient fondés sur des concepts liés empiriquement au risque (p. ex. les attitudes antisociales). Selon Walters, l'intégration de mesures d'auto-évaluation *applicables au contenu* avec les outils d'évaluation du risque actuariels/structurés pourrait ajouter à la validité de l'évaluation du risque.

Même s'il s'agit d'une première étape informative, il y a plusieurs autres domaines d'intérêt qui résultent de la méta-analyse de Walter (2006). Premièrement, seulement un certain nombre d'outils d'évaluation du risque structurés/actuariels ont été codés (HCR-20, INS-R, PCL-R, VRAG et LCFS). De plus, seules neuf valeurs de l'effet étaient disponibles pour comparer la catégorie globale des méthodes structurées/actuarielles aux mesures d'auto-évaluation sur le plan de leur capacité de prédire la récidive *avec violence*. Pour ces neuf valeurs de l'effet, la valeur de l'effet moyenne des mesures structurées/actuarielles ( $r = 0,24$ ) était plus grande que pour la catégorie générale des mesures d'auto-évaluation ( $r = 0,17$ ). Une base de données plus importante, comprenant plus de valeurs de l'effet et une fourchette plus grande de mesures, est nécessaire pour reproduire les conclusions de Walter.

De plus, Walters (2006) n'a pas indiqué les validités prédictives respectives de chacun des outils figurant dans la catégorie actuarielle/structurée. Les renseignements de ce genre seraient utiles pour les spécialistes qui choisissent parmi les outils dont ils disposent ceux qui devraient figurer dans leurs évaluations du risque de violence. Quelques méta-analyses ont été réalisées à ce sujet. Gendreau et coll. (1996) ont comparé l'INS-R, l'échelle SFS et le système de classement des cas du Wisconsin [Baird (1981), Baird, Heinz et Bemus (1979)]. Chacun de ces outils était un prédicteur modéré de la récidive générale, mais l'INS-R a produit la valeur de l'effet moyenne pondérée la plus élevée (0,33). En outre, Gendreau et coll. ont évalué la valeur du MMPI pour la prédiction de la récidive générale. Même si le MMPI n'était pas un prédicteur aussi solide que l'INS-R ou l'échelle PCL-R, il a tout de même produit une valeur de l'effet pondérée importante de 0,21. Malheureusement, Gendreau et coll. n'ont pas analysé ces outils par rapport aux résultats en matière de violence. Une méta-analyse antérieure de Gendreau, Goggin et Law (1997) sur la prédiction de l'inconduite en milieu carcéral comprenait une comparaison de l'INS-R, du MMPI, d'autres mesures du risque et des mesures autres que le MMPI de la personnalité antisociale comme prédicteurs de l'inconduite en établissement. À l'instar de Gendreau et coll. (1996), les auteurs n'ont pas présenté un rapport distinct sur les validités prédictives de ces mesures pour l'inconduite non violente et l'inconduite violente. Cette décision était fondée sur le manque de variation importante des valeurs de l'effet entre ces deux résultats pour les nombreux autres prédicteurs examinés dans leur analyse. D'après le résultat global, l'INS-R a produit les validités prédictives les plus élevées ( $r = 0,23$ ) et a surpassé les autres mesures. Enfin, une comparaison méta-analytique plus récente de Gendreau et coll. (2002) a permis d'évaluer la violence comme critère distinct et de constater que l'INS-R avait un léger avantage sur l'échelle PCL-R pour la prédiction de la violence. Par conséquent, les méta-analyses existantes qui comparent les outils d'évaluation du risque donnent à penser qu'il peut y avoir des ressemblances dans la validité prédictive des outils d'évaluation du risque en ce qui concerne la récidive avec violence et générale.

## *Résumé*

On a fait beaucoup de progrès dans l'évaluation du risque général et de la dangerosité. Cependant, il reste une incertitude concernant les outils les plus appropriés de prédiction de la violence, étant donné les variations du contenu des éléments, de l'échelle, du niveau de subjectivité de l'évaluateur permis et de l'utilité des outils d'auto-évaluation en tant que composante des protocoles d'évaluation du risque de violence. Même si plusieurs études primaires comparent l'utilité des outils d'évaluation du risque pour la prédiction de la violence [p. ex. Dahle (2006); Douglas, Yeomans et Boer (2005); Grann, Belfrage et Tengstöm (2000); Kroner et Mills (2001); Mills et Kroner (2006)], seules quelques méta-analyses [c.-à-d. Gendreau et coll. (1996); Gendreau et coll. (1997); Gendreau et coll. (2002); Walters (2006)] ont été effectuées pour faire la synthèse de cette littérature à l'intention des spécialistes, et aucune de celles-ci n'était assez exhaustive dans son estimation du risque de violence. Une synthèse de cette nature arrive à point nommé, car très peu de psychologues correctionnels disent utiliser des outils conçus expressément pour prédire le risque [ou qui sont au moins fondés sur des données empiriques pertinentes à l'estimation du risque, voir Boothby et Clements (2000)]. Par conséquent, le principal objectif de la méta-analyse actuelle consistait à déterminer les outils les plus efficaces en tant que prédicteurs valides de la violence future (non sexuelle) en milieu carcéral et dans la collectivité. Au moyen de ces renseignements, il est possible d'établir des lignes directrices concernant la sélection des outils d'évaluation du risque ayant la capacité de produire les estimations du risque les plus valides et de contribuer à la gestion des cas et à la planification de la réadaptation.





## Méthodes and procédures

### *Échantillon d'études*

Nous avons effectué une recherche documentaire concernant les études sur la prédiction pertinentes au moyen des bases de données EBSCO (Academic Search Elite, PsycARTICLES et PsycINFO). Les termes de recherche clés étaient : a) des termes ayant trait à l'évaluation (p. ex. actuariel, clinique, prédiction, INS-R, PCL-R); b) des termes ayant trait à la population carcérale (p. ex. délinquant adulte, détenu, libéré conditionnel); c) des termes visant à classer le résultat sur le plan de la violence (p. ex. récidive, inconduite). Des données non publiées ont également été obtenues à la suite d'une demande envoyée par courrier électronique à environ 33 chercheurs et 23 centres de recherche qui effectuent des recherches sur les risques que posent les délinquants. D'autres études ont été ajoutées au moyen de la méthode traditionnelle (c.-à-d. examen des bibliographies d'articles). La recherche s'est limitée aux études réalisées de 1980 à 2006.

Selon les critères d'inclusion, les études primaires devaient a) être de nature prédictive (c.-à-d. que l'évaluation précédait la mesure du résultat); b) porter sur un échantillon d'adultes (c.-à-d. un échantillon moyen de personnes de 18 ans ou plus au moment de l'évaluation) choisi dans une population générale ou une population de délinquants bénéficiant de services de psychiatrie légale; c) présenter assez de données pour permettre de calculer une valeur de l'effet (p. ex. coefficient  $r$  de Pearson, coefficient Phi ( $\Phi$ )) entre la mesure de prédiction et les résultats de l'inconduite ou de la récidive violente. Les études sur l'incarcération et la probation ont été prises en compte, peu importe la durée du suivi. Dans le cas des études sur la récidive après la libération, la période de suivi devait être d'au moins six mois pour être incluse. Pour chaque étude, les données de l'échantillon le plus important, la période de suivi la plus longue et le genre de critère le plus spécifique (c.-à-d. condamnation par rapport à arrestation) ont été enregistrés. Pour éviter la répétition de la méta-analyse récente de Hanson et Morton-Bourgon (2007) sur la validité prédictive des outils d'évaluation du risque dans le cas des délinquants sexuels, dans l'analyse actuelle, nous avons exclu les études pour lesquelles les données sur les résultats en matière de violence provenaient presque exclusivement d'un échantillon de délinquants sexuels. De même, les outils conçus spécialement pour l'évaluation de la récidive sexuelle n'ont pas été inclus dans l'analyse actuelle. Les études prises en compte figurent à l'annexe A.

### *Codage des études*

Les catégories de codage, assorties d'exemples de leurs sous-éléments, étaient les suivantes : a) caractéristiques de l'étude et de l'auteur (p. ex. genre de publication, affiliation de l'auteur, année de publication); b) variables de l'échantillon (p. ex. groupe ethnique, sexe, genre de délinquant); c) descripteurs de l'évaluation du risque (p. ex. mesure utilisée, méthode d'administration, genre de prédicteurs évalués); d) descripteurs de la valeur de l'effet (p. ex. genre de résultat, valeur de l'effet calculée). Il est possible d'obtenir des renseignements détaillés ou un exemplaire du manuel de codage en contactant la première auteure. Toutes les études ont été codées par S. French. Le coefficient d'objectivité a été établi au moyen d'un échantillon tiré au hasard de 15 études, codé anonymement par un deuxième codeur d'expérience. Selon la formule de Yeaton et Wortman (1993) :  $\sum (\text{nombre d'accords}) / \sum (\text{nombre d'accords} + \text{nombre de désaccords})$ , l'indice d'accord était de 0,82. La source des désaccords concernait les caractéristiques moins évidentes de l'échantillon (c.-à-d. détermination du niveau de risque de l'échantillon) et les aspects de la nature d'un outil d'évaluation du risque particulier (c.-à-d. genre de

contenu de l'élément, production d'un outil d'évaluation du risque). Les désaccords résultaient la plupart du temps d'un malentendu au moment de la lecture de l'étude ou d'une erreur de transcription au moment de l'entrée des codes. Les deux évaluateurs ont discuté des désaccords et ils se sont entendus sur le codage de ces éléments avant l'analyse.

### *Calcul de la valeur de l'effet*

Les coefficients Phi ( $\Phi$ ) ont été calculés pour la validité prédictive de chaque mesure au moyen des résultats concernant l'inconduite et la récidive. Lorsque des statistiques autres que  $r$  ont été présentées (c.-à-d.  $F$ ,  $t$ ,  $\chi^2$ ,  $p$ ,  $AUC$ ), la formule appropriée de conversion de  $\Phi$  a été utilisée [Rosenthal (1991); Swets (1986)]. Compte tenu des taux de base généralement faibles de l'inconduite et de la récidive avec violence, il a fallu prendre en considération cette influence possible sur les valeurs de l'effet. Les coefficients Phi ont été ajustés au moyen de la formule de Ley (1972) :  $r' = [(r_{xy})(\delta_x'/\delta_x)] / [1 - r_{xy}^2 + (r_{xy}^2)(\delta_x'^2/\delta_x^2)]^{1/2}$ , où  $r_{xy}$  était la corrélation observée,  $\delta_x$  était l'écart-type observé du taux de base,  $\delta_x'$  était l'écart-type moyen fondé sur le taux de base moyen pour les études de l'analyse et  $r_{xy}'$  était la corrélation corrigée. L'écart-type du taux de base a été calculé au moyen de la formule  $\delta = [pq/(N)(N-1)]^{1/2}$ , où  $p$  était le nombre de participants qui étaient des récidivistes en milieu carcéral ou dans la collectivité,  $q$  était le nombre de participants qui étaient des non-récidivistes en milieu carcéral ou dans la collectivité et  $N$  était la taille totale de l'échantillon.

Les paramètres utilisés afin d'estimer et d'interpréter l'ampleur de la relation entre chaque mesure du risque (p. ex. HCR-20) ou catégorie de prédicteur (p. ex. mesures de 2<sup>e</sup> génération) et les valeur de l'effet de l'inconduite ou de la récidive étaient la valeur moyenne  $r'$  ( $M_{r'}$ ) pondérée par la taille de l'échantillon [ $Z^+$ , voir Hedges et Olkin (1985)], ainsi que son intervalle de confiance associé de 95 % ( $CI_{Z^+}$ ). L'intervalle de confiance ( $CI$ ) a servi à indiquer le degré d'accord entre les variables de l'étude. S'il n'y avait pas de chevauchement du tout entre les intervalles de confiance pour deux valeurs de l'effet moyennes ou si les intervalles de confiance se rejoignaient, les valeurs étaient interprétées comme représentant des paramètres de population différents. Ce critère équivaut à la signification statistique de  $p < 0,006$ , pourvu que la taille de l'échantillon soit  $\geq 10$  et que la largeur des intervalles de confiance ne varie pas plus d'un facteur ou deux [Cumming et Finch (2005)]. Lorsque les intervalles de confiance entre deux comparaisons se chevauchaient dans une proportion d'au plus un quart de la longueur moyenne des deux intervalles, les valeurs moyennes ont également été interprétées comme représentant deux paramètres de population différents et étaient statistiquement différentes, soit environ  $p \leq 0,05$  [voir Cummings et Finch (2005)]. Si le chevauchement des intervalles de confiance dépassait le critère susmentionné, cela signifiait que les valeurs de l'effet moyennes résultaient probablement du paramètre de la population de l'échantillon et, par conséquent, qu'elles n'étaient pas statistiquement différentes. Une autre utilisation des intervalles de confiance a consisté à indiquer la précision des estimations des valeurs de l'effet, qu'on a établie en notant la largeur de l'intervalle de confiance [voir Cumming et Finch (2001); Gendreau, Goggin et Smith (2000); Schmidt (1996)]. Des intervalles plus étroits indiquent une estimation plus précise d'un paramètre de la population que des intervalles plus larges.

### *Hétérogénéité de la valeur de l'effet*

Nous avons déterminé l'influence des observations aberrantes au moyen de la statistique  $Q$  [Rosenthal (1991)]. Pour chaque valeur de l'effet, une valeur  $q$  a été calculée au moyen de la formule :  $(n - 3)(z_{r'} - Z^+)$ , où  $n$  était la taille de l'échantillon total pour chaque valeur de l'effet;  $z_{r'}$  correspondait à la valeur normalisée  $r'$  pour chaque valeur de l'effet et  $Z^+$  était la valeur pondérée de l'échantillon  $M_{r'}$ .

pour chaque catégorie de prédicteurs. Ces valeurs  $q$  ont été ensuite additionnées pour chaque catégorie de prédicteurs, ce qui a donné  $Q$ , qui est une estimation de l'hétérogénéité des valeurs de l'effet dans cette catégorie. Pour évaluer sa signification, on a évalué la statistique  $Q$  au moyen de la valeur critique de  $\chi^2$  avec  $(k - 1)$  degrés de liberté. Une statistique  $Q$  significative indique qu'il y a plus de variabilité que ce que le hasard pourrait laisser supposer. Dans ces cas, les valeurs de l'effet des observations aberrantes ont été examinées et n'ont été éliminées que s'il y avait une raison logique de les exclure : par exemple, une erreur de codage ou une caractéristique unique de l'étude (p. ex. un échantillon restrictif).

#### *Estimation du nombre garantissant la certitude des résultats*

Nous avons eu recours à la méthode du nombre garantissant la certitude des résultats pour déterminer combien de valeurs de l'effet supplémentaires il faudrait pour modifier une estimation de la valeur de l'effet obtenue. Nous avons calculé l'indice du nombre de valeurs de l'effet ( $Z^+ = 0,00$ ) nécessaire pour une mesure du risque donnée plus précise pour la prédiction de l'inconduite ou de la récidive afin d'obtenir une valeur de l'effet égale à une valeur d'exactitude moindre à l'aide de la formule  $[(k_B(Z_B^+ - Z_A^+))]/(Z_A^+ - Z_{B=0}^+)$ , où  $Z_{B=0}^+$  indique une valeur de l'effet nulle pour la mesure du risque la plus précise [voir Gendreau et coll. (2002)]. Aux fins de la présente méta-analyse, nous partons de la prémisse selon laquelle la valeur de l'effet moyenne de la mesure A était de 0,30 ( $k = 50$ ), et celle de la mesure B, de 0,35 ( $k = 40$ ). Selon cette formule, on estime à sept le nombre de prédictions B avec un coefficient  $Z^+ = 0,00$  qui devraient être nécessaires afin d'infirmer la supériorité de la mesure B par rapport à la mesure A. En d'autres termes, il faudrait trouver sept valeurs de l'effet supplémentaires de la mesure B, dont chacune aurait une importance de  $Z^+ = 0,00$  pour pouvoir conclure que les deux mesures ont une efficacité prédictive égale.



## Résultats

### *Description de la base de données*

L'ensemble de données final contenait 88 études sur la validité prédictive de diverses mesures du risque d'inconduite violente ( $k = 76$ ) et de récidive avec violence ( $k = 185$ ). Même si les études portaient sur plus de 70 mesures du risque différentes au total, seules les mesures représentées dans 10 estimations de la valeur de l'effet ou plus par résultat seront indiquées afin de mettre l'accent sur les outils pour lesquels les données les plus nombreuses étaient disponibles. Ces outils comprenaient l'échelle HCR-20 ( $k = 11$  pour l'inconduite;  $k = 11$  pour la récidive), l'INS/INS-R ( $k = 19$  pour la récidive), l'échelle PCL/PCL-R ( $k = 24$  pour la récidive), l'échelle d'ISR ( $k = 17$  pour la récidive) et le VRAG ( $k = 14$  pour la récidive). Certains outils comptant moins de 10 valeurs de l'effet sont recensés s'il y a lieu, mais leur validité prédictive doit être interprétée avec prudence étant donné le faible nombre de valeurs de l'effet.

Les renseignements utilisés pour compléter les outils d'évaluation ont été recueillis en grande partie à partir de dossiers (52,2 % des valeurs de l'effet), d'entrevues (11,2 %) ou d'une combinaison des deux méthodes (16,5 %). Des méthodes d'auto-évaluation ont été utilisées afin de recueillir des données pour 17,4 % des valeurs de l'effet et ont été combinées à une entrevue dans moins de 1 % des valeurs de l'effet. Enfin, 1,8 % des valeurs de l'effet provenaient de mesures établies à partir d'observations du comportement par le personnel des établissements. L'évaluation du risque était entièrement ou en grande partie fondée sur des facteurs de risque statiques pour 34,9 % des valeurs de l'effet et sur des facteurs dynamiques pour 51,9 % des valeurs de l'effet. Un peu plus de 8 % des valeurs de l'effet provenaient de mesures faisant appel à un nombre égal de prédicteurs statiques et de prédicteurs dynamiques. Enfin, 4,9 % provenaient d'études où les renseignements n'étaient pas suffisants pour déterminer la nature des prédicteurs utilisés. La grande majorité des mesures (85 %) s'appliquaient aux services correctionnels (c.-à-d. fondées sur une théorie du comportement criminel et/ou créées expressément pour servir d'outil d'évaluation du risque criminel). Un peu moins de 3 % des valeurs de l'effet étaient fondées sur une évaluation clinique (c.-à-d. évaluation de première génération), 52,3 % provenaient de mesures classées comme étant des mesures actuarielles de deuxième génération. Les mesures de troisième et quatrième génération étaient représentées dans 42,3 % et 2,5 % des valeurs de l'effet, respectivement.

En ce qui concerne les caractéristiques de l'étude, 63,1 % des valeurs de l'effet provenaient d'études publiées dans des livres, des revues ou des rapports gouvernementaux, 32 % de thèses de doctorat et 5 % de données non publiées obtenues directement de chercheurs (c.-à-d. données brutes, manuscrits non publiés ou exposés à des conférences ou communications affichées). Les études canadiennes et américaines ont produit 60,1 % et 24,8 % des valeurs de l'effet, respectivement. La plupart des valeurs de l'effet ont été présentées dans des études produites par des auteurs affiliés à des établissements d'enseignement (51,3 %) de la psychologie (85,4 %). Les valeurs de l'effet représentaient un total combiné de 273 734 délinquants (inconduite  $N = 232\ 790$ ; récidive  $N = 40\ 944$ ). La majorité (81,3 % des valeurs de l'effet s'appliquaient aux deux résultats) provenaient d'échantillons composés en grande partie d'hommes.

Les échantillons représentant les populations de délinquants générales ont produit 63,9 % des valeurs de l'effet tandis que les autres étaient fondés sur des échantillons de délinquants bénéficiant de services de psychiatrie légale (30,7 %) et mixtes (5 %). Les valeurs de l'effet de la violence en milieu carcéral étaient fondées sur un pourcentage équivalent d'échantillons de la population carcérale générale (50,7 %) et

d'échantillons de délinquants bénéficiant de services de psychiatrie légale (49,3 %). Par contre, la majorité des valeurs de l'effet de la récidive avec violence étaient fondées sur des échantillons de la population carcérale générale (70,0 %). Un niveau de risque d'un échantillon était défini par les auteurs de l'étude originale (6,3 % des valeurs de l'effet) ou par le principal codeur de l'étude en cours (90,6 % des valeurs de l'effet). Dans l'ensemble, la plupart des valeurs de l'effet étaient fondées sur des échantillons à risque faible ou modéré (43,6 % et 44,0 % respectivement). Seulement 7,5 % des valeurs de l'effet provenaient d'échantillons à risque élevé et un peu plus de 3 % étaient fondés sur des échantillons à risque mixte. Pour 2,1 % des valeurs de l'effet, il n'y avait pas assez de renseignements pour permettre de déterminer le niveau de risque de l'échantillon. Il était impossible d'évaluer la prédisposition à la violence chez les délinquants avec un certain degré de certitude parce que les renseignements sur les crimes de violence précédents et actuels n'étaient pas déclarés pour 67,6 % et 56,0 % des valeurs de l'effet, respectivement.

En moyenne, le taux de base pour les inconduites violentes en établissement s'établissait à 25,84 % (É.-T. = 13,61) et le taux de base moyen de la récidive avec violence était de 21,73 % (É.-T. = 12,99). Seulement 39,4 % des valeurs de l'effet de la violence en milieu carcéral étaient fondées sur des périodes de suivi de plus d'un an, et la plupart des délinquants mis en liberté étaient suivis pendant des périodes allant de 2 à 5 ans (41,7 %). L'indice le plus courant de l'inconduite en milieu carcéral était les registres des pénitenciers (74,7 %) tandis que les données sur les arrestations par suite de récidive, condamnations ou incarcérations étaient les indices les plus courants de la récidive avec violence (72,2 % des valeurs de l'effet). Dans la plupart des études (97,0 %), les récidivistes violents étaient comparés à un groupe global de délinquants (c.-à-d. les délinquants qui n'avaient pas récidivé du tout combinés à ceux qui pouvaient avoir récidivé sans violence). Par conséquent, il y avait peu de données prédictives faisant appel à un critère de résultat fondé uniquement sur l'« absence de récidive ».

#### *Mesures du risque : validité prédictive de la violence en établissement*

Le tableau 1 contient les valeurs  $Z^+$  et les intervalles de confiance connexes de 95 % des mesures du risque et des résultats de la violence en établissement. Une seule mesure était représentée par plus de 10 valeurs de l'effet (c.-à-d. HCR-20); toutefois, pour assurer l'uniformité avec les outils signalés pour les résultats de la récidive violente, des données préliminaires pour les outils correspondants sont présentées ici en dépit d'un  $k$  inférieur à 10. L'échelle HCR-20 et l'INS/INS-R affichaient les valeurs de l'effet pondérées moyennes les plus élevées pour la prédiction de la violence en établissement ( $Z^+ = 0,28$  et  $0,24$ , respectivement). L'échelle PCL-VD ( $k = 7$ ) a produit la troisième valeur de l'effet moyenne en importance ( $Z^+ = 0,22$ ) tandis que les échelles PCL/PCL-R et le guide VRAG ont produit les associations les plus faibles pour la violence en établissement ( $Z^+ = 0,14$  et  $Z^+ = 0,15$ , respectivement). Toutefois, les intervalles de confiance de 95 % pour chacune des mesures du risque ci-dessus se chevauchaient considérablement, ce qui donne à penser qu'ils s'appliquent tous au même paramètre de la population. En outre, la largeur de leurs intervalles de confiance et le petit nombre de valeurs de l'effet disponibles indiquent le manque de précision de l'estimation de la valeur de l'effet de chaque outil. Par conséquent, les interprétations fondées sur ces estimations devraient être considérées comme provisoires jusqu'à ce que d'autres études primaires soient effectuées. Étant donné qu'un critère minimum de 10 valeurs de l'effet par outil a été établi pour le calcul des analyses du nombre garantissant la certitude des résultats (pour permettre des comparaisons significatives entre les validités prédictives de chaque outil), ces mesures n'ont pas été calculées pour la violence en établissement.

**Tableau 1**

Comparaisons de la valeur de l'effet des mesures du risque pour la prédiction de l'inconduite violente en établissement<sup>a</sup> et la récidive violente<sup>b</sup>

Mesure	<i>k</i>	<i>N</i>	$Z^+$	$CI_{z^+}$	<i>Q</i>
<b>Violence en établissement</b>					
HCR-20	11	758	0,28	0,10 à 0,24	12,26
INS/INS-R	6	650	0,24	0,09 à 0,25	5,91
PCL/PCL-R	5	626	0,14	0,00 à 0,16	2,90
PCL-VD	7	504	0,22	0,07 à 0,25	5,59
VRAG	2	222	0,15	-0,08 à 0,18	1,54
Indices des antécédents criminels	4	204 132	0,26	0,24 à 0,25	12,83*
<b>Récidive violente</b>					
HCR-20	11	1 395	0,22	0,17 à 0,27	6,68
INS/INS-R	19	4 361	0,28	0,25 à 0,31	57,15*
PCL/PCL-R	24	4 757	0,27	0,24 à 0,30	48,04*
Échelle d'ISR	17	5 618	0,22	0,19 à 0,25	32,54*
VRAG	14	2 082	0,32	0,28 à 0,36	47,06*

\*  $p < 0,05$ , indique qu'il y a plus de variabilité que ce que le hasard pourrait laisser supposer.

N.B. *k* = valeurs de l'effet par mesure du risque; *N* = délinquants par mesure du risque;  $Z^+$  = valeur *r*' pondérée selon la taille de l'échantillon;  $CI_{z^+}$  = intervalle de confiance de 95% pour  $Z^+$ .

<sup>a</sup> Même si le nombre total d'estimations des valeurs de l'effet pour les mesures du risque de violence en établissement était de 76, il n'y avait qu'une catégorie dont *k* était supérieur à 10. Les autres mesures indiquées ci-dessus sont incluses pour faciliter les comparaisons provisoires de la validité prédictive de ces mesures avec les résultats de l'inconduite et de la récidive.

<sup>b</sup> Même si le nombre total d'estimations des valeurs de l'effet pour les mesures du risque de récidive s'élevait à 185, seules les mesures comptant plus de 10 validités prédictives figurent au tableau 1.

#### Mesures du risque : validités prédictives de la récidive avec violence

Les valeurs  $Z^+$  avec des intervalles de confiance associés de 95 % pour la prédiction de la récidive violente figurent dans la deuxième partie du tableau 1. La valeur  $Z^+$  la plus élevée a été enregistrée pour le guide VRAG. Il y avait un chevauchement des intervalles de confiance entre cette mesure et l'INS-R et l'échelle PCL/PCL-R, mais ses intervalles de confiance ne chevauchaient pas avec l'échelle HCR-20 ou l'échelle d'ISR. D'après ces largeurs, les intervalles de confiance pour chacune de ces mesures, l'INS-R, l'échelle PCL-R et l'échelle d'ISR ont produit des estimations ponctuelles légèrement plus précises que l'échelle HCR-20 et le guide VRAG. Des analyses garantissant la certitude des résultats ont indiqué que six valeurs de l'effet supplémentaires nulles pour le guide VRAG seraient nécessaires pour réduire sa capacité prédictive au niveau de celle de l'échelle HCR-20 et de l'échelle d'ISR. Seulement deux autres valeurs de l'effet nulles pour le guide VRAG seraient nécessaires pour que celui-ci donne des résultats semblables à ceux de l'INS-R ou de l'échelle PCL-R.

En ce qui concerne les mesures dignes de mention ayant moins de 10 valeurs de l'effet pour la récidive violente (ne figurant pas au tableau 1), le NS-IGC ( $k = 3$ ,  $N = 841$ ) a produit un niveau élevé de validité prédictive ( $Z^+ = 0,47$ ,  $CI_{Z^+} = 0,40$  à  $0,54$ ), suivi de près par l'échelle SFS ( $k = 8$ ,  $N = 1\ 094$ ,  $Z^+ = 0,37$ ,  $CI_{Z^+} = 0,31$  à  $0,43$ ). Les intervalles de confiance pour ces deux mesures ne se chevauchaient que légèrement et donnaient à penser qu'ils pouvaient, en fait, servir à estimer des paramètres de population distincts. Il convient toutefois de noter que les conclusions tirées au sujet de ces deux mesures doivent l'être à la lumière du fait que peu de valeurs de l'effet étaient disponibles pour tester cette validité prédictive pour la récidive violente. Les autres mesures dignes de mention étaient l'échelle PCL-VD ( $k = 5$ ,  $N = 641$ ,  $Z^+ = 0,20$ ,  $CI_{Z^+} = 0,12$  à  $0,28$ ) et les mesures constituées seulement des éléments des antécédents criminels ( $k = 9$ ,  $N = 2\ 230$ ,  $Z^+ = 0,23$ ,  $CI_{Z^+} = 0,19$  à  $0,27$ ). D'après les trois valeurs de l'effet de l'inventaire MMPI (selon la typologie de Megargee et l'échelle d'adaptation au pénitencier), cet outil n'a pas prédit de manière significative la récidive violente ( $Z^+ = 0,00$ ).

#### *Comparaison des valeurs de l'effet selon la génération de l'outil d'évaluation du risque*

Le tableau 2 présente des données servant à comparer des estimations des valeurs de l'effet moyennes pour différentes générations de mesures du risque. Les mesures de première et de quatrième génération comptaient moins de 10 valeurs de l'effet et ont donc été exclues du tableau, mais leurs données préliminaires sont décrites ci-dessous. Comme le montre le tableau 2, les outils de deuxième génération ont obtenu un meilleur rendement que les outils de troisième génération comme prédicteurs de la violence en établissement. Ce phénomène était attribuable au poids considérable accordé à trois études particulièrement importantes sur la deuxième génération où *ns* était supérieur à 10 000 délinquants. Selon des calculs garantissant la certitude des résultats, on a estimé qu'il faudrait ajouter 34 autres valeurs de l'effet de deuxième génération de zéro avant que sa valeur moyenne soit égale à celle des mesures de troisième génération dans la prédiction de la violence en établissement. L'avantage des outils de deuxième génération par rapport à ceux de troisième génération a été renversé lorsque le résultat était la récidive violente, c'est-à-dire que les mesures de troisième génération avaient un léger avantage par rapport à celles de deuxième génération et qu'il n'y avait pas de chevauchement de leurs intervalles de confiance. Selon l'indice garantissant la certitude des résultats, il faudrait 23 autres valeurs de l'effet nulles pour les études de troisième génération afin de réduire son effet moyen au niveau de celui des outils de deuxième génération.

Pour la première et la quatrième générations (qui ne figurent pas dans le tableau 2), les méthodes de première génération ont produit une valeur  $Z^+$  de  $0,18$  ( $k = 7$ ,  $N = 1\ 461$ ,  $CI_{Z^+} = 0,13$  à  $0,23$ ) pour la récidive violente, qui était une valeur de l'effet plus élevée que prévu selon la littérature sur le risque de récidive. Selon un examen des données, cette estimation était fondée sur seulement quatre études, dont chacune faisait appel à une approche différente, mais décrite vaguement, d'évaluation clinique du risque. De plus, quatre de ces effets provenaient d'une seule étude. Parmi toutes les générations d'outils, les mesures de quatrième génération ( $k = 5$ ,  $N = 3\ 759$ ) ont produit l'estimation prédictive la plus élevée ( $Z^+ = 0,52$ ,  $CI_{Z^+} = 0,49$  à  $0,55$ ) pour la récidive violente. Il convient de noter que la catégorie de la quatrième génération ne chevauchait pas les estimations des valeurs de l'effet de première, de deuxième et de troisième générations.



**Tableau 2**

*Comparaison des générations d'évaluation du risque pour la prédiction de la violence en établissement<sup>a</sup> et de la récidive violente<sup>b</sup>*

<i>Mesure</i>	<i>k</i>	<i>N</i>	<i>Z</i> <sup>+</sup>	<i>CI</i> <sub>Z<sup>+</sup></sub>	<i>Q</i>
<b>Violence en établissement</b>					
Deuxième génération	48	229 397	0,34	0,33 à 0,35	410,12*
Troisième génération	27	3 349	0,20	0,17 à 0,23	26,94*
<b>Récidive violente</b>					
Deuxième génération	92	19 874	0,18	0,17 à 0,19	328,13*
Troisième génération	81	15 233	0,23	0,21 à 0,25	247,38*

\*  $p < 0,05$ , indique qu'il y a plus de variabilité que ce que le hasard pourrait laisser supposer.

*N.B.* *k* = valeurs de l'effet par génération; *N* = délinquants par génération; *Z*<sup>+</sup> = valeur *r*' pondérée selon la taille de l'échantillon; *CI*<sub>Z<sup>+</sup></sub> = intervalle de confiance de 95% pour *Z*<sup>+</sup>.

<sup>a</sup> Il faut noter que seulement 75 des 76 valeurs de l'effet pour l'inconduite sont représentées. Une valeur de l'effet, produite par une mesure de quatrième génération, ne figure pas dans le tableau.

<sup>b</sup> Il faut noter que seulement 173 des 185 valeurs de l'effet pour la récidive sont représentées. Sept valeurs de l'effet produites par des mesures de première génération et cinq valeurs de l'effet produites par une mesure de quatrième génération ne figurent pas dans le tableau.

*Comparaisons fondées sur le contenu de l'outil : statique par rapport à dynamique*

Le tableau 3 résume l'exactitude prédictive des outils qui contenaient des éléments surtout statiques, des éléments surtout dynamiques et des éléments comprenant une combinaison égale d'éléments statiques et dynamiques. Dans le cas de la violence en établissement, les outils statiques ont produit l'effet moyen le plus important ( $Z^+ = 0,32$  comparativement à l'élément dynamique ( $Z^+ = 0,21$ ) et aux outils combinés ( $Z^+ = 0,23$ ), sans chevauchement des intervalles de confiance entre les catégories surtout statiques et surtout dynamiques. Selon l'indice garantissant la certitude des résultats, il faudrait 14 autres valeurs de l'effet statiques avec un  $r = 0$  pour ramener l'ampleur prédictive des outils statiques au niveau des outils dynamiques. De plus, il faudrait 10 études supplémentaires pour réduire l'estimation prédictive des outils statiques à celle produite par les outils combinés. En ce qui concerne la prédiction de la récidive violente, les outils dynamiques ont un léger avantage prédictif par rapport aux outils statiques comme en témoigne le chevauchement très minime des intervalles de confiance entre ces facteurs (c.-à-d.  $p < 0,05$ ). L'effet moyen des outils dynamiques était légèrement supérieur à celui des outils combinés, avec un très léger chevauchement des deux intervalles de confiance. Selon les calculs garantissant la certitude des résultats, 13 autres valeurs de l'effet dynamiques de zéro seraient nécessaires pour réduire sa validité prédictive au niveau de celle des mesures statiques. De plus, 24 autres valeurs de l'effet nulles seraient nécessaires pour réduire la capacité prédictive des mesures dynamiques au niveau de celle des mesures combinées.

**Tableau 3**

Comparaison des outils statiques et dynamiques pour la violence en établissement<sup>a</sup> et la récidive avec violence<sup>b</sup>

Mesure	$k$	$N$	$Z^+$	$CI_{z^+}$	$Q$
<b>Violence en établissement</b>					
Statique	26	226 026	0,32	0,316 à 0,324	210,48*
Dynamique	37	5 616	0,21	0,18 à 0,24	165,50*
Combinaison	12	1 029	0,23	0,17 à 0,29	14,36
<b>Récidive violente</b>					
Statique	64	13 409	0,22	0,20 à 0,24	152,17*
Dynamique	96	21 913	0,25	0,24 à 0,26	512,45*
Combinaison	13	1 697	0,20	0,15 à 0,25	28,53*

\*  $p < 0,05$ , indique qu'il y a plus de variabilité que ce que le hasard pourrait laisser supposer.

B.  $k$  = valeurs de l'effet par domaine de prédicteur;  $N$  = délinquants par domaine de prédicteur;  $Z^+$  = valeur  $r'$  pondérée selon la taille de l'échantillon;  $CI_{Z^+}$  = intervalle de confiance de 95 % pour  $Z^+$ .

<sup>a</sup> Il convient de noter que seulement 75 des 76 résultats de la violence en établissement sont représentés. On n'a pas pu déterminer la nature des prédicteurs pour l'une des valeurs de l'effet.

<sup>b</sup> Il convient de noter que seulement 173 des 185 des valeurs de l'effet de la récidive sont représentées. On n'a pas pu déterminer la nature des prédicteurs pour douze des valeurs de l'effet.

#### Comparaisons fondées sur la méthode d'administration des mesures

Des comparaisons des validités prédictives entre différentes méthodes d'administration sont présentées au tableau 4. En ce qui concerne la violence en établissement, la valeur  $Z^+$  la plus importante (0,34) a été attribuée à la catégorie de l'extraction de dossiers (c.-à-d. examen des dossiers). De plus, l'intervalle de confiance associé à cet effet moyen ne chevauchait pas du tout celui des catégories de l'auto-évaluation, de l'entrevue ou de la catégorie combinée dossier/entrevue. Le calcul garantissant la certitude des résultats a révélé qu'il faudrait ajouter 36 autres valeurs de l'effet nulles pour l'extraction des dossiers afin de réduire son effet moyen pour qu'il corresponde à celui de la catégorie de l'auto-évaluation; de plus, il faudrait 17 valeurs de l'effet nulles supplémentaires pour obtenir la parité avec la méthode combinée de l'examen du dossier et de l'entrevue et 46 valeurs de l'effet nulles pour obtenir l'égalité avec la méthode de l'entrevue. Quant aux résultats de la récidive, l'examen de la deuxième partie du tableau 4 a révélé que la méthode d'examen du dossier et de l'entrevue présentait la validité prédictive la plus élevée ( $Z^+ = 0,30$ ). L'intervalle de confiance pour cette catégorie ne concernait que les méthodes d'extraction de dossiers et ne partageait aucun chevauchement avec les deux autres méthodes. Pour réduire l'exactitude prédictive de la méthode dossiers/entrevue à celle des seules méthodes de l'examen des dossiers, des entrevues ou de l'auto-évaluation, on aurait besoin de 4, 47 ou 41 effets nuls de l'extraction de dossiers respectivement.

**Tableau 4**

*Comparaison des méthodes d'administration pour la prédiction de l'inconduite violente en établissement<sup>a</sup> et de la récidive violente<sup>b</sup>*

<i>Mesure</i>	<i>k</i>	<i>N</i>	<i>Z</i> <sup>+</sup>	<i>CI</i> <sub>Z<sup>+</sup></sub>	<i>Q</i>
<b>Violence en établissement</b>					
Extraction des dossiers	32	22 3071	0,34	0,336 à 0,344	209,14*
Entrevue	6	635	0,14	0,06 à 0,22	2,52
Auto-évaluation	13	2 505	0,16	0,12 à 0,20	21,63*
Dossier + entrevue	13	1 352	0,22	0,17 à 0,27	20,18
<b>Récidive violente</b>					
Extraction des dossiers	97	24 648	0,26	0,25 à 0,27	591,04*
Entrevue	21	2 921	0,11	0,07 à 0,15	24,58
Auto-évaluation	29	5 029	0,12	0,09 à 0,15	53,31*
Dossier + entrevue	27	5 741	0,30	0,27 à 0,33	100,13*

\*  $p < 0,05$ , indique qu'il y a plus de variabilité que ce que le hasard pourrait laisser supposer.

*N. B.*  $k$  = valeurs de l'effet par méthode;  $N$  = délinquants par méthode;  $Z^+$  = valeur  $r'$  pondérée selon la taille de l'échantillon;  $CI_{Z^+}$  = intervalle de confiance de 95 % pour  $Z^+$ .

<sup>a</sup> Il convient de noter que seulement 64 des 76 résultats de la violence en établissement sont représentés. On n'a pas pu déterminer la nature des méthodes d'administration pour douze des valeurs de l'effet.

<sup>b</sup> Il convient de noter que seulement 174 des 185 des valeurs de l'effet de la récidive sont représentées. On n'a pas pu déterminer la nature des méthodes d'administration pour onze des valeurs de l'effet.

#### *Comparaisons fondées sur la pertinence des outils d'évaluation pour les services correctionnels*

Une autre comparaison digne d'intérêt était la pertinence d'un outil d'évaluation pour les services correctionnels. Chaque valeur de l'effet a été codée pour déterminer si la mesure provenait d'une théorie criminologique et/ou si elle avait été créée expressément pour servir d'outil d'évaluation du risque. Par exemple, une mesure comme l'INS-R aurait été codée comme pertinente pour les services correctionnels parce qu'elle provenait de théories de la criminalité et qu'elle était créée pour servir d'outil d'évaluation du risque. Le VRAG était également codé comme pertinent parce que, même s'il n'avait pas été créé à partir d'une théorie, il avait été créé pour l'évaluation du risque.

Parmi les exemples d'outils non pertinents, il y avait ceux conçus pour évaluer des concepts comme les niveaux d'alphabétisation, l'estime de soi et les maladies psychiatriques. Le tableau 5 présente les résultats des outils pertinents par rapport aux outils non pertinents pour la prédiction des résultats violents. Les outils pertinents étaient de meilleurs prédicteurs de la violence en établissement et de la récidive, aucun chevauchement n'ayant été observé dans les intervalles de confiance avec les outils non pertinents. Selon les analyses garantissant la certitude des résultats, pour la violence en établissement, il faudrait 19 valeurs de l'effet supplémentaires nulles pour les outils pertinents afin de réduire leur performance prédictive à celle des mesures non pertinentes. Dans le cas de la récidive violente, il faudrait 415 valeurs de l'effet nulles supplémentaires pour les outils pertinents pour égaliser la validité prédictive entre les deux catégories.

**Tableau 5**

*Comparaison des mesures pertinentes par rapport aux mesures non pertinentes pour la prédiction de l'inconduite et de la récidive violentes<sup>a</sup>*

<i>Mesure</i>	<i>k</i>	<i>N</i>	<i>Z</i> <sup>+</sup>	<i>CI</i> <sub>Z<sup>+</sup></sub>	<i>Q</i>
<b>Violence en établissement</b>					
Pertinence	63	214 444	0,35	0,346 à 0,354	286,10*
Non-pertinence	13	18 346	0,27	0,26 à 0,28	144,25*
<b>Récidive violente</b>					
Pertinence	153	33 031	0,26	0,25 à 0,27	647,86*
Non-pertinence	25	5 835	0,07	0,04 à 0,10	56,93*

$p < 0,05$ , indique qu'il y a plus de variabilité que ce que le hasard pourrait laisser supposer.

*N. B.*  $k$  = valeurs de l'effet par catégorie de prédicteur;  $N$  = délinquants par catégorie de prédicteur;  $Z^+$  = valeur  $r'$  pondérée selon la taille de l'échantillon;  $CI_{Z^+}$  = intervalle de confiance de 95 % pour  $Z^+$ .

<sup>a</sup> Il convient de noter que seulement 178 des 185 valeurs de l'effet pour la récidive sont représentés. On n'a pas pu déterminer la pertinence des mesures pour sept des valeurs de l'effet.

## Discussion

La prédiction de la violence s'est révélée être une tâche difficile pour les spécialistes des services correctionnels et médico-légaux [Hanson (2005)]. Des prédicteurs assez robustes de la violence ont été définis [p. ex. Bonta et coll. (1998)] et des outils fondés sur ces facteurs ont été conçus pour aider à prédire le risque. Même si les spécialistes ont accès à une foule d'outils, ils font face à des difficultés lorsqu'ils tentent de déterminer lesquels de ces outils conviennent le mieux à l'évaluation du risque. Conformément aux principes du risque, du besoin et de la réceptivité [Andrews et Bonta (2003)], l'idéal serait de trouver un outil qui maximise la prédiction du risque tout en contribuant à la gestion des cas, à la planification de la réadaptation et aux progrès de la réduction du risque. Pour aider au processus décisionnel, la présente méta-analyse a servi à faire la synthèse des études portant sur les validités prédictives de divers outils utilisés pour prédire la violence. À partir des 88 études qui répondaient à nos critères d'inclusion, nous avons produit 185 valeurs de l'effet pour la récidive avec violence et 76 valeurs de l'effet pour l'inconduite violente en milieu carcéral. Groupée d'un outil à l'autre, la capacité modérée de prédire la récidive avec violence et l'inconduite violente en établissement était conforme aux estimations présentées dans d'autres méta-analyses sur la prédiction du risque [p. ex. Gendreau et coll. (1996); Gendreau et coll. (1997); Walters 2006)].

Il faut envisager la discussion qui suit en tenant compte des limites de la présente méta-analyse. Le premier ensemble de limites a trait à l'absence de renseignements nécessaires dans les études primaires visant à coder les variables importantes comme modératrices potentielles. Il s'agit d'un élément de frustration familial pour la plupart des méta-analystes. Dans l'étude en cours, 68 % des valeurs de l'effet étaient fondées sur des études dans lesquelles il n'y avait pas assez de renseignements pour coder ou définir le niveau des antécédents violents dans un échantillon particulier. De plus, aucune des études sur la violence en établissement n'a fourni de détails sur le niveau préexistant de violence en établissement de leur échantillon. Par conséquent, l'étendue de l'omission de données sur les antécédents de violence empêchait tout examen de cette variable comme modératrice de la valeur de l'effet. De plus, pour 56 % des valeurs de l'effet, il n'y avait pas assez de renseignements sur la nature des infractions à l'origine de la peine de l'échantillon (crimes de violence par rapport à crimes sans violence), ce qui a également empêché l'examen des effets modérateurs de la gravité de l'infraction à l'origine de la peine sur la validité prédictive. Il était également intéressant de noter que 21 valeurs de l'effet ont été obtenues à partir d'études qui n'indiquaient pas le sexe des membres de l'échantillon. Lorsque le sexe était indiqué, il était clair que la plupart des valeurs de l'effet provenaient d'échantillons d'hommes. Par conséquent, la généralisation des résultats actuels aux délinquantes ainsi qu'à d'autres sous-groupes de délinquants mal représentés (p. ex. les délinquants autochtones) est limitée. Une dernière question méthodologique, mais importante, était le fait que plus de 88 % des valeurs de l'effet provenaient d'échantillons définis comme présentant un risque faible ou modéré. Il est donc difficile de généraliser les conclusions actuelles aux échantillons à risque élevé à n'importe quel degré de certitude tant que d'autres données concernant cette population n'auront pas été obtenues.

### *Outils actuariels et structurés et échelle de psychopathie*

*Récidive violente.* Les outils fondés surtout sur les éléments de risque dynamiques ont produit la valeur de l'effet la plus élevée pour la prédiction de la récidive violente ( $Z^+ = 0,25$ , voir le tableau 3), et les intervalles de confiance pour cette catégorie ne concernaient que la valeur de l'effet moyenne pour la plupart des mesures des éléments de risque statiques ( $Z^+ = 0,22$ ). Combinés au résultat de l'indice

garantissant la certitude des résultats, ces constatations indiquent un léger avantage des outils d'évaluation du risque dynamiques par rapport aux outils statiques lorsqu'il s'agit de prédire la récidive violente. De même, les outils de troisième génération ont produit une meilleure estimation du risque de récidive violente que les mesures de deuxième génération (voir le tableau 2). Compte tenu des limites associées à seulement cinq valeurs de l'effet pour les mesures de quatrième génération, cette catégorie d'outils a produit les validités prédictives les plus fortes ( $Z^+ = 0,52$ ). Fait intéressant, la validité prédictive des évaluations cliniques non structurées de la récidive violente était plus élevée que prévu ( $Z^+ = 0,18$ ) étant donné l'opinion négative de cette approche dans la littérature. Cependant, il convient de noter que cet effet moyen était fondé sur sept valeurs de l'effet dans seulement quatre études différentes. De plus, quatre de ces valeurs de l'effet provenaient d'une seule étude [Rowe (1995)]. Il est possible que certains de ces jugements aient été portés en tenant compte de facteurs de risque criminogènes pertinents, car ces études avaient été réalisées au cours d'une période où des renseignements importants sur la prédiction du risque étaient disponibles (surtout dans les années 90). Cependant, les auteurs de ces quatre études n'ont pas suffisamment mentionné la façon de formuler une évaluation clinique ou une évaluation du risque. Il serait donc prématuré de conclure que les évaluations cliniques non structurées sont valables à ce moment-ci. Cela est particulièrement vrai à la lumière de la récente méta-analyse de Hanson et Morton-Bourgon (2007), qui démontre que la validité prédictive des prédictions cliniques non structurées du risque de violence chez les délinquants sexuels est généralement faible par rapport à celle des outils de prédiction actuariels.

L'examen des valeurs de l'effet moyennes pour ces outils d'évaluation avec au moins dix valeurs de l'effet disponibles montre clairement que chacune pouvait prédire la récidive violente avec au moins un degré modéré de succès (voir le tableau 1). Les valeurs de l'effet variaient de 0,22 pour l'échelle HCR-20 et l'échelle d'ISR à 0,32 pour le guide VRAG). L'INS/INS-R, l'échelle PCL-R et l'échelle d'ISR ont fourni les estimations ponctuelles les plus précises (c.-à-d. qu'ils ont produit l'intervalle de confiance le plus étroit), mais aucune mesure n'a ressorti comme étant la plus efficace pour la prédiction de la récidive violente. Le guide VRAG a donné de bons résultats, mais l'intervalle de confiance de la valeur de l'effet chevauchait avec ceux de l'INS-R et de l'échelle PCL-R. Selon les indices garantissant la certitude des résultats, seulement deux valeurs de l'effet nulles du guide VRAG seraient nécessaires pour réduire sa valeur moyenne au niveau de celle de l'INS-R et de l'échelle PCL-R. Par conséquent, elles s'appliquent probablement au même paramètre de population. De plus, la validité prédictive de l'INS-R équivalait à celle de l'échelle PCL-R et, dans une moindre mesure, à celle de l'échelle HCR-20 et de l'échelle d'ISR. Le guide VRAG a donné de meilleurs résultats que l'échelle HCR-20 et l'échelle d'ISR. Par conséquent, la plupart des mesures figurant au tableau 1 semblent avoir une capacité prédictive semblable, l'exception étant le guide VRAG, qui a un avantage prédictif seulement par rapport à l'échelle HCR-20 et à l'échelle d'ISR.

Les résultats actuels sont conformes aux études antérieures, selon lesquelles il existe une intercorrélation allant de modérée à élevée entre bon nombre des outils d'évaluation du risque fréquemment utilisés [p. ex. Glover et coll. (2002)]. Cela implique que ces mesures partagent une partie importante de la variance même si elles ne se chevauchent pas complètement. La ressemblance entre les outils a été examinée davantage dans une étude de Kroner, Mills et Reddon (2005), qui ont produit au hasard quatre mesures du risque hybrides fondées sur le contenu des éléments de l'échelle PCL-R, de l'INS-R, du guide VRAG et de l'ISGR. Les tests de chacune de ces mesures effectués en fonction de leur capacité de prédire la récidive générale ont révélé que les outils hybrides avaient un aussi bon rendement que chacun des outils d'origine respectifs. La présente méta-analyse a également permis de mettre à jour une conclusion

antérieure de Gendreau et coll. (2002), qui a signalé un léger avantage de l'INS-R par rapport à l'échelle PCL-R pour la prédiction de la récidive violente. L'inclusion de valeurs de l'effet supplémentaires publiées depuis la collecte des données de Gendreau et coll. donne à penser que l'échelle PCL-R et l'INS-R sont assez comparables en tant que prédicteurs de la récidive violente.

En général, malgré les préoccupations justifiables au sujet de la prédiction de la violence future et le débat permanent au sujet de la meilleure mesure, il y avait encore très peu de valeurs de l'effet disponibles pour examiner ces questions (c.-à-d. que le nombre le plus élevé a été obtenu pour l'échelle PCL-R à  $k = 24$ ). Les auteurs actuels font une mise en garde : il est probablement peu utile de produire de nouvelles mesures du risque pour le moment. La dernière chose dont a besoin le domaine de l'évaluation du risque, c'est de reproduire les efforts gaspillés dans la littérature sur la prédiction de la réhospitalisation en psychiatrie, dans laquelle 419 échelles ont été produites, dont seulement 3 ont permis d'établir plus de 10 estimations de la validité prédictive (Smith, Gendreau et Goggin, sous presse). Les études doivent plutôt porter sur la validation approfondie des mesures du risque actuelles dans différents contextes médico-légaux et sous-groupes de délinquants. Plus précisément, la majorité des valeurs de l'effet utilisées dans l'analyse actuelle étaient fondées sur un groupe générique d'échantillons de délinquants non psychiatriques, et la généralisabilité des constatations actuelles à des groupes de délinquants spécialisés exige une étude supplémentaire. Ces renseignements devraient montrer davantage les points forts et les points faibles d'une mesure individuelle. À notre avis, les chercheurs devraient tripler le nombre de valeurs de l'effet actuellement disponibles avant de continuer le débat sur la suprématie d'une mesure par rapport à une autre dans la prédiction de la récidive avec violence.

*Violence en établissement.* En ce qui concerne la violence en établissement, les outils de deuxième génération ont un avantage sur le plan de la prédiction par rapport aux mesures de troisième génération ( $Z^+ = 0,34$  contre  $0,20$ , respectivement). Par conséquent, les outils fondés sur les antécédents criminels et d'autres variables statiques étaient plus informatifs que les autres mesures pour l'estimation du risque de violence en établissement. En fait, selon l'étude actuelle, les outils fondés surtout sur des facteurs statiques étaient de meilleurs prédicteurs de la violence en établissement ( $Z^+ = 0,32$ ) que ceux fondés surtout sur des facteurs dynamiques ( $Z^+ = 0,21$ ), mais tous deux ont produit des estimations plus précises du risque que les outils qui comptaient des facteurs statiques et dynamiques combinés. L'ampleur  $Z^+$  pour l'outil combiné ne différait pas de façon significative de celle des outils fondés sur des facteurs dynamiques. Les échantillons de délinquants et de personnes nécessitant des services de psychiatrie légale étaient également représentés dans les valeurs de l'effet de la violence en établissement. En soi, la nature de l'échantillon ne contribue pas nécessairement à la force des outils de prédiction statiques pour ce résultat. Il est possible que les facteurs statiques soient plus utiles comme éléments de risque pour l'évaluation de la violence en établissement en raison de la durée à court terme de ces évaluations. La plupart de ces valeurs de l'effet étaient fondées sur des études dont les périodes de suivi étaient de moins d'un an tandis que l'inclusion des facteurs de risque dynamiques peut être plus pertinente pour les prédictions à long terme de la violence en établissement (comme dans le cas de la récidive, pour laquelle les périodes de suivi étaient plus longues).

Contrairement au cas de la prédiction de la récidive violente, la capacité des outils d'évaluation du risque individuels de prédire la violence en établissement varie beaucoup plus. Les indices des antécédents criminels ( $Z^+ = 0,26$ ) étaient de meilleurs prédicteurs de la violence en établissement que n'importe quelle des autres mesures, mais il s'agissait d'une catégorie passe-partout des diverses mesures ayant trait à la criminalité antérieure et il est donc difficile d'en interpréter la valeur. En ce qui concerne les mesures du risque uniformisées, le nombre le plus élevé de valeurs de l'effet pour la prédiction de la violence en

établissement a été obtenu pour l'échelle HCR-20 ( $Z^+ = 0,28$ ). Cependant, il convient de noter que malgré le rendement de l'échelle HCR-20, son application clinique pose des défis qui sont examinés plus loin. De plus, les données de l'échelle HCR-20 étaient surtout fondées sur des échantillons de personnes bénéficiant de services de psychiatrie légale, ce qui limite sa généralisabilité à la violence dans les établissements correctionnels non psychiatriques. En ce qui concerne les autres mesures, l'échelle PCL-VD ( $Z^+ = 0,22$ ) et l'INS-R ( $Z^+ = 0,24$ ) étaient modérément prédictifs de la violence en établissement tandis que le guide VRAG ( $k = 2$ ) et l'échelle PCL-R ( $k = 5$ ) qui étaient mal représentés, ont tous deux enregistré de petites associations avec ce résultat ( $Z^+ = 0,15$  et  $0,14$ , respectivement). Quelques études primaires sur la validité prédictive du guide VRAG pour la violence en établissement ont été publiées depuis l'achèvement de nos analyses (p. ex. Nadeau, Nadeau, Smiley et McHattie (1999); McDermott, Edens, Quanbeck, Busse et Scott, sous presse). L'inclusion de ces valeurs de l'effet supplémentaires dans les futures méta-analyses pourrait permettre de préciser le rôle du guide VRAG dans la prédiction de la violence en établissement. Conformément aux résultats actuels concernant l'échelle PCL-R, Guy, Edens, Anthony et Douglas (2005) ont constaté que l'échelle PCL-R a produit une valeur de l'effet pondérée moyenne de 0,17 pour l'agression physique en établissement. Dans leur analyse, l'échelle PCL-R était un meilleur prédicteur de l'agression verbale en milieu carcéral. Étant donné la popularité de ces deux dernières mesures, on espère que d'autres résultats concernant leur efficacité dans les établissements seront publiés. Par conséquent, sauf pour l'échelle HCR-20, l'INS-R et l'échelle PCL-VD, il est justifié de faire preuve de beaucoup de prudence dans le choix de l'outil utilisé pour prédire le risque en milieu carcéral jusqu'à ce que d'autres études prospectives aient été réalisées.

*Autres constatations pertinentes.* En général, les données actuelles corroborent l'inclusion de mesures d'auto-évaluation dans l'évaluation du risque de violence, mais non comme le seul moyen de prédiction. Plus précisément, la valeur de l'effet moyenne pour la catégorie générale des mesures d'auto-évaluation était petite pour la récidive violente et la violence en établissement ( $Z^+ = 0,12$  et  $0,16$ , respectivement) tandis que l'examen des dossiers et la méthode dossier/entrevue ont produit les validités prédictives les plus élevées pour les deux résultats. Parmi les mesures d'auto-évaluation, un outil qui a fait l'objet récemment d'un examen dans la littérature est le QAE. La prédiction par le QAE de la violence en établissement était fondée sur une seule valeur de l'effet, mais il pourrait avoir une certaine utilité pour la prédiction de ce résultat [ $r = 0,27$ , Loza et Loza-Fanous (2002)]. La valeur de l'effet moyenne pour le QAE était très prometteuse dans le cas de la récidive violente ( $Z^+ = 0,37$ ). L'avantage du QAE est qu'il contient six échelles qui permettent d'évaluer bon nombre des facteurs de risque définis empiriquement pour la récidive générale et la récidive violente [Bonta et coll. (1998); Gendreau et coll. (1996)], y compris les attitudes antisociales, les caractéristiques du trouble de la personnalité antisociale, les problèmes de comportement précoces, le comportement criminel antérieur, la toxicomanie et les fréquentations antisociales. Cet outil contient également un indice de validité et une sous-échelle de la colère [Loza et Loza-Fanous (2003)]. Comme l'a constaté Walters (2006), ce sont les mesures d'auto-évaluation pertinentes au contenu qui sont le plus susceptibles de produire des estimations plus précises du risque de violence et le QAE se classe dans cette catégorie. Néanmoins, étant donné le nombre limité de valeurs de l'effet du QAE, il faut effectuer des études primaires supplémentaires pour documenter sa capacité de prédire de manière prospective la violence et le soumettre à un test dans différents sous-groupes de délinquants (p. ex. les délinquantes).

Tout comme la méta-analyse de Walters (2006), les mesures servant à évaluer des concepts non pertinents ou sans rapport avec les résultats de la violence (p. ex. l'anxiété) sont généralement de piètres prédicteurs de la récidive violente dans la méta-analyse en cours. Même si la catégorie générale des mesures non



pertinentes a permis de mieux prédire la violence en établissement, elles étaient encore inférieures aux mesures pertinentes au contenu dans cette catégorie. De plus, relativement peu de données prospectives sont disponibles sur leur contribution à la violence en établissement. Une constatation digne de mention concernant l'ensemble de données actuel était le peu d'attention accordée à l'inventaire MMPI-2 en tant que prédicteur des résultats futurs sur la violence dans les études prospectives récentes. Cela est d'autant plus surprenant que l'inventaire MMPI était l'un des outils d'évaluation utilisés le plus souvent par les psychologues travaillant dans les établissements correctionnels aux États-Unis, même si peu de ces spécialistes s'employaient activement à effectuer une évaluation du risque [Boothy et Clements (2000)]. Bien qu'il n'existe pas de statistiques précises sur la fréquence de l'utilisation de l'inventaire MMPI-2 dans les évaluations du risque de violence au Canada, la pratique professionnelle indique que ce n'est certainement pas un phénomène rare. Nous avons recensé une seule étude, qui est maintenant vieillie, sur la validité prédictive de l'inventaire MMPI (typologie de Megargee) comme indice de la violence future [Motiuk (1991)]. Selon cette étude, cet inventaire était un piètre prédicteur de la récidive violente et son rendement était à peine meilleur comme prédicteur de la violence en établissement. Par conséquent, les évaluateurs doivent être prudents lorsqu'ils utilisent cet outil pour prendre des décisions sur le risque, étant donné le manque de données récentes sur la validité prédictive de la violence. Si l'on devait utiliser l'inventaire MMPI-2 pour une évaluation du risque de violence, il faudrait se limiter à la compréhension de la dynamique de la personnalité potentielle et aux problèmes de santé mentale qui peuvent s'appliquer aux préoccupations en matière de réceptivité.

#### *Recommandations pour guider la sélection des outils d'évaluation de la violence future*

Une question pratique importante qui se pose pour les évaluateurs spécialisés du risque et les spécialistes de la réadaptation est la sélection des meilleurs outils pour leur travail auprès des délinquants [Bonta (2002)]. Même si l'analyse actuelle indique qu'il n'y avait pas beaucoup de différence entre les validités prédictives des outils actuariels/structurés d'évaluation de la récidive avec violence, cela ne signifie pas qu'ils produiraient la même quantité de renseignements pour la planification des cas et l'établissement de stratégies d'intervention lorsque l'objectif est la réduction du risque. Compte tenu des défis que pose la sélection des outils appropriés pour l'évaluation du risque de violence, plusieurs recommandations ont été formulées. Ces recommandations étaient fondées sur la méta-analyse actuelle et les paramètres de la pratique professionnelle pertinents à la sélection des outils aux fins de l'évaluation du risque de violence [voir également Bonta (2002); Quinsey et coll. (1998)]. Ces recommandations font ressortir l'importance de tenir compte du contexte et de l'objectif de l'évaluation du risque demandée, du contenu et de la structure d'un outil d'évaluation du risque particulier dont l'utilisation est envisagée et la prise en considération de l'incorporation d'autres mesures pour l'évaluation des outils d'évaluation du risque uniformisés et l'élaboration de stratégies de réduction du risque.

*Première recommandation : déterminer le contexte et l'objet de l'évaluation du risque.* La sélection des outils devrait dépendre de l'objectif et du contexte du milieu auquel l'évaluation du risque doit s'appliquer. Par exemple, l'évaluation vise-t-elle à déterminer la possibilité qu'un délinquant soit traité, mis en liberté conditionnelle, placé dans un établissement ou que son niveau de sécurité soit déterminé? Ou bien vise-t-elle à évaluer les variations des facteurs de risque et criminogènes du délinquant au fil du temps? De plus, l'évaluation a-t-elle pour but d'informer les agents de gestion des cas du comportement d'un délinquant en milieu carcéral ou dans la collectivité? Par conséquent, il importe de choisir les outils conçus et testés expressément pour le contexte dans lequel ils doivent être utilisés. Par exemple, la prometteuse échelle ERV [Wong et Gordon (2006)] vise à aider à prendre des décisions concernant la candidature d'un délinquant à risque élevé à un programme de traitement des délinquants violents; le

Système de gestion interne des adultes [AIMS; Quay (1984)] vise à recenser et à gérer les délinquants qui risquent de recourir à la violence en milieu carcéral et l'INSR-R [Andrews et Bonta (1995)] a pour but de classer les délinquants incarcérés et dans la collectivité d'une manière qui permet de mieux faire correspondre l'intensité des services de surveillance et de réadaptation au niveau de risque et de besoin du délinquant. Par conséquent, pour déterminer les outils d'évaluation du risque pertinents, les évaluateurs doivent d'abord définir clairement le contexte et l'objectif de l'évaluation du risque que pose le délinquant.

La sélection des outils d'évaluation du risque soulève une question connexe : il faut déterminer si l'objectif de l'évaluation est la prédiction du risque ou l'évaluation des effets de la réadaptation. Si l'objectif final est la prédiction pure et qu'on s'intéresse peu à la planification des cas au-delà des décisions concernant le niveau de surveillance (p. ex. les décisions relatives au niveau de sécurité), alors presque tous les outils de deuxième génération conviendraient (p. ex. VRAG ou échelle d'ISR). Cependant, comme l'a soutenu Heilbrun (1997), l'évaluation du risque ne devrait pas se limiter à la prédiction. Par conséquent, si l'objectif consiste à prédire le risque *et* à contribuer à la planification de la réadaptation, un outil de troisième ou quatrième génération (p. ex. INS-R, NS-IGC, HCR-20) convient mieux.

*Deuxième recommandation : tenir compte du contenu et de la structure des outils d'évaluation du risque.* Du point de vue de l'évaluation, ce sont les concepts liés au risque qu'un outil est censé mesurer qui donnent une signification à l'évaluation du risque et fournissent un contexte pour interpréter cette évaluation [voir Kroner et coll. (2005)]. Par conséquent, l'évaluateur devrait choisir un outil fondé sur des concepts théoriques solides du risque et de la violence (concept du risque) tout en tenant compte des propriétés psychométriques de l'outil sur le plan de sa fiabilité interne, de la validité du concept et de la validité prédictive. Parmi les outils qui ne se limitent pas à la prédiction du risque de base, l'échelle HCR-20 et l'INS-R ont permis de prédire la violence à un degré modéré dans l'analyse actuelle, et tous deux ont été conçus pour s'appliquer théoriquement et empiriquement à l'évaluation du risque.

Même si l'INS-R a été conçu à l'origine pour l'évaluation de la récidive générale, et il s'agit de son point fort [pour un examen méta-analytique, voir Gendreau et coll. (2002)], la méta-analyse actuelle fait ressortir sa pertinence pour l'évaluation du risque de violence. Cela est probablement attribuable au chevauchement considérable des facteurs de risque et de besoins des délinquants violents et non violents [Bonta et coll. (1998)]. De plus, l'INS-R et l'échelle HCR-20 font appel à de multiples sources d'information pour l'évaluation des éléments individuels (c.-à-d. examens des dossiers, informateurs auxiliaires et entrevue avec le délinquant), ce qui est idéal pour une évaluation éclairée [Bonta (2002)]. Un autre avantage est que les deux outils peuvent tenir compte des variations du risque, car ils contiennent certains facteurs de risque dynamiques. Par conséquent, tous deux peuvent être utiles pour la réévaluation du risque au fil du temps.

Les éléments relatifs au risque et au besoin dans l'échelle HCR-20 et l'INS-R sont combinés pour permettre une estimation générale du risque, mais c'est à ce moment qu'apparaît une différence importante entre les deux mesures. Alors que l'INS-R utilise une estimation numérique du risque (examinée plus bas), l'échelle HCR-20 encourage l'évaluateur à éviter d'utiliser une totalisation numérique des résultats par élément pour obtenir l'estimation du risque. Selon Webster et coll. (1997b), la présence d'un seul élément de risque parmi les dix éléments historiques (p. ex. problèmes de toxicomanie, psychopathie), les cinq éléments cliniques (p. ex. manque d'intéroception, impulsivité) et les cinq éléments relatifs à la gestion du risque (p. ex. plans difficiles à réaliser, non-conformité aux tentatives de

correction) inclus dans leur plan structuré pouvait suffire à rendre un jugement de risque élevé dans un cas précis. Cela serait impossible dans un système numérique. Par conséquent, on conseille aux évaluateurs utilisant l'échelle HCR-20 de l'utiliser comme guide et non comme un outil actuariel absolu. Néanmoins, Webster et coll. (1997b) reconnaissent que le nombre plus élevé de facteurs de risque définis coïncide généralement avec un risque plus grand de violence [voir également Andrews et Bonta (2003)].

On a critiqué l'échelle HCR-20 parce qu'elle n'est rien d'autre qu'un retour aux évaluations cliniques du risque de première génération [Andrews et coll. (2006)] et, par conséquent, elle est sujette aux mêmes erreurs et biais [Edens et Otto (2001); Hilton et coll. (2006)]. Les études sur cet outil, toutefois, font ressortir un point de vue moins négatif. Selon la présente méta-analyse, les évaluations du risque numériques produites par l'échelle HCR-20 (bien qu'établies aux fins de recherche) étaient modérément prédictives de la violence future en milieu carcéral et dans la collectivité. D'après des études primaires antérieures, les évaluations structurées du risque (faible, moyen, élevé) de l'échelle HCR-20 ont également révélé des coefficients d'objectivité acceptables dans les analyses rétrospectives, au moins dans le cas des malades psychiatriques [de Vogel, de Ruitter, Hildebrand, Bos et Van de Ven (2004); Douglas, Ogloff et Hart (2003)]. Cela n'est pas surprenant, étant donné que l'évaluation du risque structurée dans l'échelle HCR-20 est censée être intégrée aux facteurs de risque historiques (c.-à-d. stables dans le temps). L'évaluation structurée peut ensuite être ajustée en fonction des facteurs dynamiques cliniques et de gestion du risque qui peuvent être présents [Webster et coll. (1997b)].

Comparativement à l'échelle HCR-20, l'INS-R exige une analyse plus exhaustive des facteurs de risque et de besoins. L'évaluateur doit évaluer 54 éléments, qui sont regroupés en 10 domaines de besoins et de facteurs de risque criminogènes, dont les antécédents criminels, les questions relatives à l'éducation et à l'emploi, les problèmes financiers, les préoccupations familiales/matrimoniales, le logement, les loisirs et activités récréatives, les compagnons, les problèmes d'alcool et de drogue, les problèmes affectifs et personnels et l'attitude ou l'orientation du délinquant [Andrews et Bonta (1995)]. Les évaluations des éléments individuels sont totalisées pour produire un score global des facteurs de risque et de besoins de l'INSR, qui est ensuite converti en percentile et comparé à des données normatives fournies dans le manuel pour les délinquants et les délinquantes. Selon le score du délinquant par rapport à ces données normatives, il est classé dans une catégorie descriptive allant de risque-besoins faibles à risque-besoins élevés [Andrews et Bonta (1995)]. On a modifié le contenu de l'INS-R dans le NS-IGC en réduisant le nombre de domaines de besoins et de facteurs de risque criminogènes à huit et le nombre d'éléments à 43 et en ajoutant le domaine « tendance antisociale », qui porte sur des aspects des caractéristiques psychopathiques et du trouble de la personnalité antisociale [Andrews et coll. (2004)]. Il y a également un espace pour noter les facteurs de protection ou les points forts du délinquant dans le NS-IGC et un nouveau formulaire sommaire sur la libération. Les nouveaux formulaires sur le contenu et l'évaluation du NS-IGC sont conçus pour intégrer plus efficacement l'évaluation du risque au processus de classification des délinquants et à la planification de la gestion des cas et à l'évaluation des progrès des délinquants.

Selon l'examen des éléments contenus dans l'échelle HCR-20 et l'INS-R (ainsi que le NS-IGC), ceux-ci se chevauchent dans le cas de certains éléments (p. ex. antécédents criminels, attitudes antisociales/négatives, problèmes d'emploi, toxicomanie). Toutefois, il y a des différences dignes de mention (p. ex. l'échelle HCR-20 met l'accent sur les questions cliniques). Une autre différence entre les deux mesures concerne le rôle de l'échelle PCL-R. L'échelle HCR-20 comprend une évaluation de la psychopathie au moyen de l'échelle PCL-R/VD contrairement à l'INS-R. Les outils fondés sur l'échelle PCL sont des tests psychologiques protégés et, à ce titre, ils ne sont pas mis à la disposition de

tout le personnel correctionnel et médico-légal qui peut être obligé de procéder à des évaluations du risque (p. ex. les agents de libération conditionnelle). Cela limite l'utilisation à des fins professionnelles de l'échelle HCR-20. Il est possible de ne pas tenir compte des éléments de l'échelle PCL-R/VD dans le mécanisme d'évaluation du risque de l'échelle HCR-20, mais les études ne sont pas constantes en ce qui concerne la façon dont cette exclusion pourrait influencer sur la validité prédictive du mécanisme. Plus particulièrement, Grann et coll. (2000) ont constaté que la suppression de l'élément PCL n'avait beaucoup d'impact sur la validité prédictive de la partie historique de l'échelle HCR-20 dans un suivi rétrospectif des délinquants souffrant d'un trouble de la personnalité et de troubles mentaux. Par contre, selon une étude rétrospective de Strand, Belfrage, Fransson et Lavender (1999), sauf pour l'élément PCL-R, la validité prédictive de la composante historique de l'échelle HCR-20 concernant le comportement violent chez les malades bénéficiant de services de psychiatrie légale a beaucoup diminué. Par conséquent, on ne sait pas clairement si le pouvoir prédictif de l'échelle HCR-20 est réduit par suite de l'exclusion de l'échelle PCL-R/VD. Un avantage du NS-IGC est que le domaine des tendances antisociales contient certains des éléments associés aux personnalités psychopathiques et antisociales, mais qu'il n'exige pas l'administration de mesures de l'échelle PCL.

En conclusion, les études effectuées jusqu'à maintenant donnent à penser que l'échelle HCR-20 est aussi efficace que l'INS-R pour la prédiction du comportement violent dans une population de délinquants générale ou bénéficiant de services de psychiatrie légale. Toutefois, les valeurs de l'effet incluses dans la présente méta-analyse étaient fondées sur les scores numériques de l'échelle HCR-20 et non sur les prédictions cliniques structurées dont l'utilisation est préconisée dans son application clinique. Même si les évaluateurs peuvent être tentés d'utiliser les scores numériques pour calculer les estimations du risque, il ne convient pas d'interpréter ces scores sans un groupe normatif approprié auquel ils peuvent être comparés. Par conséquent, il faudrait limiter l'utilisation clinique de l'échelle HCR-20 pour les populations de délinquants jusqu'à ce que d'autres études sur la validité prédictive de ses évaluations structurées soient effectuées. De plus, peu d'études prospectives de l'échelle HCR-20 ont été réalisées avec des échantillons de délinquants qui ne souffraient pas de troubles mentaux. En soi, l'applicabilité de l'échelle HCR-20 aux délinquants ne souffrant pas de troubles mentaux n'est pas claire. Enfin, bon nombre des études de validation de l'échelle HCR-20 ne font que tester les domaines historiques et de gestion du risque plutôt que les facteurs cliniques [Andrews et Bonta (2003)]. Par conséquent, il faut poursuivre les études sur les sous-composantes des échantillons de délinquants de la population générale et de délinquants bénéficiant de services de psychiatrie légale.

L'INS-R constitue une solution de rechange viable à l'échelle HCR-20 pour les évaluations générales du risque de violence qui mettent l'accent sur l'évaluation des facteurs criminogènes, la planification de la réadaptation et la mesure des progrès des délinquants dans la réduction du risque. Fait important, l'estimation ponctuelle de l'INS-R était plutôt concise pour la récidive avec violence relative à d'autres mesures. Une valeur de l'INS-R, relevée par Gendreau et coll. (2002), est la possibilité de l'appliquer aux délinquants qui ont commis des crimes de violence ou des infractions non violentes. Par conséquent, il y a un avantage sur le plan du rapport coût-efficacité et du temps à utiliser un outil qui permet d'évaluer la récidive générale et d'apporter une contribution significative à la prédiction de la violence. Malheureusement, l'INS-R et le NS-IGC ne contiennent que des comparaisons normatives pour la récidive générale, et non pour le risque de violence. Par conséquent, l'inclusion de normes distinctes pour la récidive avec violence ajouterait une valeur à la nature exhaustive de l'INS-R et du NS-IGC. Jusqu'à ce que ces normes soient établies, les évaluateurs devraient restreindre à la récidive générale leurs estimations des probabilités de risque fondées sur le NS-IGC et l'INS-R. Ils peuvent se servir de leur

connaissance des études empiriques pour indiquer comment ce risque peut aussi inclure la violence dans leurs descriptions qualitatives de l'estimation du risque. L'évaluateur devrait s'assurer de noter les limites pour l'estimation du risque de violence lorsqu'il communique des renseignements sur le risque à des tiers.

*Troisième recommandation : considérer les mesures pertinentes au contenu comme une source de renseignements pour les outils d'évaluation du risque et des besoins.* À un niveau plus général que la sélection de l'outil d'évaluation du risque, les évaluateurs devraient également décider de la pertinence des renseignements qui serviront à évaluer le risque et les besoins. Conformément au point de vue selon lequel les approches faisant appel à de multiples méthodes (p. ex. examen des dossiers, entrevue, contact avec des informateurs auxiliaires) accroissent l'exhaustivité des renseignements utilisés pour compléter les outils d'évaluation du risque et des besoins [Andrews et Bonta (2003); Andrews et coll. (2006)], la présente analyse donne à penser qu'il peut être avantageux de considérer les renseignements obtenus à partir de mesures d'auto-évaluation pertinentes au contenu comme faisant partie du processus d'évaluation. Cet argument est aussi conforme aux constatations de Walters (2006). Le recours à des mesures d'auto-évaluation pertinentes au contenu peut fournir des renseignements sur l'évaluation du risque qu'il est difficile d'évaluer à l'aide d'autres moyens, comme les attitudes antisociales. En particulier, le QAE peut ajouter de la crédibilité aux impressions cliniques subjectives d'un évaluateur concernant les distorsions cognitives et les schèmes de pensée antisociale d'un délinquant. Des indices de validité sur le QAE aident à minimiser les préoccupations concernant la gestion des impressions sur cette échelle. D'autres études, des mesures d'auto-évaluation pertinentes au contenu, comme le QAE, pourraient aussi servir d'outils de sélection dans le cadre de programmes d'intervention et de mesures préliminaires et postérieures pour la documentation des changements dans les domaines du risque et des besoins dissimulés. Seulement huit valeurs de l'effet étaient disponibles pour une analyse, mais il semble que le QAE puisse apporter une grande contribution à l'évaluation du risque. Malgré sa capacité prédictive, il est recommandé de faire preuve de prudence lorsqu'on utilise le QAE pour prédire le risque. Tout comme dans le cas de tout outil psychométrique, il faut interpréter son score dans le contexte d'autres renseignements connus sur l'individu en question. Les données actuelles indiquent clairement l'avantage de la méthode de l'examen des dossiers et de la méthode dossier/entrevue par rapport à l'auto-évaluation et aux entrevues pour la prédiction du risque. Par conséquent, le QAE peut fournir des renseignements très pertinents sur l'évaluation du risque et sur les facteurs de risque criminogènes contenus dans les autres outils d'évaluation du risque et des besoins.

Une autre mesure psychologique qui pourrait contribuer à l'évaluation et à l'interprétation de l'évaluation du risque est l'échelle PCL-R. Cette mesure ne vise pas à servir d'outil d'évaluation du risque, même si on l'utilise souvent à cette fin [Douglas, Vincent et Edens (2006)]. Elle a été conçue pour mesurer un concept de personnalité précis qui est lié modérément à la violence comme le montrent la présente méta-analyse et les précédentes [p. ex. Salekin, Rogers et Sewell (1996)]. C'est pourquoi les scores fondés sur l'échelle PCL ont été inclus comme éléments dans des outils de prédiction du risque comme l'échelle HCR-20 et le VRAG. Fait intéressant, selon les données actuelles, les échelles PCL-R et PCL-VD ont une capacité différente de prédire la violence en établissement. L'échelle PCL-VD, peut-être en raison du fait qu'elle met moins l'accent sur les éléments des antécédents criminels et davantage sur les traits de personnalité psychopatique, semblait se prêter mieux à l'évaluation de la violence en établissement que l'échelle PCL-R. L'échelle PCL-R a produit une estimation plus précise de la valeur de l'effet que l'échelle PCL-VD pour la récidive avec violence, mais la capacité prédictive de ces deux mesures était semblable pour ce critère. Elles ont produit un résultat semblable aux outils d'évaluation du risque standard (p. ex. VRAG, INS-R, HCR-20) pour la prédiction de la récidive avec violence. Par

conséquent, les échelles PCL-R et PCL-VD n'ajoutent peut-être pas de validité à la capacité des mesures actuarielles de prédire la violence [(Gendreau et coll. (2002)], mais elles peuvent contribuer aux stratégies de gestion de cas relatives aux délinquants psychopathes et à la solution de leurs problèmes de réceptivité plus difficiles [p. ex. égocentrisme, manipulation; Bonta (2002); Douglas et coll. (2006); Harris et Rice (2006)].

#### *Autres questions qui pourraient faire l'objet d'études à l'avenir*

En plus de la nécessité de poursuivre la validation empirique des mesures du risque actuelles et de la production de normes en matière de risque de violence pour l'INS-R et le NS-IGC, un autre domaine de recherche future est la définition des facteurs prédictifs de la nature et du contexte du comportement violent d'un délinquant. Une telle recherche pourrait permettre de déterminer les facteurs de risque grave/transitoire pertinents pour la détermination de l'imminence de la violence ou aider à porter un jugement sur la possibilité de diverses formes de comportement agressif [p. ex. réactif par rapport à instrumental; voir Quinsey et coll. (1998)]. Les aspects détaillés du risque de violence et les conditions dans lesquelles la violence est le plus susceptible de se produire sont sans doute plus utiles pour les surveillants des cas qu'un vague énoncé sur l'estimation générale du risque de violence. Compte tenu de la valeur de renseignements qualitatifs de ce genre pour la gestion du risque de violence, ces éléments qualitatifs sont souvent incorporés dans les formulations descriptives du risque dans les rapports spécialisés présentés aux intervenants. Malheureusement, malgré un degré raisonnable de compétence pour la prédiction de la probabilité générale d'un comportement violent avec l'aide d'outils appropriés, les évaluateurs ont plus de difficulté à prédire avec exactitude la probabilité de diverses dimensions du comportement violent [p. ex. gravité de l'agression, imminence probable du crime de violence, utilisation d'une arme; Douglas et coll. (2003)]. Par conséquent, il faut poursuivre les recherches sur les prédictions concernant la nature et la description qualitative de la violence afin de contribuer à la validité des descriptions qualitatives du risque fournies aux intervenants.

Un autre aspect souvent négligé de l'évaluation du risque de violence est l'incidence des facteurs de protection, qui peuvent modifier légèrement ou modérer l'association entre les facteurs criminogènes et le risque de violence [voir Rogers (2000)]. Cela est en partie attribuable aux difficultés de définir les facteurs de protection et de déterminer leurs interactions potentielles avec d'autres variables qui, ensemble, mènent à la réduction du risque. Hoge, Andrews et Leschied (1996) ont réalisé une des rares études sur cette question. Au moyen d'un échantillon de jeunes délinquants en probation, Hoge et ses collègues ont constaté que la présence de relations positives avec les pairs, de bonnes notes à l'école, une réponse positive à l'autorité et une utilisation efficace des temps libres étaient associées à des estimations plus faibles du risque dans une version pour les jeunes de l'INS-R. Les jeunes affichant ces facteurs de protection avaient également tendance à se conformer davantage aux conditions de leur probation, et l'effet de ces facteurs était uniforme pour les délinquants à faible risque et les délinquants à risque élevé. Par conséquent, les outils d'évaluation du risque qui incorporent des facteurs de protection permettraient une évaluation plus équilibrée du risque [voir Rogers (2000)]. Conçus comme des points forts, les facteurs de protection pourraient servir d'éléments constitutifs favorisant des changements menant à la réduction du risque. On s'intéresse de plus en plus à la reconnaissance des points forts des délinquants comme composante des outils d'évaluation du risque de quatrième génération [Andrews et coll., 2006]. Certains outils de prédiction structurés (p. ex. HCR-20) incitent également à tenir compte des facteurs de protection pour les évaluations finales du risque. Jusqu'à maintenant, il n'y a aucune ligne directrice précise sur la façon de procéder sans surestimer ou sous-estimer le risque. En outre, la littérature sur le

risque que constituent les jeunes délinquants et les délinquants adultes est assez limitée en ce qui concerne la capacité de définir des facteurs de protection précis.

### *Conclusion*

D'après les études disponibles sur sa validité prédictive prospective et sa capacité de contribuer à la gestion des cas et à la planification de la réadaptation, l'INS-R (et probablement le NS-IGC) se sont avérés être une option viable pour l'évaluation du risque de violence. Toutefois, ces outils ne contiennent pas de normes concernant le risque de violence comme dans le cas de la récidive générale. Par conséquent, il faut faire preuve de prudence lorsqu'on présente des énoncés sur la probabilité de violence pour la communication des résultats de l'évaluation à d'autres. Il faut poursuivre les recherches afin d'examiner cette limite. L'échelle HCR-20 a également révélé un potentiel important en tant qu'outil d'évaluation du risque, surtout auprès des populations nécessitant des services de psychiatrie légale. Cependant, il faut poursuivre les recherches sur la validité prospective des évaluations prédictives structurées de cet outil et ses prédictions numériques concernant les populations de délinquants générales. L'échelle PCL-R était un prédicteur modéré de la récidive violente tandis que l'échelle PCL-VD était un meilleur prédicteur de la violence en établissement. Le guide VRAG et l'échelle d'ISR permettent de prédire la récidive violente, mais ils ne sont pas sensibles aux changements du niveau de risque qui se produisent avec le temps et(ou) aux efforts de réadaptation en raison de leur caractère statique. En ce qui concerne les mesures d'auto-évaluation, il peut être utile d'inclure le QAE comme outil dans le protocole d'évaluation du risque pour contribuer à l'évaluation des facteurs criminogènes. Néanmoins, il ne faut pas recourir uniquement à l'auto-évaluation et aux entrevues pour évaluer le risque. Enfin, les évaluateurs doivent se méfier des outils qui n'ont pas été entièrement validés comme prédicteurs significatifs de la violence *future* dans la population de délinquants où ils projettent de les appliquer. Nous encourageons également les chercheurs à continuer d'établir une base de données sur les études prospectives concernant la prédiction de la violence dans divers milieux carcéraux et diverses populations.





## Bibliographie

Les entrées marquées d'un astérisque (\*) figuraient dans la méta-analyse.

- ANDREWS, D. A. Il est possible de prévoir et d'influencer la récidive : utiliser des outils de prédiction du risque afin de réduire la récidive, *Forum – Recherche sur l'actualité correctionnelle*, vol. 1, p. 11-18, 1989.
- ANDREWS, D. A. et J. BONTA. *Level of Service Inventory-Revised*, Toronto, Ontario, Canada, Multi-Health Systems, 1995.
- ANDREWS, D. A. et J. BONTA. *The psychology of criminal conduct (3<sup>rd</sup> edition)*, Cincinnati, Ohio, Anderson, 2003.
- ANDREWS, D. A., J. BONTA et R. D. HOGE. « Classification for effective rehabilitation », *Rediscovering psychology, Criminal Justice and Behavior*, vol. 17, p. 19-52, 1990.
- ANDREWS, D. A., J. BONTA et J. S. WORMITH. *The Level of Service/Case Management Inventory (LS/CMI)*, Toronto, Ontario, Canada, Multi-Health Systems, 2004.
- ANDREWS, D. A., J. BONTA et J. S. WORMITH. « The recent past and near future of risk and/or need assessment », *Crime and Delinquency*, vol. 52, p. 7-27, 2006.
- ANDREWS, D. A. et J. S. WORMITH. *The Criminal Sentiments Scale*, Ottawa, Ontario, Canada, ministère des Services correctionnels du Canada, 1984.
- BAIRD, S. C. « Classifying juveniles: Making the most of an important management tool », *Corrections Today*, vol. 47, p. 32-38, 1981.
- BAIRD, S. C., R. C. HEINZ et B. J. BEMUS. *The Wisconsin Case Classification/Staff development project: A two-year follow-up report*, Madison, Wisconsin, Wisconsin Bureau of Community Corrections, 1979.
- \*BELFRAGE, H., G. FRANSSON et S. STRAND. « Prediction of violence using the HCR-20: A prospective study in two maximum-security correctional institutions », *Journal of Forensic Psychiatry*, vol. 11, p. 167-175, 2000.
- \*BLANCHETTE, K. *Field-test of a gender-informed security re-classification scale for female offenders*, thèse de doctorat non publiée, Carleton University, Ottawa, Ontario, Canada, 2005.
- \*BLANCHETTE, K., P. VERBRUGGE et C. G. WICHMANN. *Les femmes purgeant une peine de ressort fédéral, Échelle de classement par niveau de sécurité et placement pénitentiaire*, rapport de recherche n° 127, Ottawa (Ontario), Canada, Service correctionnel du Canada, 2002.
- BOER, D. P., S. D. HART, P. R. KROPP et C. D. WEBSTER. *Manual for the Sexual Violence Risk-20: Professional guidelines for assessing risk for sexual violence*, Vancouver, Colombie-Britannique, Canada, The British Columbia Institute Against Family Violence, 1997.

- \*BONTA, J. « Native inmates: Institutional response, risk, and needs », *Revue canadienne de criminologi*, vol. 31, p. 49-61, 1989.
- BONTA, J. « Offender risk assessment: Guidelines for selection and use », *Criminal Justice and Behavior*, vol. 29, p. 355-379, 2002.
- \*BONTA, J., W. G. HARMAN, R. G. HANN et R. B. CORMIER. « The prediction of recidivism among federally sentenced offenders: A re-validation of the SIR scale », *Revue canadienne de criminology*, vol. 38, p. 61-79, 1996.
- BONTA, J., M. LAW et K. HANSON. « The prediction of criminal and violent recidivism among mentally disordered offenders: A meta-analysis », *Psychological Bulletin*, vol. 123, p. 123-142, 1998.
- \*BONTA, J. et L. L. MOTIUK. *Use of the Level of Supervision Inventory for assessing incarcerates*, document présenté à l'assemblée annuelle de l'American Psychological Association, Washington (D.C.), 1986.
- \*BONTA, J. et A. K. YESSINE. *Système national de repérage : identification des délinquants violents à risque élevé*, rapport n° 2005-04, Sécurité publique et Protection civile Canada, Ottawa (Ontario), Canada, 2005.
- BOOTHY, J. L. et C. B. CLEMENTS. « A national survey of correctional psychologists », *Criminal Justice and Behavior*, vol. 27, p. 715-731, 2000.
- BORUM, R. « Improving the clinical practice of violence risk assessment: Technology, guidelines, and training », *American Psychologist*, vol. 51, p. 945-956, 1996.
- CAPALDI, D. M. et G. R. PATTERSON. *The violent adolescent male: Specialist or generalist?*, document présenté à la 60<sup>e</sup> réunion biennale de la Society for Research in Child Development, Nouvelle-Orléans (Louisiane), du 25 au 28 mars 1993.
- \*COLLIE, R. M. et D. L. L. POLASCHEK. « Sorting women's risk: New Zealand women prisoners' misconducts and internal security risk », *New Zealand Journal of Psychology*, vol. 32, p. 101-109, 2003.
- \*COOKE, D. « Predicting offending in prison: The predictive validity of the prison behavior rating scales », *Legal and Criminological Psychology*, vol. 1, p. 65-82, 1996.
- CUMMING, G. et S. FINCH. « A primer on the understanding, use, and calculation of confidence intervals that are based on central and noncentral distributions », *Educational and Psychological Measurement*, vol. 61, p. 532-574, 2001.
- \*CUNNINGHAM, M. D et J. R. SORENSEN. « Actuarial models for assessing prison violence risk: Revisions and extensions of the Risk Assessment Scale for Prison (RASP) », *Assessment*, vol. 13, p. 253-265, 2006

- \*CUNNINGHAM, M. D., J. R. SORENSEN et T. J. REIDY. « An actuarial model for assessment of prison violence risk among maximum security inmates », *Assessment*, vol. 12, p. 40-49, 2005.
- \*DAFFERN, M., J. R. P. OGLOFF, M. FERGUSON et L. THOMSON. « Assessing risk for aggression in a forensic psychiatric hospital using the *Level of Service Inventory-Revised*,: Screening Version », *International Journal of Forensic Mental Health*, vol. 4, p. 201-206, 2005.
- \*DAHLE, K-P. « Strengths and limitations of actuarial prediction of criminal reoffence in a German prison sample: A comparative study of LSI-R, HCR-20, and PCL-R », *International Journal of Law and Psychiatry*, vol. 29, p. 431-442, 2006.
- DE VOGEL, V., C. DE RUITER, M. HILDEBRAND, B. BOS ET P. VAN DE VEN. « Type of discharge and risk of recidivism measured by the HCR-20: A retrospective study in a Dutch sample of treated forensic psychiatric patients », *International Journal of Forensic Mental Health*, vol. 3, p. 149-165, 2004.
- \*DERNEVIK, M., M. GRANN et S. JOHANNSON. « Violent behaviour in forensic psychiatric patients: Risk assessment and different risk-management levels using the *HCR-20* », *Psychology, Crime, & Law*, vol. 8, p. 93-111, 2002.
- \*DOLAN, M. et A. KHAWAJA. « The HCR-20 and post-discharge outcome in male patients discharged from medium security in the UK », *Aggressive Behaviour*, vol. 30, p. 469-483, 2004.
- DOUGLAS, K. S., S. D. HART et P. R. KROPP. « Validity of the Personality Assessment Inventory for forensic assessments », *International Journal of Offender Therapy and Comparative Criminology*, vol. 45, p. 183-197, 2001.
- \*DOUGLAS, K. S., J. R. P. OGLOFF et S. D. HART. « Examination of a model of violence risk assessment among forensic psychiatric patients », *Psychiatric Services*, vol. 54, p. 1372-1379, 2003.
- DOUGLAS, K. S. et J. SKEEM. « Violence risk assessment: Getting specific about being dynamic », *Psychology, Public Policy, and Law*, vol. 11, p. 347-383, 2005.
- DOUGLAS, K. S., G. M. VINCENT et J. F. EDENS. « Risk for criminal recidivism: The role of psychopathy », sous la direction de C. J. Patrick, *Handbook of psychopath*, New York, Guilford, p. 533-554, 2006.
- DOUGLAS, K. S., N. YEOMANS et D. P. BOER. Comparative validity analysis of multiple measures of violence risk in a sample of criminal offenders, *Criminal Justice and Behavior*, vol. 32, p. 479-510, 2005.
- DOWDEN, C. et D. A. ANDREWS. « Effective correctional treatment and violent re-offending: A meta-analysis », *Revue canadienne de criminologie*, p. 449-467, 2000.
- \*DOYLE, M., M. DOLAN et J. MCGOVERN. « The validity of North American risk assessment tools in predicting in-patient violent behaviour in England », *Legal and Criminological psychology*, vol. 7, p. 141-154, 2002.

- EDENS, J. F. et R. K. OTTO. « Release decision-making and planning », sous la direction de J. B. Ashford, B. D. Sales et W. Reid, *Treating adult and juvenile offenders with special needs*, p. 335-371, Washington (D.C.), American Psychological Association, 2001.
- \*EDENS, J. F. et M. A. RUIZ. « On the validity of validity scales: The importance of defensive responding in the prediction of institutional misconduct », *Psychological Assessment*, vol. 18, p. 220-224, 2006.
- FRENCH, S. et P. GENDREAU. « Reducing prison misconducts: What works! », *Criminal Justice and Behavior*, vol. 33, p. 185-218, 2006.
- \*FUJII, D. E., A. B. TOKIOKA, A. I. LICHTON et E. HISHINUMA. « Ethnic differences in prediction of violence risk with the HCR-20 among psychiatric inpatients », *Psychiatric Services*, vol. 56, p. 711-716, 2005.
- \*GAGLIARDI, G. J., D. LOVELL, P. D. PETERSON et R. JEMELKA. « Forecasting recidivism in mentally ill offenders released from prison », *Law and Human Behavior*, vol. 28, p. 133-155, 2004.
- GENDREAU, P., C. GOGGIN, S. FRENCH et P. SMITH. « Practicing psychology in correctional settings », sous la direction de I. B. Weiner et A. K. Hess, *The handbook of forensic psychology*, troisième édition, p. 722-750, Hoboken (New Jersey), John Wiley & Sons, 2006.
- GENDREAU, P., C. GOGGIN et M. A. LAW. « Predicting prison misconducts », *Criminal Justice and Behavior*, vol. 24, p. 414-431, 1997.
- GENDREAU, P., C. GOGGIN et P. SMITH. « Generating rational correctional policies: An introduction to advances in cumulating knowledge », *Corrections Management Quarterly*, vol. 4, p. 52-60, 2000.
- GENDREAU, P., C. GOGGIN et P. SMITH. « Is the PCL-R really the “unparalleled” measure of offender risk? A lesson in knowledge cumulation », *Criminal Justice and Behavior*, vol. 29, p. 397-426, 2002.
- GENDREAU, P., T. LITTLE et C. GOGGIN. « A meta-analysis of the predictors of adult recidivism: What works! », *Criminology*, vol. 34, p. 401-433, 1996.
- \*GIRARD, L. *The Level of Service Inventory-Ontario Revision: Risk/need assessment and recidivism*, thèse de doctorat non publiée, Université d’Ottawa (Ontario), Canada, 1999.
- \*GLOVER, A. J. J., D. E. NICHOLSON, T. HEMMATI, G. A. BERFELD et V. L. QUINSEY. « A comparison of predictors of general and violent recidivism among high risk federal offenders », *Criminal Justice and Behavior*, vol. 29, p. 235-249, 2002.
- GRANN, M., H. BELFRAGE et A. TENGSTRÖM. « Actuarial assessment of risk for violence: Predictive validity of the VRAG and the historical part of the HCR-20 », *Criminal Justice and Behavior*, vol. 27, p. 97-114, 2000.

- \*GRANN, M., N. LÅNGSTRÖM, A. TENGSTÖM et G. KULLGREN. « Psychopathy (PCL-R) predicts violent recidivism among criminal offenders with personality disorders in Sweden », *Law and Human Behavior*, vol. 23, p. 205-218, 1999.
- \*GRAY, N. S., C. HILL, A. MCGLEISH, D. TIMMONS, M. J. MACCULLOCH et R. J. SNOWDEN. « Prediction of violence and self-harm in mentally disordered offenders: A prospective study of the efficacy of the HCR-20, PCL-R, and psychiatric symptomatology », *Journal of Consulting and Clinical Psychology*, vol. 71, p. 443-451, 2003.
- \*GRAY, N. S., R. J. SNOWDEN, S. MACCULLOCH, H. PHILLIPS, J. TAYLOR et T. J. MACCULLOCH. « Relative efficacy of criminological, clinical, and personality measures of future risk of offending in mentally disordered offenders: A comparative study of HCR-20, PCL:SV, and OGRS », *Journal of Consulting and Clinical Psychology*, vol. 72, p. 523-530, 2004.
- \*GREVATT, M., B. THOMAS-PETER, et G. HUGHES. « Violence, mental disorder and risk assessment: Can structured clinical assessments predict the short-term risk of inpatient violence? », *Journal of Forensic Psychiatry and Psychology*, vol. 15, p. 278-292, 2004.
- GROVE, W. M., D. H. ZALD, B. S. LEBOW, B. E. SNITZ et C. NELSON. « Clinical versus mechanical prediction: A meta-analysis », *Psychological Assessment*, vol. 12, p. 19-30, 2000.
- \*HALSALL, S. et S. VAN DINE. « Security reclassification risk instrument for males: Development and methodology », *Ohio Corrections Research Compendium*, Columbus (Ohio), Ohio Department of Rehabilitation and Correction, vol. 1, p. 95-96, 2002.
- HANSON, R. K. « Twenty years of progress in violence risk assessment », *Journal of Interpersonal Violence*, vol. 20, p. 212-217, 2005.
- HANSON, R. K. et K. E. MORTON-BOURGON. *L'exactitude des évaluations du risque de récidive chez les délinquants sexuels : une méta-analyse*, Sécurité publique et Protection civile Canada, 2007.
- \*HANSON, R. K. et S. WALLACE-CAPRETTA. *Prédire la récidive chez les hommes violents envers leur partenaire*, rapport pour spécialistes 2000-06, Ottawa (Ontario), Solliciteur général Canada, 2000.
- HARE, R. D. *The Hare Psychopathy Checklist - Revised*, Toronto (Ontario), Canada, Multi-Health Systems, 1991.
- HARE, R. D. *The Hare Psychopathy Checklist – Revised*, 2<sup>e</sup> édition, Toronto (Ontario), Canada, Multi-Health Systems, 2003.
- \*HARER, M. D. et N. P. LANGAN. « Gender differences in predictors of prison violence: Assessing the predictive validity of a risk classification system », *Crime and Delinquency*, vol. 47, p. 513-536, 2001.

- HARRIS, G. T. et M. E. RICE. « Treatment of psychopathy: A review of empirical findings », sous la direction de C. J. Patrick, *Handbook of Psychopathy*, p. 555-572, New York, Guilford, 2006.
- \*HARRIS, G. T., M. E. RICE et C. A. CORMIER. « Psychopathy and violent recidivism », *Law and Human Behavior*, vol. 15, p. 626-637, 1991.
- \*HARRIS, G. T., M. E. RICE et C. A. CORMIER. « Prospective replication of the Violence Risk Appraisal Guide in predicting violent recidivism among forensic patients », *Law and Human Behavior*, vol. 26, p. 377-394, 2002.
- \*HARRIS, G. T., M. E. RICE et V. L. QUINSEY. « Violent recidivism of mentally disordered offenders: The development of a statistical prediction instrument », *Criminal Justice and Behavior*, vol. 20, p. 315-335, 1993.
- HATHAWAY, S. R. et J. C. MCKINLEY. *Inventaire multiphasique de la personnalité Minnesota*, Institut de recherches psychologiques, Montréal, 1977.
- HEDGES, L. V. et I. OLKIN. *Statistical methods for analysis*, Toronto (Ontario), Canada, Academic Press, 1985.
- HEILBRUN, K. « Prediction versus management models relevant to risk assessment: The importance of legal decision-making context », *Law and Human Behavior*, vol. 21, p. 347-359, 1997.
- \*HEILBRUN, K., S. D. HART, R. D. HARE, D. GUSTAFSON, C. NUNEZ et A. J. WHITE. « Inpatient and post-discharge aggression in mentally disordered offenders: The role of psychopathy », *Journal of Interpersonal Violence*, vol. 13, p. 514-552, 1998.
- \*HEMPHILL, J. F. *Psychopathy, criminal history, and recidivism*, thèse de doctorat non publiée, Université de la Colombie-Britannique, Canada, 1998.
- HEMPHILL, J. F., R. D. HARE et S. WONG. « Psychopathy and recidivism: A review », *Legal and Criminological Psychology*, vol. 3, p. 139-170, 1998.
- \*HEMPHILL, J. F., R. TEMPLEMAN, S. WONG et R. D. HARE. « Psychopathy and crime: Recidivism and criminal careers », sous la direction de D. J. Cooke, A. E. Forth et R. D. Hare, *Psychopathy: Theory, research and implications for society*, p. 375-399, Dordrecht (Pays-Bas), Kluwer Academic, 1998.
- \*HILDEBRAND, M., C. DE RUITER et H. NIJMAN. « PCL-R psychopathy predicts disruptive behavior among male offenders in a Dutch forensic psychiatric hospital », *Journal of Interpersonal Violence*, vol. 19, p. 13-29, 2004.
- HILTON, N. Z., G. T. HARRIS et M. E. RICE. « Sixty-six years of research on clinical versus actuarial prediction of violence », *The Counseling Psychologist*, vol. 34, p. 400-409, 2006.
- HOFFMAN, P. B. « Screening for risk: A revised salient factor score (SFS 81) », *Journal of Criminal Justice*, vol. 11, p. 539-547, 1983.

- HOGUE, R. D. et D. A. ANDREWS. *Assessing the youthful offender: Issues and techniques*, New York, Plenum, 1996.
- HOGUE, R. D., D. A. ANDREWS et A. W. LESCHIEDA. « An investigation of risk and protective factors in a sample of youthful offenders », *Journal of Child Psychology and Psychiatry*, vol. 37, p 419-424, 1996
- \*HOLLAND, T. R., N. HOLT, M. LEVI et G. E. BECKETT. « Comparison and combination of clinical and statistical predictions of recidivism among adult offenders », *Journal of Applied Psychology*, vol. 68, p. 203-211, 1983.
- \*JAYJOHN, J. et S. VAN DINE. « Standard initial security classification risk instrument for males: Development and methodology », *Ohio Corrections Research Compendium*, Ohio Department of Rehabilitation and Correction, Columbus (Ohio), vol. 1, p. 93-94, 2002.
- \*JEMELKA, R. P., G. A. WIEGAND, E. A. WALKER et E. W. TRUPIN. « Computerized offender assessment: Validation study », *Psychological Assessment*, vol. 4, p. 138-144, 1992.
- \*KRONER, D. G. et W. LOZA. « Evidence for the efficacy of self-report in predicting nonviolent and violent criminal recidivism », *Journal of Interpersonal Violence*, vol. 16, p. 168-178, 2001.
- \*KRONER, D. G. et J. F. MILLS. « The accuracy of five risk appraisal instruments in predicting institutional misconduct and new convictions », *Criminal Justice and Behavior*, vol. 28, p. 471-489, 2001.
- KRONER, D. G., J. F. MILLS et J. R. REDDON. « A coffee can, factor analysis, and prediction of antisocial behaviour: The structure of criminal risk », *International Journal of Law and Psychiatry*, vol. 28, p. 360-374, 2005.
- KROPP, P. R., S. D. HART, C. D. WEBSTER et D. EAVES. *Manual for the Spousal Assault Risk Assessment Guide*, 2<sup>e</sup> édition, Vancouver (Colombie-Britannique), Canada, British Columbia Institute of Family Violence, 1995.
- \*LAW, M. A. *A longitudinal follow-up of federally sentenced women in the community: Assessing the predictive validity of the dynamic characteristics of the community intervention scale*, thèse de doctorat non publiée, Carleton University, Ottawa (Ontario), Canada, 2004.
- LEY, B. A. *Quantitative aspects of psychological assessment: An introduction*, London, Duckworth, 1972.
- LITWACK, T. R. et L. B. SCHLESINGER. « Dangerousness risk assessments: Research, legal, and clinical considerations », sous la direction de A. K. Hess et I. B. Weiner, *The handbook of forensic psychology*, p. 171-217, New York, Wiley, 1999.
- \*LOVELL, D., G. GAGLIARDI et P. PHIPPS. *Washington's dangerous mentally-ill offender law: Was community safety increased?*, Report #05-03-10901, Olympia (Washington), Washington State Institute for Public Policy, 2005.

- LOZA, W., G. K. DHALIWAL, D. G. KRONER et A. LOZA-FANOUS. « Reliability and concurrent validity of the Self-Appraisal Questionnaire (SAQ): A tool for assessing violent and non-violent recidivism », *Criminal Justice and Behavior*, vol. 27, p. 356-374, 2000.
- \*LOZA, W. et K. GREEN. « The Self Appraisal Questionnaire: A self-report measure for predicting recidivism versus clinician-administered measures: A five-year follow-up study », *Journal of Interpersonal Violence*, vol. 18, p. 781-797, 2003.
- \*LOZA, W. et A. LOZA-FANOUS. *Predictive validity of the LSI: A tool for assessing violent and non-violent recidivism*, manuscrit non publié, 1999b.
- \*LOZA, W. et A. LOZA-FANOUS. « Predictive validity of the Self-Appraisal Questionnaire (SAQ): A tool for assessing violent and nonviolent release failures », *Journal of Interpersonal Violence*, vol. 15, p. 1183-1191, 2000.
- \*LOZA, W. et A. LOZA-FANOUS. « The effectiveness of the Self-Appraisal Questionnaire in predicting offenders' post-release outcome: A comparison study », *Criminal Justice and Behavior*, vol. 28, p. 105-121, 2001.
- \*LOZA, W. et A. LOZA-FANOUS. « The effectiveness of the Self-Appraisal Questionnaire (SAQ) as an offenders' classification measure », *Journal of Interpersonal Violence*, vol. 17, p. 3-14, 2002.
- \*LOZA, W. et A. LOZA-FANOUS. « More evidence for the validity of the Self Appraisal Questionnaire for predicting violent and nonviolent recidivism: A five-year follow-up study », *Criminal Justice and Behavior*, vol. 30, p. 709-721, 2003.
- \*LOZA, W., D. B. VILLENEUVE et A. LOZA-FANOUS. « Predictive validity of the Violence Risk Appraisal Guide: A tool for assessing violent offenders' recidivism », *International Journal of Law and Psychiatry*, vol. 25, p. 85-92, 2002.
- \*LUCIANI, F., L. MOTIUK et M. NAFEKH. *Examen opérationnel de la fiabilité, de la validité et de l'utilité pratique de l'Échelle de classement par niveau de sécurité*, rapport de recherche n° R-47, Ottawa (Ontario), Service correctionnel Canada, 1996.
- MARTIN, G. et J. PEAR. *Behaviour modification: What it is and how to do it*, 8<sup>e</sup> édition, Upper Saddle River (New Jersey), Pearson/Prentice Hall, 2007.
- \*MCHATTIE, L., A. HILLS, C. SMILEY et R. MACKENZIE. *Does the CATS predict recidivism as well as the PCL-R?*, Abbotsford (Colombie-Britannique), Canada, Centre régional de santé, 1999.
- MCGRAW, K. O. et S. P. WONG. « A common language effect size statistic », *Psychological Bulletin*, vol. 111, p. 361-365, 1992.
- MEGARGEE, E. et M. J. BOHN Jr. *Classifying criminal offenders: A new system based on the MMPI*, Beverly Hills (Californie), Sage, 1999.



- MEGARGEE, E. I. et J. L. CARBONELL. « Use of the MMPI-2 in correctional settings », sous la direction de Y. S. Ben-Porath, J. R. Graham, G. C. N. Hall, R. D. Hirschman et M. S. Zaragoza, *Forensic applications of the MMPI-2*, p. 127-159, Thousand Oaks (Californie), Sage, 1995.
- \*MILLS, J. F., M. N. JONES et D. G. KRONER. An examination of the generalizability of the LSI-R and VRAG probability bins, *Criminal Justice and Behavior*, vol. 32, p. 565-585, 2005.
- \*MILLS, J. F. et D. G. KRONER. « The effect of discordance among violence and general recidivism risk estimates on predictive accuracy », *Criminal Behaviour and Mental Health*, vol. 16, p. 155-166, 2006.
- \*MILLS, J. F., D. G. KRONER et T. HEMMATI. « The Measures of Criminal Attitudes and Associates (MCAA): The prediction of general and violent recidivism », *Criminal Justice and Behavior*, vol. 31, p. 717-733, 2004.
- \*MILLS, J. F., W. LOZA et D. G. KRONER. « Predictive validity despite social desirability: Evidence of the robustness of self-report among offenders », *Criminal Behaviour and Mental Health*, vol. 13, p. 140-150, 2003.
- MONAHAN, J. et H. STEADMAN (dir.). *Violence and mental disorder: Developments in risk assessment*, Chicago, University of Chicago Press, 1994.
- MOREY, L. C. *Personality Assessment Inventory: A professional manual*, Odessa (Floride), Psychological Assessment Resources, 1991.
- MOREY, L. C. et B. D. QUIGLEY. « The use of the Personality Assessment Inventory (PAI) in assessing offenders », *International Journal of Offender Therapy and Comparative Criminology*, vol. 46, p. 333-349, 2002.
- \*MOTIUK, L. L. *Antecedents and consequences of prison adjustment: A systematic assessment and reassessment approach*, thèse de doctorat non publiée, Université Carleton, Ottawa (Ontario), Canada, 1991.
- \*MOTIUK, M. S., L. L. MOTIUK et J. BONTA. « A comparison between self-report and interview-based inventories in offender classification », *Criminal Justice and Behavior*, vol. 19, p. 143-159, 1992.
- \*MÜLLER-ISBERNER, R., J. SOMMER, K. ÖZOKYAY et R. FREESE. *Clinical use of the HCR-20 for predicting violence in a German Forensic Psychiatric Hospital*, document présenté à l'International Conference on Risk Assessment and Management: Implications for Prevention of Violence, Vancouver (Colombie-Britannique), Canada, novembre 1999.
- \*NICHOLLS, T. *Violence risk assessments with female NCRMD acquittees: Validity of the HCR-20 and PCL-SV*, thèse de doctorat non publiée, Université Simon Fraser, Colombie-Britannique, Canada, 2001.

- \*NICHOLLS, T. L., G. M. VINCENT et K. E. WHITTEMORE. *Assessing risk of impatient violence in a sample of forensic psychiatric patients: Comparing the PCL:SV, HCR-20, and VRAG*, document présenté à l'International Conference on Risk Assessment and Management: Implications for Prevention of Violence, Vancouver (Colombie-Britannique), Canada, novembre 1999.
- NUFFIELD, J. *La libération conditionnelle au Canada : recherches en vue d'une normalisation des décisions*, Ottawa (Ontario), Ministère des Approvisionnement et Services Canada, 1982.
- \*NUGENT, P. M. *The use of detention legislation: Factors affecting detention decisions and recidivism among high-risk federal offenders in Ontario*, thèse de doctorat non publiée, Université Queen's, Kingston (Ontario), Canada, 2000.
- OSBERG, T. M. et D. L. POLAND. « Validity of the MMPI-2 basic and Harris-Lingoes subscales in a forensic sample », *Journal of Clinical Psychology*, vol. 57, p. 1369-1380, 2001.
- PHILLIPS, H. K., N. S. GRAY, S. I. MACCULLOCH, J. TAYLOR, S. C. MOORE, P. HUCKLE et M. J. MACCULLOCH. « Risk assessment in offenders with mental disorders: Relative efficacy of personal demographic, criminal history, and clinical variables », *Journal of Interpersonal Violence*, vol. 20, p. 833-847, 2005.
- PIQUERO, A. « Frequency, specialization, and violence in offending careers », *Journal of Research in Crime and Delinquency*, vol. 37, p. 392-418, 2000.
- \*POLVI, N. H. *The prediction of violence in pre-trial forensic patients: The relative efficacy of statistical versus clinical predictions of dangerousness*, thèse de doctorat non publiée, Université Simon Fraser, Colombie-Britannique, Canada, 2001.
- QUAY, H. *Managing adult inmates: Classification for housing and program assignments. Adult Internal Management System (AIMS) Classification manual*, College Park (Maryland), American Correctional Association, 1984.
- QUINSEY, V. L., G. T. HARRIS, M. E. RICE et C. A. CORMIER. *Violent offenders: Appraising and Managing Risk*, 1<sup>ère</sup> édition, Washington (D.C.), American Psychological Association, 1998.
- \*RAYNOR, P. *Implementing the Level of Service Inventory-Revised (LSI-R) in Britain: Report 3, risk and need assessment in the five pilot areas*, Centre for Criminal Justice and Criminology, Department of Social Policy and Applied Social Studies, University of Wales, Swansea (Royaume-Uni), de concert avec la Cognitive Centre Foundation, Dinas Powys Vale of Glamorgan (R.-U.), 1998.
- \*RETTINGER, L. J. *A recidivism follow-up study investigating risk and need within a sample of provincially sentenced women*, thèse de doctorat non publiée, Université Carleton, Ottawa (Ontario), Canada, 1998.
- RICE, M. E. « Violent offender research and implications for the criminal justice system », *American Psychologist*, vol. 52, p. 414-423, 1997.

- \*RICE, M. E. et G. T. HARRIS. « A comparison of criminal recidivism among schizophrenic and nonschizophrenic offenders », *International Journal of Law and Psychiatry*, vol. 15, p. 397-408, 1992.
- \*RICE, M. E. et G. T. HARRIS. « Psychopathy, schizophrenia, alcohol abuse, and violent recidivism », *International Journal of Law and Psychiatry*, vol. 18, p. 333-342, 1995.
- ROGERS, R. « The uncritical acceptance of risk assessment in forensic practice », *Law and Human Behavior*, vol. 24, p. 595-605, 2000.
- ROSENTHAL, R. *Meta-analytic procedures for social research*, Beverley Hills (Californie), Sage, 1991.
- \*ROWE, R. *The utilization of an interview-based classification instrument for parole board decision making in Ontario*, thèse de doctorat non publiée, Université Carleton, Ottawa (Ontario), Canada, 1995.
- \*ROWE, R. *Applying risk-needs assessments to parole decisions*, manuscrit en cours de rédaction, 1997.
- SALEKIN, R. T., R. ROGERS et K. W. SEWELL. « A review and meta-analysis of the Psychopathy Checklist - Revised: Predictive validity of dangerousness », *Clinical Psychology: Science and Practice*, vol. 3, p. 203-215, 1996.
- SCHMIDT, F. L. « Statistical significance testing and cumulative knowledge in psychology: Implications for training researchers », *Psychological Methods*, vol. 1, p. 115-129, 1996.
- \*SERIN, R. T. Violent recidivism in criminal psychopaths, *Law and Human Behavior*, vol. 20, p. 207-217, 1996.
- \*SERIN, R. C. et S. L. BROWN. *Empirical guidelines for clinicians: Enhancing the practice of offender risk assessment*, Ottawa (Ontario), Canada, Direction de la recherche, Service correctionnel du Canada, 1998.
- SIMOURD, D. J. « The Criminal Sentiments Scale-Modified and Pride in Delinquency scale: Psychometric properties and construct validity of two measures of criminal attitude », *Criminal Justice and Behavior*, vol. 24, p. 52-70, 1997.
- \*SIMOURD, D. J. « Use of a dynamic risk/need assessment instruments among long-term incarcerated offenders », *Criminal Justice and Behavior*, vol. 31, p. 306-323, 2004.
- \*SIMOURD, D. J. et J. VAN DE VEN. « Assessment of criminal attitudes: Criterion-related validity of the Criminal Sentiments Scale-Modified and the Pride in Delinquency Scale », *Criminal Justice and Behavior*, vol. 26, p. 90-106, 1999.
- SMITH, P., P. GENDREAU et C. G. GOGGIN (sous presse). « “What works” in predicting psychiatric hospitalization and relapse: The specific responsivity dimension of effective correctional treatment for mentally disordered offenders », sous la direction de R. Ax et T. Faggan, *Corrections, mental health, and social policy: International perspectives*, Springfield (Illinois), Charles C. Thomas & Ltd.

- \*ST. AMAND, M. D. *The coping relapse model of recidivism and the prediction of failure after relapse*, thèse de doctorat non publiée, Université Queen's, Kingston (Ontario), Canada, 2003.
- \*STADTLAND, C., N. KLEINDIENST, C. KRÖNER, M. EIDT et N. NEDOPIL. « Psychopathic traits and risk of criminal recidivism in offenders with and without mental disorders », *International Journal of Forensic Mental Health*, vol. 4, p. 89-97, 2005.
- \*STRAND, S., H. BELFRAGE, G. FRANSSON et S. LAVENDER. « Clinical and risk management factors in risk prediction of mentally disordered offenders – more important than historical data? A retrospective study of 40 mentally disordered offenders assessed with the HCR-20 Violence Risk Assessment Scheme », *Legal and Criminological Psychology*, vol. 4, p. 67-76, 1999.
- \*STRIBLING, A. L. *Exploring differences in clinical judgments of dangerousness of North Carolina psychiatric inpatients: Utility of the HCR-20 risk assessment scheme*, thèse de doctorat non publiée, University of Tennessee, 2003.
- SWETS, J. A. « Indices of discrimination and diagnostic accuracy: Their ROCs and implied models », *Psychological Bulletin*, vol. 99, p. 100-117, 1986.
- \*TENGSTRÖM, A. « Long-term predictive validity of historical factors in two risk assessment instruments in a group of violent offenders with schizophrenia », *Nordic Journal of Psychiatry*, vol. 55, p. 243-249, 2001.
- \*TENGSTRÖM, A., M. GRANN, N. LÅNGSTRÖM et G. KULLGREN. « Psychopathy (PCL-R) as a predictor of violent recidivism among criminal offenders with schizophrenia », *Law and Human Behavior*, vol. 24, p. 45-58, 2000.
- \*URBANIOK, F., T. NOLL, S. GRUNEWALD, J. STEINBACH et J. ENDRASS. « Prediction of violent and sexual offences: A replication study of the VRAG in Switzerland », *Journal of Forensic Psychiatry and Psychology*, vol. 17, p. 23-31, 2006.
- \*VILLENEUVE, D. B., N. OLIVER et W. LOZA. « Cross-validation of the Self Appraisal Questionnaire with a maximum-security psychiatric population », *Journal of Interpersonal Violence*, vol. 18, p. 1325-1334, 2003.
- \*VILLENEUVE, D. B. et V. L. QUINSEY. « Predictors of general and violent recidivism among mentally disordered inmates », *Criminal Justice and Behavior*, vol. 22, p. 397-410, 1995.
- WALTERS, G. D. « The Psychological Inventory of Criminal Thinking Styles (PICTS). Part I: Reliability and preliminary validity », *Criminal Justice and Behavior*, vol. 22, p. 307-325, 1995.
- WALTERS, G. D. « The Psychological Inventory of Criminal Thinking Styles (PICTS). Part III: Predictive validity », *International Journal of Offender Therapy and Comparative Criminology*, vol. 40, p. 105-112, 1996.
- WALTERS, G. D. « Risk-appraisal versus self-report in the prediction of criminal justice outcomes », *Criminal Justice and Behavior*, vol. 33, p. 179-304, 2006.

- \*WALTERS, G. D., S. A. DUNCAN et M. D. GEYER. « Predicting disciplinary adjustment in inmates undergoing forensic evaluation: A direct comparison of the PCL-R and PAI », *Journal of Forensic Psychiatry and Psychology*, vol. 14, p. 382-393, 2003.
- \*WALTERS, G. D. et W. MANDELL. « Incremental validity of the Psychological Inventory of Criminal Thinking and Psychopathy Checklist: Screening Version in predicting disciplinary outcome », *Law and Human Behavior*, sous presse.
- WALTERS, G. D., T. W. WHITE et D. DENNEY. « The Lifestyle Criminality Screening Form: Preliminary data », *Criminal Justice and Behavior*, vol. 18, p. 406-418, 1991.
- WEBSTER, C. D., K. S. DOUGLAS, D. EAVES et S. D. HART. « Assessing risk of violence to others », sous la direction de C. D. Webster et M. A. Jackson, *Impulsivity: Theory, assessment, and treatment*, p. 251-277, New York, Guilford, 1997a.
- WEBSTER, C. D., K. DOUGLAS, D. EAVES et D. HART. *HCR-20: Assessing risk for violence (Version 2)*, Vancouver (Colombie-Britannique), Canada, Mental Health, Law, & Policy Institute, Université Simon Fraser, 1997b.
- WEBSTER, C. D., G. T. HARRIS, M. E. RICE, C. CORMIER et V. L. QUINSEY. *The Violence Prediction Scheme: Assessing dangerousness in high risk men*, Toronto (Ontario), Canada, Université de Toronto, Centre of Criminology, 1994.
- \*WINTRUP, A. M. *Assessing risk of violence in mentally disordered offenders with the HCR-20*, thèse de doctorat non publiée, Université Simon Fraser, Burnaby (Colombie-Britannique), Canada, 1996.
- \*WONG, S. C. P. et A. GORDON. « The Violence Risk Scale ». *Forensic Update*, vol. 67, p. 19-23, 2001.
- \*WONG, S. C. P. et A. GORDON. The validity and reliability of the Violence Risk Scale: A treatment friendly violence risk assessment tool, *Psychology, Public Policy, and the Law*, vol. 12, p. 279-309, 2006.
- \*WORMITH, J. S., M. E. OLVER, H. E. STEVENSON et L. GIRARD. *The long term prediction of offender recidivism using diagnostic, personality, and risk/need approaches to offender assessment*, manuscrit en cours de rédaction, 2006.
- YEATON, W.H. et P. M. WORTMAN. « On the reliability of meta-analytic reviews: The role of intercoder agreement », *Evaluation Review*, vol. 17, p. 292-309, 1993.



## Annexe A

### Les études figurant dans la méta-analyse et leurs estimations de la valeur de l'effet respectives pour la prédiction de la violence en établissement (iv) et de la récidive (récid)

No.	Étude	Mesure	N <sub>iv</sub>	N <sub>récid</sub>	R <sub>iv</sub>	r <sub>récid</sub>
1	Belfrage et coll. (2000)	Échelle des variables historiques, cliniques et de gestion du risque-20	41	-	0,44	-
		Échelle de psychopathie : version de dépistage	41	-	0,33	-
2	Blanchette (2005)	Échelle de réévaluation du niveau de sécurité pour les délinquants	400	-	0,27	-
		Cote de sécurité des délinquants	400	-	0,24	-
3	Blanchette et coll. (2002)	ECNS : adaptation à la vie en établissement	61	-	0,39	-
		ECNS : adaptation à la vie en établissement	230	-	0,19	-
		Échelle de classement par niveau de sécurité : risque pour la sécurité	61	-	0,01	-
		Échelle de classement par niveau de sécurité : risque pour la sécurité	230	-	0,18	-
4	Bonta (1989)	INS/INS-R	49	-	0,16	-
		INS/INS-R	71	-	0,25	-
5	Bonta et coll. (1996)	Échelle d'ISR	-	3267	-	0,15
6	Bonta et Yessine (2005)	Échelle d'ISR	-	159	-	0,43
		VRAG	-	48	-	0,31
		VRAG-approximation	-	207	-	0,33
		INS:SV	-	235	-	0,27
		STATIQUE-99	-	154	-	0,14
7	Bonta et Motiuk (1986)	INS/INS-R	119	-	0,36	-
8	Collie et Polaschek (2003)	Outil de classement par niveau de sécurité de l'Australie	889	-	0,25	-
9	Cooke (1996)	PBRS: Anti-authority	220	-	0,15	-
		PBRS: Anti-authority	220	-	0,08	-
		PBRS: Anti-authority	220	-	0,10	-
10	Cunningham et Sorensen (2006)	PBRS: Anxious-depressed	13341	-	0,22	-
11	Cunningham et coll. (2005)	PBRS: Anxious-depressed	2505	-	0,24	-
12	Daffern et coll. (2005)	INS: Version de dépistage	232	-	0,15	-
13	Dahle (2006)	INS/INS-R	-	307	-	0,23
		Échelle des variables historiques, cliniques et de gestion du risque-20	-	307	-	0,31
		PCL/PCL-R	-	307	-	0,32
14	Dernevik et coll. (2002)	Échelle des variables historiques, cliniques et de gestion du risque-20	54	-	0,46	-
		Échelle de psychopathie : version de dépistage	54	-	0,17	-
15	Dolan et Khawaja (2004)	Échelle des variables historiques, cliniques et de gestion du risque-20	-	70	-	0,44
16	Douglas et coll. (2003)	Échelle des variables historiques, cliniques et de gestion du risque-20	-	100	-	0,26
17	Doyle et coll. (2002)	Échelle de psychopathie : Version de dépistage	87	-	0,38	-
		VRAG avec l'échelle PCL-VD	87	-	0,22	-
		Échelle des variables historiques-10	87	-	0,25	-

No.	Étude	Mesure	N <sub>iv</sub>	N <sub>récid</sub>	R <sub>iv</sub>	r <sub>récid</sub>
18	Fujji et coll. (2005)	Échelle des variables historiques, cliniques et de gestion du risque-20	41	-	0,09	-
		Échelle des variables historiques, cliniques et de gestion du risque-20	38	-	0,18	-
		Échelle des variables historiques, cliniques et de gestion du risque-20	29	-	0,35	-
19	Gagliardi et coll. (2004)	Divers éléments des antécédents criminels	-	333	-	0,42
20	Girard (1999)	INS/INS-R	-	630	-	0,28
		INS/INS-R	-	36	-	0,24
21	Glover et coll. (2002)	Échelle d'ISR	-	106	-	0,34
		VRAG	-	106	-	0,32
		VRAG-Child Adolescent Taxon	-	106	-	0,29
		Diagnostic de troubles	-	106	-	0,28
		Échelle d'ISR (violence)	-	106	-	0,27
		Formule de présélection pour l'évaluation psychologique	-	106	-	0,25
		Child Adolescent Taxon	-	106	-	0,18
		PCL/PCL-R	-	106	-	0,10
		DSM III/IV – Trouble de la personnalité antisociale	-	106	-	0,03
22	Grann et coll. (1999)	PCL/PCL-R	-	352	-	0,36
23	Gray et coll. (2003)	Échelle des variables cliniques-15	34	-	0,53	-
		PCL/PCL-R	34	-	0,35	-
		Échelle de désespoir de Beck	34	-	0,18	-
		Brief Psychiatric Rating Scale	34	-	0,61	-
24	Gray et coll. (2004)	Échelle de psychopathie: version de dépistage	-	316	-	0,11
		Échelle des variables historiques, cliniques et de gestion du risque-20	-	316	-	0,08
		Échelle de récidive des groupes de délinquants	-	316	-	0,08
25	Grevatt et coll. (2004)	Échelle des variables historiques et cliniques-15	44	-	0,10	-
		Échelle d'évaluation du risque de violence	44	-	0,05	-
26	Hanson et Wallace-Capretta (2000)	INS/INS-R	-	275	-	0,32
27	Harer et Langen (2001)	Divers éléments des antécédents criminels	177767	-	0,18	-
		Divers éléments des antécédents criminels	24675	-	0,15	-
28	Harris et coll. (1991)	PCL/PCL-R	-	169	-	0,42
		INS/INS-R	-	169	-	0,24
29	Harris et coll. (2002)	VRAG	-	133	-	0,48
		Évaluation clinique	-	383	-	0,17
30	Harris et coll. (1993)	PCL/PCL-R	-	618	-	0,34
		INS/INS-R	-	618	-	0,23
		Diagnostic de schizophrénie	-	618	-	-0,17
		Diagnostic du trouble de la personnalité	-	618	-	0,26
		VRAG	-	618	-	0,42
		Divers éléments des antécédents criminels	-	618	-	0,36
31	Heilbrun et coll. (1998)	PCL/PCL-R	218	181	0,14	0,16
32	Hemphill (1991)	PCL/PCL-R	-	106	-	0,06
		Échelle d'ISR	-	106	-	0,00
		Échelle des facteurs prépondérants	-	106	-	0,10
33	Hemphill et coll. (1998)	PCL/PCL-R	-	274	-	0,20
34	Hildebrand et coll. (2004)	PCL/PCL-R	92	-	0,03	-



No.	Étude	Mesure	N <sub>iv</sub>	N <sub>récid</sub>	R <sub>iv</sub>	r <sub>récid</sub>
35	Holland et coll. (1983)	Évaluation clinique	-	198	-	0,37
		Échelle des facteurs prépondérants	-	198	-	0,19
36	Jemelka et coll. (1992)	Diagnostic du trouble de la personnalité	500		0,65	-
37	Kroner et Loza (2001)	Questionnaire d'auto-évaluation	-	78	-	0,30
		PCL/PCL-R	-	78	-	0,21
		Échelle d'ISR	-	78	-	0,30
		VRAG	-	78	-	0,14
38	Kroner et Mills (2001)	PCL/PCL-R	97	87	0,14	0,12
		INS/INS-R	97	87	0,20	0,19
		Échelle des variables historiques, cliniques et de gestion du risque-20	97	87	0,11	0,16
		VRAG	97	87	0,26	0,11
		Lifestyle Criminality Screening Form	97	87	0,13	0,12
39	Law (2004)	Échelle d'intervention dans la collectivité	-	497	-	0,11
40	Lovell et coll. (2005)	INS/INS-R	-	100	-	0,21
		Divers éléments des antécédents criminels	-	100	-	0,37
41	Loza et Green (2003)	Échelle d'ISR	-	91	-	0,35
		INS/INS-R	-	91	-	0,22
		PCL/PCL-R	-	91	-	0,22
		Questionnaire d'auto-évaluation	-	91	-	0,30
		VRAG	-	91	-	0,19
42	Loza et Loza-Fanous (1999)	INS/INS-R	-	140	-	0,11
43	Loza et Loza-Fanous (2000)	Questionnaire d'auto-évaluation	-	153	-	0,32
44	Loza et Loza-Fanous (2001)	Échelle d'ISR	-	68	-	0,32
		INS/INS-R	-	68	-	0,23
		PCL/PCL-R	-	68	-	0,19
		Questionnaire d'auto-évaluation	-	68	-	0,32
		VRAG	-	68	-	0,22
45	Loza et Loza-Fanous (2002)	Questionnaire d'auto-évaluation	303	-	0,27	-
46	Loza et Loza-Fanous (2003)	Questionnaire d'auto-évaluation		305	-	0,34
47	Loza et coll. (2002)	VRAG		124	-	0,05
48	Luciani et coll. (1996)	Échelle de classement par niveau de sécurité : score total	2187	-	0,19	-
49	McHattie et coll. (1999)	Child Adolescent Taxon	-	42	-	0,06
		PCL/PCL-R	-	42	-	0,27
50	Mills et coll. (2005)	VRAG	-	209	-	0,26
51	Mills et Kroner (2006)	PCL/PCL-R	-	209	-	0,18
		INS/INS-R	-	209	-	0,26
		Échelle d'ISR	-	209	-	0,30
52	Mills et coll. (2004)	MCAA-Partie A	-	144	-	0,18
		MCAA-Partie B	-	144	-	0,30
		Échelle d'ISR	-	144	-	0,38
53	Mills et coll. (2003)	Questionnaire d'auto-évaluation	-	77	-	0,37
		Échelle d'ISR	-	77	-	0,29
		QRS –Gestion des impressions	-	77	-	0,38
		QRS–Illusion sur soi-même	-	77	-	0,24

No.	Étude	Mesure	N <sub>iv</sub>	N <sub>récid</sub>	R <sub>iv</sub>	r <sub>récid</sub>
54	Motiuk (1991)	MMPI/MMPI-II	215	215	0,14	0,02
		MMPI/MMPI-II	215	215	0,14	0,06
		MMPI/MMPI II	215	215	0,04	0,06
		Évaluation du risque du client selon le système de classement des cas du Wisconsin Illinois Initial Risk Scale	215	215	0,19	0,08
		LS:SRI	99	-	0,28	-
55	Motiuk et coll. (1992)	INS/INS-R	99	-	0,32	-
		Échelle des variables historiques, cliniques et de gestion du risque-20	220	-	0,23	-
56	Müller-Isberner et coll. (1999)	Échelle des variables historiques, cliniques et de gestion du risque-20	31	39	0,26	0,26
		Échelle des variables historiques, cliniques et de gestion du risque-20	30	39	0,43	0,15
		Échelle de psychopathie : version de dépistage	31	39	0,27	0,32
		Échelle de psychopathie : version de dépistage	30	39	0,32	0,37
		Échelle des variables historiques, cliniques et de gestion du risque-20	125	-	0,31	-
58	Nicholls et coll. (1999)	VRAG	125	-	0,08	-
		Échelle de psychopathie : version de dépistage	125	-	0,14	-
		Échelle des variables historiques, cliniques et de gestion du risque-20	-	120	-	0,18
59	Nugent (2000)	Échelle d'ISR	-	120	-	0,23
		INS/INS-R	-	120	-	0,21
		VRAG	-	120	-	0,23
		PCL/PCL-R	-	120	-	0,17
		Child Adolescent Taxon	-	120	-	0,09
		Divers éléments des antécédents criminels	800	-	0,35	-
60	Jayjohn et Van Dine (2002)	Divers éléments des antécédents criminels	800	-	0,39	-
61	Halsall et Van Dine (2002)	Divers éléments des antécédents criminels	800	-	0,39	-
62	Polvi (1999)	Évaluation clinique	-	215	-	0,12
		Dangerous Behavior Rating Scale	-	215	-	0,14
		Échelle des variables historiques, cliniques et de gestion du risque-20	-	215	-	0,09
		Échelle de psychopathie : version de dépistage	-	215	-	0,25
		VRAG	-	215	-	0,31
		Échelle d'ISR	-	157	-	0,10
		PCL/PCL-R	-	164	-	0,23
63	St. Amand (2002)	Divers éléments des antécédents criminels	-	159	-	0,09
		Divers éléments des antécédents criminels	-	159	-	0,27
		Divers éléments des antécédents criminels	-	159	-	0,14
		Questionnaire révisé sur la résolution de problèmes sociaux	-	164	-	0,27
		Access to criminal resources items	-	121	-	0,16
		CSLQ	-	130	-	0,06
		Violent Beliefs Inventory	-	164	-	0,19
		Échelle d'ISR	-	234	-	0,16
		Child Adolescent Taxon	-	233	-	0,06
		PCL/PCL-R	-	233	-	0,05
		Divers éléments des antécédents criminels	-	234	-	0,00

No.	Étude	Mesure	N <sub>iv</sub>	N <sub>récid</sub>	R <sub>iv</sub>	r <sub>récid</sub>
		Divers éléments des antécédents criminels	-	234	-	0,17
		Divers éléments des antécédents criminels	-	234	-	0,20
		Problem Survey Checklist	-	234	-	0,01
		Questionnaire sur l'impulsivité	-	233	-	0,02
		Échelle de l'affect positif-affect négatif	-	224	-	0,02
		Échelle de l'affect positif-affect négatif	-	224	-	0,09
		Éléments relatifs à l'anxiété	-	229	-	0,09
		Perceived Problem Index	-	232	-	0,07
		Éléments relatifs à l'adaptation	-	231	-	0,02
		Éléments relatifs à l'adaptation	-	231	-	0,04
		Questionnaire révisé sur la résolution de problèmes sociaux	-	231	-	0,04
		Criminal Insensitivity/Irresponsibility Scale	-	224	-	0,10
		Access to criminal resources items	-	217	-	0,01
		CSLQ	-	227	-	0,00
		Social Support Scheme	-	231	-	0,09
		Questionnaire sur l'emploi du temps	-	220	-	0,07
		Violent Beliefs Inventory --	-	233	-	0,01
		QRS –Gestion des impressions	-	207	-	0,14
		QRS–Illusion sur soi-même deception	-	164	-	0,16
64	Serin (1996)	PCL/PCL-R	-	81	-	0,28
		Échelle d'ISR	-	81	-	0,08
		Échelle des facteurs prépondérants	-	81	-	0,15
		Base Expectancy Score	-	81	-	0,21
65	Serin et Brown (1998)	PCL/PCL-R	-	263	-	0,31
66	Simourd (2004)	INS/INS-R	-	129	-	0,26
67	Simourd et Van De Ven (1999)	ESC/ESC-M	-	87	-	0,17
		ESC/ESC-M	-	54	-	0,07
		Échelle de la fierté d'être délinquant	-	87	-	0,03
		Échelle de la fierté d'être délinquant	-	54	-	0,04
68	Stadtland et coll. (2005)	PCL/PCL-R	-	258	-	0,25
69	Strand et coll. (1999)	Échelle des variables historiques, cliniques et de gestion du risque-20	-	40	-	0,51
		Échelle de psychopathie : version de dépistage	-	40	-	0,35
70	Stribling (2003)	Échelle des variables historiques, cliniques et de gestion du risque-20	52	-	,51	-
71	Tengström (2001)	Échelle des variables historiques-10	-	106	-	0,41
		VRAG	-	106	-	0,29
72	Tengström et coll. (2000)	PCL/PCL-R	-	202	-	0,36
73	Urbaniok et coll. (2006)	VRAG	-	79	-	0,34
74	Villeneuve et coll. (2003)	Questionnaire d'auto-évaluation	-	49	-	0,28
		Questionnaire d'auto-évaluation	-	273	-	0,33
75	Villeneuve et Quinsey (1995)	Échelle d'ISR	-	117	-	0,25
		Violence Recidivism Scale (VRISK)	-	117	-	0,43
76	Walters et Mandell (in press)	Échelle de psychopathie: version de dépistage	136	-	0,14	-
		PICTS	136	-	0,19	-
		PICTS: échelle P	136	-	0,24	-
		PICTS: échelle R	136	-	0,16	-
77	Wintrup (1996)	PCL/PCL-R	-	70	-	0,20
		Échelle des variables historiques, cliniques et de gestion du risque-20	-	70	-	0,20

No.	Étude	Mesure	N <sub>iv</sub>	N <sub>récid</sub>	R <sub>iv</sub>	r <sub>récid</sub>
78	Wong et Gordon (2001)	Échelle de risque de violence (ERV)	-	2000	-	0,46
79	Wong et Gordon (2006)	Échelle de risque de violence (ERV)	-	918	-	0,40
80	Wormith et coll. (2006)	NS-IGC	-	60	-	0,31
		PCL/PCL-R	-	60	-	0,28
		DSM III/IV Trouble de la personnalité antisociale	-	60	-	0,35
81	Raynor (1998)	INS/INS-R	-	147	-	0,26
82	Rettinger (1998)	INS/INS-R	-	441	-	0,44
		NS-IGC	-	441	-	0,44
83	Rice et Harris (1992)	Diagnostic de schizophrénie	-	190	-	0,10
		INS/INS-R	-	618	-	0,11
		Diagnostic du trouble de la personnalité	-	618	-	0,11
84	Rice et Harris (1995)	PCL/PCL-R	-	190	-	0,27
		Éléments relatifs à la toxicomanie	-	190	-	0,10
85	Rowe (1995)	INS/INS-R	-	389	-	0,34
		Échelle d'ISR	-	389	-	0,24
		Échelle des facteurs prépondérants	-	289	-	0,14
		Évaluation clinique	-	262	-	0,15
		Évaluation clinique	-	145	-	0,11
		Évaluation clinique	-	112	-	0,11
		Évaluation clinique	-	146	-	0,13
86	Rowe (1997)	NS-IGC	-	340	-	0,30
87	Walters et coll. (2003)	PCL/PCL-R	185	-	0,11	-
		IEP – Sous-échelle sur l'agression	149	-	0,17	-
		IEP – Sous-échelle antisociale	149	-	0,12	-
88	Edens et Ruiz (2006)	IEP – Sous-échelle antisociale	349	-	0,15	-
		IEP – Gestion des impressions positives	349	-	0,02	-

Nota : QRS = Questionnaire sur les réactions souhaitables; ECNS = Échelle de classement par niveau de sécurité; CSLQ = Criminal Socialisation and Lifestyle Questionnaire; ESC/ESC-M = échelle de sentiments criminels / échelle de sentiments criminels - modifiée; NS-IGC = niveau de service-inventaire de gestion des cas; INS/INS-R = inventaire du niveau de service ou inventaire du niveau de service – révisé; MCAA = Measures of Criminal Attitudes and Associates; IEP = Inventaire d'évaluation de la personnalité; PBRS = Prison Behavior Rating Scale; PICTS = Psychological Inventory of Criminal Thinking Styles; PCL/PCL-R = Échelle de psychopathie / Échelle de psychopathie – révisée; échelle d'ISR = Échelle d'information statistique sur la récidive; VRAG = guide d'évaluation du risque de violence (*Violence Risk Appraisal Guide*).