

**Lignes directrices sur l'évaluation des
recherches sur l'efficacité des programmes
de traitement pour délinquants sexuels du
Comité de collaboration sur les données
collectives relatives aux résultats**

Partie 2 : Lignes directrices du CCDCR

2007-03

Membres du Comité de collaboration sur les données collectives
relatives aux résultats (par ordre alphabétique) :

Anthony Beech, Guy Bourgon, R. Karl Hanson, Andrew J. R. Harris,
Calvin Langton, Janice Marques, Michael Miner, William Murphy,
Vernon Quinsey, Michael Seto, David Thornton, Pamela M. Yates

N° de cat. : PS3-1/2007-3F
ISBN : 978-0-662-09390-9

Table des matières

Introduction	1
Définition du modèle	3
I. Contrôle administratif des variables indépendantes	4
1. Définition du traitement	6
2. Définition de la comparaison	10
3. Divers facteurs secondaires	14
II. Attentes de l'expérimentateur	17
4. Intérêt de l'expérimentateur dans le résultat	18
5. Évaluation à l'aveugle de la gestion des données	21
III. Taille de l'échantillon	24
6. Taille de l'échantillon du ou des groupes expérimentaux	26
7. Taille de l'échantillon du ou des groupes témoins	29
8. Taille de l'échantillon d'établissements : modèles fondés sur les établissements	32
IV. Attrition	35
9. Sélection des sujets	38
10. Abandon du programme	42
11. Analyse des sujets retenus au début de l'étude	45
12. Attribution pendant le suivi	48
V. Équivalence des groupes	51
13. Équivalence des groupes a priori	53
14. Qualité de la recherche de différences préexistantes	60
15. Constatations sur l'équivalence des groupes	63
VI. Variables des résultats	67
16. Durée du suivi	69
17. Validité et fiabilité des renseignements sur la récidive	72
18. Équivalence du suivi	75
VII. Comparaisons correctes établies	78
19. Recherche d'une association à tout prix	80
20. Efficacité des procédures statistiques de contrôle des biais	83
21. Évaluation de la comparaison qui introduit le moins de biais	87
VIII. Évaluation globale	91
Feuille sommaire du guide d'évaluation de la qualité de l'étude	92
Bibliographie	95
Membres du Comité	97

Introduction

Les lignes directrices suivantes du Comité de collaboration sur les données collectives relatives aux résultats (CCDCR) servent à l'évaluation de la qualité des études sur les résultats du traitement des délinquants sexuels. Dans le cas d'une étude de qualité supérieure, on peut être à peu près certain que l'estimation de l'effet du traitement comporte un biais minimum. Les présentes lignes directrices devraient être utiles pour les lecteurs et les examinateurs de documents spécialisés ainsi que les chercheurs qui réalisent de nouvelles études ou évaluent des programmes en vigueur. Des renseignements de base figurent dans *Lignes directrices sur l'évaluation des recherches sur l'efficacité des programmes de traitement pour délinquants sexuels du Comité de collaboration sur les données collectives relatives aux résultats (Lignes directrices du CCDCR) Partie 1 : Introduction et aperçu*.

Directives concernant l'utilisation des lignes directrices du CCDCR avec les études actuelles

Première étape – Définir le genre de modèle. Au moyen de l'organigramme fourni, définir la méthode d'affectation des sujets qui ressemble le plus à l'étude examinée. Ce classement initial est important parce que certaines évaluations ne s'appliquent qu'à des modèles précis.

Deuxième étape – Définir la variable des résultats. Le choix de la variable des résultats a des conséquences pour l'évaluation de la qualité de l'étude. Les lignes directrices qui suivent sont fondées sur la récidive sexuelle comme variable des résultats – un événement dont le taux de base est faible. Si la violence générale ou toute récidive criminelle sert de critère des résultats, il est possible d'obtenir un coefficient d'efficacité statistique équivalent assorti de périodes de suivi plus courtes et des tailles d'échantillons plus petites parce que le taux de base sera plus élevé. De plus, les évaluateurs voudront utiliser des variables de contrôle différentes selon le résultat visé. Les prédicteurs de la violence générale et de toute récidive criminelle ne sont pas identiques aux prédicteurs de la récidive sexuelle.

Troisième étape – Évaluer chacun des éléments. Les évaluations portent sur la confiance (p. ex. la taille de l'échantillon) ou le biais (p. ex. l'équivalence du suivi). Les évaluations de la confiance et du biais sont cotées sur des échelles de trois points (confiance faible, moyenne ou élevée; biais considérable, moyen ou négligeable).

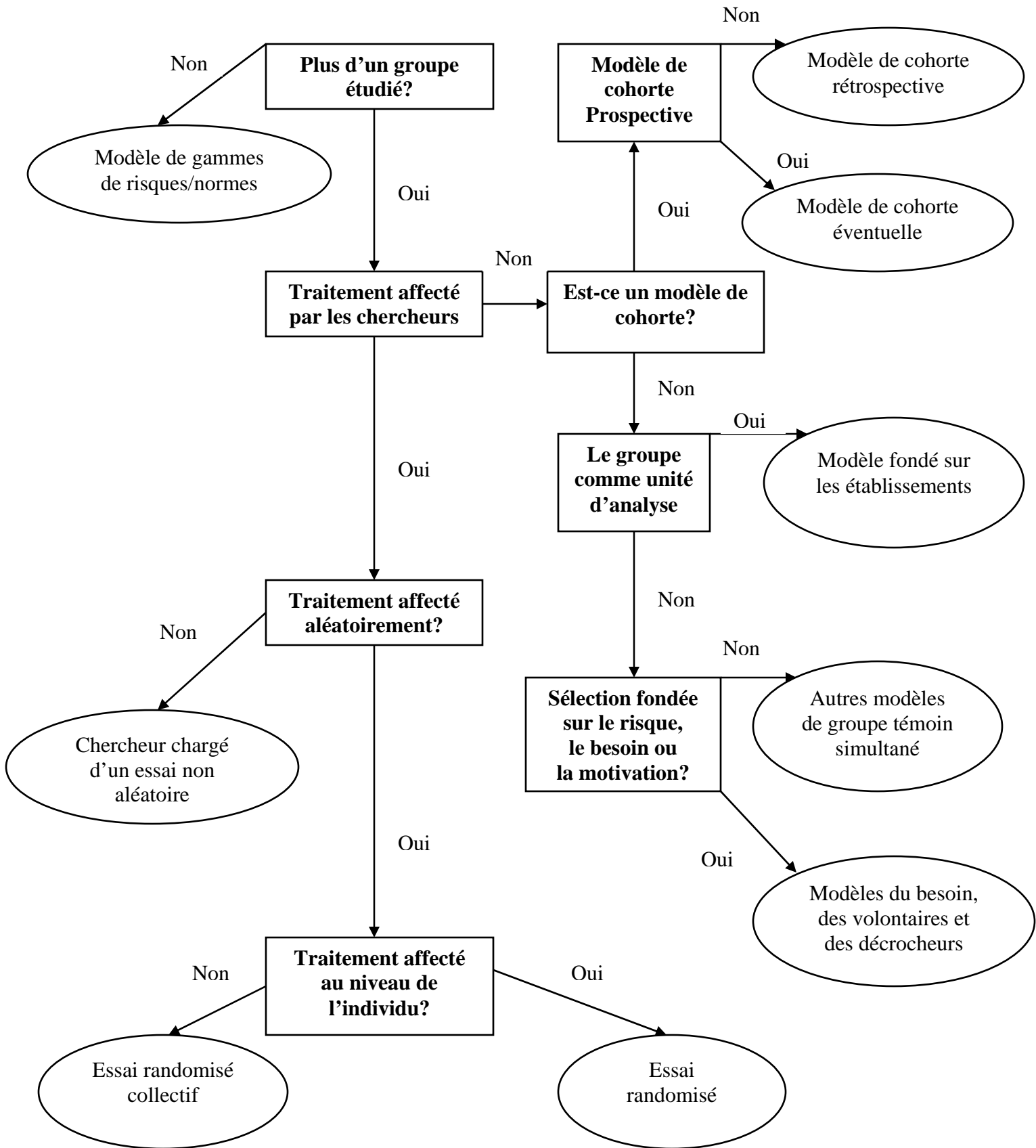
L'élément doit d'abord être évalué d'après les renseignements présentés dans l'article. À la page suivante, l'élément peut être évalué une deuxième fois selon de nouvelles données obtenues à partir d'autres sources ou au moyen de nouvelles analyses. Nous encourageons les évaluateurs à obtenir ces renseignements lorsque ceux-ci amélioreraient la qualité des évaluations ou la qualité de l'étude. Des exemples de nouveaux renseignements seraient des manuels de traitement ou la correspondance échangée avec les auteurs de l'étude indiquant que la personne qui analyse les données ne connaît pas le résultat. Des exemples de nouvelles analyses seraient la réorganisation des données présentées dans une analyse des sujets retenus au début de l'étude ou la réalisation de nouvelles analyses effectuées à partir des données brutes originales.

Quatrième étape – Procéder à une évaluation de la qualité générale. L'évaluation de la qualité finale est fondée sur les évaluations générales du « biais » et de la « confiance ». Les examinateurs utilisent leur jugement professionnel pour procéder à l'évaluation générale du « biais » et de la « confiance ». Même si l'on s'attendait à ce que les évaluations générales soient proportionnelles aux « scores totaux », nous ne recommandons pas d'utiliser un algorithme numérique particulier pour porter un jugement général sur le « biais » et la « confiance ». Après que le biais et la confiance ont été évalués, toutefois, des critères explicites sont fournis pour traduire ces évaluations sur une échelle de quatre points de la qualité de l'étude : rejetée, faible, bonne et excellente.

Les Lignes directrices du CCDCR comprennent 21 éléments, trois évaluations sommaires de la confiance, de la quantité de biais et du sens des biais et une évaluation générale de la qualité. Un élément (n° 8) s'applique aux modèles interétablissements et n'est évalué que pour ce genre d'étude. Les critères d'évaluation de certains des autres éléments varient selon le mode de sélection des groupes expérimentaux et témoins. Le tableau ci-dessous indique les modèles, les éléments et les facteurs propres au modèle qui servent à évaluer ce facteur.

Modèle	Élément
Modèles fondés sur les établissements	8. <u>Taille de l'échantillon des établissements : Modèles fondés sur les établissements</u> <i>Cet élément est coté seulement pour les modèles fondés sur les établissements (voir page 32).</i>
Modèles d'affectation aléatoire	6. <u>Taille de l'échantillon du ou des groupes expérimentaux</u> <i>La taille de l'échantillon requise est plus petite dans le cas des modèles d'affectation aléatoire que pour les autres modèles (voir page 26).</i> 7. <u>Taille de l'échantillon du ou des groupes témoins</u> <i>La taille de l'échantillon requise est plus petite dans le cas des modèles d'affectation aléatoire que pour les autres modèles (voir page 29).</i> 13. <u>Équivalence des groupes a priori</u> <i>Il faut tenir compte de facteurs particuliers des procédures d'affectation pour évaluer cet élément (voir page 55).</i>
Modèles de gammes de risques/normes	13. <u>Équivalence des groupes a priori</u> <i>Il faut tenir compte de facteurs particuliers de la validité des normes ou des gammes de risques pour évaluer cet élément (voir page 56).</i>
Modèles de cohorte	13. <u>Équivalence des groupes a priori</u> <i>Il faut tenir compte de facteurs particuliers des effets possibles de la cohorte pour évaluer cet élément (voir page 57).</i>

Définition du modèle



I. Contrôle administratif des variables indépendantes

Le contrôle administratif des variables indépendantes se rapporte au contenu et à l'intégrité des conditions de traitement et de référence. Sans connaissance du traitement donné, il est impossible de savoir ce qui a été évalué. L'évaluation d'interventions vagues et non exprimées peut intéresser les administrateurs qui se préoccupent uniquement de l'efficacité d'un traitement particulier dans un endroit particulier. Les chercheurs ont souvent des projets plus ambitieux, cependant, et souhaitent appliquer les résultats à d'autres cadres et échantillons. Pour transposer les résultats à d'autres endroits, il faut avoir une définition claire du traitement donné. À cette fin, il faut normalement des manuels de traitement et des descriptions claires de ce qui est arrivé au groupe témoin.

En plus de définir clairement le contenu du traitement, il importe de déterminer si celui-ci a été appliqué comme prévu (intégrité du traitement). On peut y arriver en formant les thérapeutes et par des « vérifications de la manipulation » qui permettent de surveiller la prestation réelle du traitement (p. ex. observer ce qui se passe pendant des séances de traitement réelles). Dans les études prospectives, les chercheurs doivent prendre des décisions concernant le degré prévu de « rigidité » des protocoles de traitement. Une étude peut exiger que les thérapeutes appliquent de manière rigide un manuel ou elle peut permettre de s'en écarter selon les besoins, le niveau de risque ou les progrès des participants. Si de nombreux ajustements sont permis, il devient difficile de décrire clairement le programme de traitement. D'autre part, les situations cliniques réelles peuvent exiger une certaine souplesse (p. ex. un programme de traitement communautaire peut obliger à intensifier le traitement si le niveau de risque d'un individu augmente et met les victimes éventuelles en danger). Même s'il peut s'agir d'une bonne pratique clinique, cette souplesse peut entraîner des écarts importants par rapport au protocole de traitement ou même un traitement croisé. Par exemple, si les participants du groupe témoin sont plus susceptibles de poser un risque accru et, par conséquent, ont besoin d'interventions pour des raisons de sécurité, le modèle original consistant à comparer un groupe expérimental à un groupe non traité est affaibli. Les chercheurs doivent établir à priori les critères qui justifient tout écart par rapport au protocole de traitement.

Les études rétrospectives présentent des défis de taille lorsqu'on tente de comprendre les interventions utilisées. Les documents disponibles sont souvent rares. Si les thérapeutes et les clients pouvaient être interrogés, il est peu probable que leurs réponses donneraient les renseignements détaillés souhaités par les chercheurs. Même si des documents exhaustifs sont disponibles, la version officielle d'un programme de traitement peut différer de ce qui s'est réellement passé.

Les questions d'intégrité s'appliquent également au(x) groupe(s) témoin(s). Afin d'évaluer les différences éventuelles entre les groupes expérimentaux et les groupes témoins, il faut connaître les politiques, les procédures et les expériences du groupe témoin. Il arrive souvent que les délinquants sexuels aient accès à certains services ou qu'ils fassent l'objet de dispositions légales et(ou) de pratiques de surveillance spéciales même s'ils n'ont pas bénéficié du traitement évalué.

Il se peut que les délinquants non traités fassent l'objet d'une attention particulière parce qu'ils n'ont pas été traités. Pour procéder à une évaluation critique d'une étude, il importe de connaître ces conditions spéciales et de savoir comment elles peuvent influencer sur la récidive du ou des groupes témoins.

Les Lignes directrices du CCDRC contiennent trois éléments servant à évaluer le contrôle administratif des variables indépendantes. Le premier élément, Définition du traitement, permet d'évaluer dans quelle mesure on est sûr que les délinquants ayant subi le traitement ont réellement reçu le traitement décrit et si le traitement peut être répété. Les renseignements pertinents comprennent des données détaillées comme la durée du traitement, le nombre de séances, le nombre d'heures de traitement, l'orientation théorique, le contenu du programme de traitement (p. ex. manuel de traitement), le lieu du traitement et les interventions et(ou) mesures utilisées pour assurer l'intégrité du traitement. Les évaluations de cet élément traduisent la confiance du codeur, d'après les données disponibles, dans le fait que le traitement a été donné selon la description fournie, avec intégrité, et pourrait être répété par d'autres.

Le deuxième élément, Définition de la comparaison, sert à évaluer la mesure dans laquelle on a confiance que les délinquants du groupe témoin ont été exposés aux conditions décrites (et non à quelque chose d'autre) et que ces conditions peuvent être reproduites. Les renseignements pertinents comprennent les services, les méthodes de surveillance, les facteurs spéciaux et la possibilité de recevoir d'autres formes de traitement. L'évaluation de cet élément traduit la mesure dans laquelle le codeur a confiance que le groupe témoin a été exposé aux conditions décrites et non à des pratiques ou à des services spéciaux.

Le troisième élément, Divers facteurs secondaires, sert à évaluer le biais qui peut être introduit par l'exposition à des facteurs extérieurs qui ne se rapportent pas aux méthodes de comparaison ou au programme de traitement. Les renseignements pertinents comprennent l'emplacement physique du traitement (p. ex. hôpital psychiatrique isolé en milieu fermé, prison/pénitencier) et les différences éventuelles dans les méthodes de surveillance entre les délinquants traités et les délinquants non traités. Les cotes pour cet élément traduisent l'évaluation par le codeur des biais possibles qui peuvent être introduits à cause de facteurs extérieurs qui pourraient modifier le résultat du ou des groupes.

1. Définition du traitement

Concept : Le concept général consiste à déterminer si le traitement pourrait ou non être répété. Le programme de traitement a-t-il été décrit de manière suffisamment détaillée pour permettre de le recréer? Dans quelle mesure êtes-vous sûr que le programme décrit était le programme offert aux participants? Il y a deux facteurs à prendre en considération : le contenu et l'intégrité. Y avait-il assez de renseignements pour donner le même traitement, y compris le contenu et l'intensité (c.-à-d. fréquence du contact), le dosage et le mode d'exécution? Quelles procédures ont été mises en place pour que les participants bénéficient réellement des interventions prévues? Plus il y a de données décrivant ce que les animateurs et les participants ont fait pendant le traitement, plus le niveau de confiance que le traitement a été donné comme prévu est élevé.

Indicateurs : Les indicateurs qui pourraient renforcer le degré de confiance que le traitement est reproductible et qu'il a été donné comme prévu comprendraient des éléments comme un manuel de traitement exhaustif, des renseignements précis sur la durée et la fréquence des séances, y compris la durée générale du traitement, la formation et la surveillance des animateurs et des vérifications de la manipulation prouvant que le traitement a été donné comme prévu (p. ex. le nombre de séances auxquelles les participants ont réellement assisté). Lorsque la plupart ou la totalité de ces indicateurs sont absents et que des énoncés généraux de l'orientation théorique (p. ex. approche cognitivo-comportementale) sont utilisés pour décrire le traitement, le niveau de confiance est peu élevé. Lorsqu'il y a certains indicateurs, comme un manuel, mais peu de formation ou aucune vérification de la manipulation, on a une certaine confiance que le traitement est reproductible et donné comme prévu.

Modèles fondés sur les établissements : Les modèles fondés sur les établissements sont uniques en ce sens qu'ils utilisent les établissements plutôt que les délinquants comme unité d'analyse. Dans ces modèles, le chercheur doit constituer un échantillon d'établissements et évaluer des facteurs théoriques particuliers aux établissements pour déterminer leur incidence sur la récidive. La sélection et la définition opérationnelle des facteurs particuliers aux établissements sont essentielles et doivent être des concepts bien définis ayant des liens théoriques avec la récidive.

Cote	Description
0	<i>Il n'y a pas de manuel de traitement; le contenu du traitement est décrit vaguement. Il peut y avoir des renseignements fournis qui indiquent la durée du programme ou la fréquence et la durée des séances, mais le nombre total d'heures du contact entre le thérapeute et le client n'est pas précisé. On ne sait pas si ou comment les animateurs se sont écartés du programme. La formation des animateurs est officieuse, et il n'y a à peu près pas de surveillance. Aucune vérification de la manipulation n'a été effectuée pour déterminer le respect du programme.</i>
1	<i>Le contenu et la structure du programme de traitement peuvent être déterminés au moyen de documents comme un calendrier du programme qui indique les sujets des séances, les documents à distribuer et(ou) les objectifs. Les renseignements indiquent le nombre total d'heures, la fréquence des séances et la durée du programme. Des procédures officielles de formation et de surveillance sont en vigueur. Les procédures de vérification de la manipulation sont officielles ou considérées comme faisant partie de la surveillance.</i>
2	<i>Un manuel de traitement exhaustif est disponible. Des procédures de formation officielles ont été utilisées, et les animateurs ont donné un traitement sous surveillance. Les vérifications de la manipulation ont permis de déterminer que le traitement a été exécuté selon la description prévue.</i>

1. Définition du traitement

Cet élément concerne le niveau de confiance que le traitement décrit correspondait au traitement que le groupe a réellement reçu. Les renseignements à prendre en considération sont la durée du traitement, le nombre de séances, le nombre d'heures de traitement, l'orientation théorique, le contenu du programme de traitement (p. ex. manuel de traitement) et les procédures ou mesures employées pour en assurer l'intégrité.

<i>Renseignements sur le traitement provenant de l'étude :</i>			
Cote de confiance			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Peu de confiance que le traitement a été donné selon la description</i>	<i>Certaine confiance que le traitement a été donné selon la description, mais pourrait être plus convaincant / certaines réserves</i>	<i>Grande confiance que le traitement a été donné selon la description</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Définition du traitement

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance que le traitement prescrit a été donné selon la description

1

Certaine confiance que le traitement prescrit a été donné selon la description, mais pourrait être plus convaincant / certaines réserves

2

Grande confiance que le traitement a été donné selon la description

Raison(s) de la cote :

2. Définition de la comparaison

Concept : Le concept général est la mesure dans laquelle les conditions de la comparaison pourraient être reproduites. Cet élément vise à déterminer la confiance du codeur que l'auteur ou les auteurs ont décrit avec suffisamment de détails les conditions de la comparaison pour recréer celles-ci et que ces conditions étaient en fait celles auxquelles le groupe témoin a été exposé. Tout comme pour l'élément précédent, il y a deux facteurs à prendre en considération. Un facteur est la qualité de la description des conditions de la comparaison. Cela comprend les services que les délinquants du groupe témoin devaient recevoir. Le deuxième facteur est l'intégrité des conditions de la comparaison : Ont-ils en fait reçu ces services et seulement ces services? Si le groupe témoin a reçu un autre traitement ou s'il a fait l'objet d'une méthode de surveillance différente, il faut évaluer les renseignements disponibles sur les interventions de rechange.

Indicateurs : Les indicateurs qui pourraient renforcer la confiance que les conditions de la comparaison sont reproductibles et correspondent à la description comprendraient les éléments suivants : descriptions claires des services que les participants ont reçus, politiques et procédures de surveillance et vérifications de la manipulation pour déterminer si ces participants n'ont pas bénéficié de services de traitement spéciaux ou différents. Lorsque la plupart ou la totalité de ces indicateurs sont absents ou qu'il n'y a qu'un énoncé général selon lequel le groupe témoin n'a pas reçu le traitement à l'étude, le niveau de confiance est faible. On a confiance que les conditions de la comparaison sont reproductibles lorsqu'il y a des indicateurs, comme une description de la nature du contenu des conditions de la comparaison ainsi que la vérification que les participants n'ont pas reçu un traitement ou des services de surveillance supplémentaires ou imprévus.

Modèles de gammes de risques/normes : Dans ces études, le groupe témoin constitue un échantillon normatif. Par conséquent, la cote des modèles de gammes de risques/normes seraient généralement « 2 – Grande confiance ». Les cotes pourraient être inférieures à 2, toutefois, si vous croyez que les normes ne sont pas fiables et que les chercheurs devraient utiliser d'autres échantillons normatifs s'ils tentent de reproduire l'étude.

Cote	Description
0	<i>Le groupe témoin est simplement décrit comme les délinquants qui n'ont pas reçu le traitement en question, et aucun autre renseignement n'est présenté concernant les conditions auxquelles ils ont été exposés.</i>
1	<i>Certaine description des conditions de la comparaison, notamment les méthodes et(ou) activités de surveillance pendant l'incarcération, mais aucun renseignement n'est présenté pour garantir que ces sujets n'ont pas reçu un traitement ou d'autres services différents non indiqués autrement.</i>
2	<i>Description claire et détaillée des conditions de la comparaison, y compris les méthodes de surveillance standards ou les activités types pendant l'incarcération et la preuve d'une vérification de la manipulation que les sujets du groupe témoin vérifiés n'ont pas reçu le traitement ou d'autres services différents non indiqués autrement.</i>

2. Définition de la comparaison

Cet élément concerne les procédures applicables au(x) groupe(s) témoin(s) et le niveau de confiance que les conditions de ce(s) groupe(s) sont reproductibles et conformes à la description. Les renseignements à prendre en considération comprennent les services et/ou la surveillance dont ils ont fait l'objet, y compris les interventions équivalentes ou semblables dont a bénéficié le groupe expérimental, et l'intégrité du groupe témoin (p. ex. ils n'ont pas reçu le traitement). Si le groupe témoin a reçu un autre traitement, il faut évaluer les renseignements disponibles sur cet autre traitement.

<i>Renseignements sur la comparaison provenant de l'étude :</i>			
Cote de confiance			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Peu de confiance que le groupe témoin a été exposé conformément à la description</i>	1 <i>Certaine confiance que le groupe témoin a reçu les services ou fait l'objet d'une surveillance conformément à la description, mais pourrait être plus convaincant / certaines réserves</i>	2 <i>Grande confiance que le groupe témoin a été exposé aux conditions décrites</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Définition de la comparaison

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0	1	2
<i>Peu de confiance que le groupe témoin a été exposé conformément à la description</i>	<i>Certaine confiance que le groupe témoin a reçu les services ou fait l'objet d'une surveillance conformément à la description, mais pourrait être plus convaincant / certaines réserves</i>	<i>Grande confiance que le groupe témoin a été exposé aux conditions décrites</i>

Raison(s) de la cote :

3. Divers facteurs secondaires

Concept : Le concept général consiste à déterminer s'il y a ou non divers facteurs secondaires qui peuvent introduire un biais dans les résultats par suite de l'exposition différente du groupe expérimental et du groupe témoin à des facteurs qui peuvent être liés à la récidive. Cet élément vise à évaluer l'ampleur et le sens du biais qui résultent de ces facteurs. Les facteurs secondaires différents importants peuvent être présents pendant la prestation du traitement ou les conditions de la comparaison (p. ex. le traitement a eu lieu dans un hôpital psychiatrique en milieu fermé éloigné tandis que les sujets du groupe témoin ont été incarcérés dans la population générale d'un établissement à sécurité maximale) ou après la prestation de services (p. ex. les délinquants traités ont été exposés à des méthodes de surveillance dans la collectivité différentes de celles applicables aux délinquants du groupe témoin).

Indicateurs : Les indicateurs qui pourraient accroître la possibilité d'un biais seraient les endroits très différents ou les méthodes de surveillance différentes pendant l'incarcération ou dans la collectivité, en particulier les méthodes qui peuvent influencer sur la récidive.

Modèles fondés sur les établissements : Les divers facteurs secondaires sont les variables dignes d'intérêt dans les modèles fondés sur les établissements, car ce sont ces variables qui sont évaluées. Il importe en particulier que le chercheur ait examiné et testé rigoureusement des hypothèses de rechange, car il est peu probable qu'il ait évalué tous les facteurs secondaires.

Cote	Description
0	<i>Le cadre et les services offerts au groupe expérimental et au groupe témoin diffèrent pour au moins un facteur qui devrait modifier les taux de récidive. Dans le cas des modèles fondés sur les établissements, il est impossible de dissocier ces facteurs secondaires des autres caractéristiques dignes d'intérêt.</i>
1	<i>Le groupe expérimental et le groupe témoin diffèrent pour au moins un facteur qui peut avoir un lien avec la récidive. L'effet de ce facteur secondaire est inconnu. Dans le cas des modèles fondés sur les établissements, il est impossible de dissocier complètement ce facteur secondaire des autres caractéristiques dignes d'intérêt. L'effet de ce facteur secondaire est inconnu.</i>
2	<i>Il n'y a pas de facteurs secondaires inhérents à la prestation de services ou pendant le suivi qui devraient modifier la récidive des groupes. Les éléments importants du cadre du groupe expérimental et du groupe témoin sont les mêmes.</i>

3. Divers facteurs secondaires

Cet élément concerne les divers facteurs secondaires qui pourraient introduire un biais dans les résultats. Il importe de saisir et d'évaluer la ressemblance des facteurs secondaires entre le groupe expérimental et le groupe témoin. Parmi les exemples de facteurs secondaires figurent le lieu physique du traitement (p. ex. hôpital psychiatrique isolé en milieu fermé, prison/pénitencier) et les différences possibles dans les méthodes de surveillance entre les délinquants traités et les délinquants non traités.

<i>Renseignements sur divers facteurs secondaires provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Divers facteurs secondaires

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote révisée des biais

0	1	2
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>

Sens révisé des biais

?	+1	0	-1
<i>Impossible d'évaluer le sens du biais</i>	<i>Le biais accroît probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Le biais diminue probablement le degré d'efficacité du traitement</i>

Raison(s) de la cote :

II. Attentes de l'expérimentateur

Les attentes de l'expérimentateur peuvent avoir un effet important sur les résultats des essais cliniques même randomisés (voir Zaza et coll., 2000; Juni et coll., 1999), ce qui explique pourquoi des procédures en double aveugle sont souvent utilisées. Les essais en double aveugle ne peuvent pas être appliqués directement au traitement des délinquants sexuels parce qu'il n'est pas réaliste de s'attendre à ce que les sujets et les expérimentateurs ne sachent pas qui reçoit le traitement. Certains examinateurs considèrent même comme avantageux pour l'expérimentateur de participer à l'élaboration et à l'exécution du programme, car cela peut donner des niveaux plus élevés d'intégrité du traitement (voir Gendreau, Goggin et Smith, 1999). Néanmoins, ceux qui ont un intérêt direct dans un programme risquent de prendre des décisions favorables à l'obtention d'un résultat positif du traitement.

Les essais à l'aveugle sont également importants pour la collecte et l'analyse des données. La connaissance du groupe auquel appartient un délinquant peut influencer par inadvertance sur les décisions en matière d'évaluation et de codage. Dans les études sur les résultats du traitement des délinquants sexuels, il est possible et souhaitable que les codeurs des données et les gestionnaires ne connaissent pas l'appartenance du groupe de sujets et les résultats.

Les Lignes directrices du CCDCR contiennent deux éléments qui permettent d'évaluer les attentes des expérimentateurs. Le premier élément, Intérêt de l'expérimentateur dans le résultat, sert à déterminer si les expérimentateurs introduisent un biais en participant aux travaux de recherche et à la prestation des services. L'expérimentateur risque fort d'introduire un biais lorsqu'il a un intérêt direct dans le résultat de l'évaluation et qu'il peut influencer sur les décisions quotidiennes concernant l'exécution du programme et la collecte des données. On s'attend à ce que les expérimentateurs aient un intérêt direct dans le résultat lorsqu'ils évaluent leurs propres programmes ou les programmes avec lesquels ils ont des liens étroits (p. ex. dans le cadre de leur affiliation à un établissement). Les renseignements pertinents comprennent les efforts particuliers ou les activités qui peuvent avoir été entreprises parce que l'expérimentateur participe au traitement et à d'autres procédures pour le groupe témoin et si les attentes de l'expérimentateur pourraient influencer sur la collecte des données. Il faut évaluer l'ampleur et le sens du biais.

Le deuxième élément, Évaluation à l'aveugle de la gestion des données, permet d'évaluer dans quelle mesure les gestionnaires des données peuvent avoir introduit un biais parce qu'ils connaissent la composition d'un groupe lorsqu'ils recueillent et analysent les données (en particulier le codage des renseignements sur la récidive).

4. Intérêt de l'expérimentateur dans le résultat

Concept : Le concept général est la mesure dans laquelle l'expérimentateur a un intérêt direct dans le traitement et dans laquelle il a la capacité d'influer sur le résultat de l'évaluation (p. ex. son niveau de participation à la prestation quotidienne des services). L'objectif consiste à évaluer le biais qui peut être introduit par sa participation au traitement et à son évaluation.

Indicateurs : Les indicateurs possibles du biais introduit par la participation de l'expérimentateur sont le fait qu'il ait un intérêt direct dans le programme de traitement et que l'évaluateur participe directement aux activités quotidiennes du programme et à la prestation des services. Même si la participation de l'expérimentateur peut influencer sur l'intégrité du traitement (évaluée par l'élément Définition du traitement), cet élément sert à évaluer l'influence que l'expérimentateur peut avoir sur les résultats de l'évaluation dans les décisions quotidiennes qui peuvent introduire un certain nombre de biais dans les résultats (p. ex. les décisions relatives à l'admission et les interventions supplémentaires qui accroissent la motivation et la conformité qui n'ont pas été décrites ou prises en note). Le chercheur peut introduire certains biais lorsqu'il évalue son propre travail.

Cote	Description
0	<i>L'expérimentateur a un intérêt direct dans le programme et l'évaluation et il participe directement à l'administration quotidienne du programme, y compris, par exemple, aux décisions relatives à l'admission, l'attrition et la prestation directe du traitement.</i>
1	<i>L'expérimentateur fait partie de l'équipe d'évaluation qui est liée indirectement, en raison de son rôle en matière de consultation et de surveillance, aux activités du programme. Le principal chercheur ne participe pas à la prestation directe des services. On ne sait pas clairement si son rôle influencerait sur les décisions quotidiennes. Une cote de 1 serait accordée pour les évaluations effectuées par une section distincte de la même organisation (p. ex. direction de la recherche ou de la vérification du même ministère qui exécute le programme).</i>
2	<i>L'expérimentateur est un évaluateur tiers du programme. Le principal chercheur ne participe pas à la gestion ou à la prestation des services. Il n'y a pas de raison de croire que l'expérimentateur a un intérêt direct dans le résultat, en raison de sa participation personnelle ou de son affiliation à l'établissement.</i>

4. Intérêt de l'expérimentateur dans le résultat

Cet élément porte sur le biais qui peut résulter de l'intérêt tout particulier de l'expérimentateur dans la gestion et la prestation des services et(ou) de sa participation directe à celles-ci, surtout quand il est responsable des décisions prises quotidiennement.

<i>Renseignements sur la participation de l'expérimentateur provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Intérêt de l'expérimentateur dans le résultat			
<i>Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?</i>			
<i>Si de nouveaux renseignements ont été obtenus, précisez.</i>			
Cotes révisées des biais			
0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>	
Sens révisé des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

5. Évaluation à l'aveugle de la gestion des données

Concept : Le concept général est la mesure dans laquelle les responsables de la gestion et du codage des données connaissent, ou auraient pu raisonnablement connaître, l'appartenance au groupe et le résultat (c.-à-d. la récidive) des participants à l'étude et savoir si cette connaissance pouvait introduire un biais.

Indicateurs : Les indicateurs possibles d'un biais dans les procédures de gestion des données seraient le moment où les codeurs ou gestionnaires des données connaissent a priori l'appartenance du délinquant au groupe et(ou) le statut de la récidive pendant les périodes de collecte des données. Même si ces renseignements ne sont pas essentiels pour la simple transcription ou la photocopie des données, ils sont cruciaux lorsqu'il faut évaluer ou coder des décisions. Par exemple, le fait de savoir au préalable que le délinquant fait partie du groupe expérimental peut influencer sur les décisions en matière de codage pendant l'évaluation rétrospective du risque. Idéalement, le responsable de la collecte, de la gestion et de l'analyse des données ne connaîtrait pas l'appartenance au groupe et le résultat. Lorsque aucun renseignement n'est fourni concernant le mode de collecte et d'analyse des données, cet élément ne doit pas être coté (c.-à-d. « renseignements insuffisants pour permettre une évaluation »).

Cote	Description
0	<i>Les chercheurs/gestionnaires des données connaissent a priori l'appartenance au groupe ou le statut de la récidive pendant les procédures relatives aux données exigeant de prendre des décisions ou de porter des jugements.</i>
1	<i>On ne sait pas clairement si les chercheurs ou gestionnaires des données connaissent a priori l'appartenance au groupe ou le statut de la récidive pendant les procédures relatives aux données exigeant de prendre des décisions ou de porter des jugements.</i>
2	<i>Les chercheurs/gestionnaires des données ne connaissent pas l'appartenance au groupe et le statut de la récidive pendant les procédures relatives aux données exigeant de prendre des décisions ou de porter des jugements.</i>

5. Évaluation à l'aveugle de la gestion des données

Cet élément consiste à déterminer si ceux qui ont recueilli et enregistré les données sur les participants à l'étude pouvaient introduire des biais dans les résultats. Les données comme les renseignements sur l'évaluation des risques et la récidive sont essentielles à l'évaluation de l'effet du traitement. Il convient de prendre note des renseignements qui indiqueraient si la ou les personnes qui a recueilli, évalué ou enregistré les renseignements connaissaient ou non le statut du traitement et(ou) de la récidive des participants.

<i>Renseignements sur l'évaluation à l'aveugle de la gestion des données provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Évaluation à l'aveugle de la gestion des données

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cotes révisées des biais

0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>	
Sens révisé des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>

Raison(s) de la cote :

III. Taille de l'échantillon

Plus la taille de l'échantillon d'une étude est grande, plus on a confiance dans les résultats. L'intervalle de confiance concernant l'ampleur de l'effet diminue à mesure que la taille de l'échantillon augmente tout comme l'efficacité statistique de la détection des différences préexistantes entre les groupes. Dans le cas des études à affectation aléatoire, on s'attend à ce que l'erreur entre les groupes s'annule (c.-à-d. qu'elle soit aléatoire). Par conséquent, on exige moins des études à affectation aléatoire que dans le cas des autres genres pour qu'il n'y ait pas de différences préexistantes entre les groupes. Lorsque les participants sont appariés en fonction du risque avant la randomisation, il est encore moins nécessaire de vérifier la possibilité d'une différence préexistante que dans le cas des autres études à affectation aléatoire. Pour toutes les études, cependant, il est souhaitable que l'échantillon soit d'une grande taille. Même les études bien conçues échouent, et une efficacité statistique adéquate permet aux examinateurs d'explorer des écarts possibles par rapport au protocole prévu.

Dans le cas des études à affectation non aléatoire, il faut absolument vérifier l'équivalence prétest des groupes expérimentaux et témoins. À mesure que la taille de l'échantillon diminue, il devient de plus en plus difficile de déterminer les différences préexistantes importantes (c.-à-d. les biais inhérents à la composition des groupes). Pour l'obtention de niveaux de confiance égaux, les études à affectation non aléatoire exigent des échantillons d'une taille plus grande que dans le cas des études à affectation aléatoire. Les échantillons de taille plus petite dans le cas des études à affectation aléatoire, toutefois, ne conviennent que lorsque la procédure d'affectation aléatoire est bien appliquée. En cas de défaillance du processus d'affectation aléatoire, les examinateurs doivent examiner l'équivalence des groupes, et des échantillons de taille plus grande que d'habitude sont nécessaires.

L'efficacité statistique est optimisée lorsque le l'échantillon total de l'étude est réparti également entre les groupes (c.-à-d. 50 % de l'échantillon total de l'étude). La perte d'efficacité est minime lorsque la taille des échantillons est inégale, cependant, pourvu que la répartition se situe dans la fourchette allant de 30 à 70 % de l'échantillon total de l'étude. Par exemple, on obtient une efficacité maximale dans le cas d'une étude dont la taille de l'échantillon total est de 600 sujets lorsqu'il y a 300 sujets dans le groupe expérimental et 300 sujets dans le groupe témoin. La perte d'efficacité est minimale lorsqu'un groupe compte 180 sujets (30 % de l'échantillon total de 600 sujets de l'étude) et que l'autre groupe en compte 420. Lorsque l'écart se situe en dehors de cette fourchette, la capacité de déterminer les différences statistiques entre les groupes est réduite (Cohen, 1988).

Dans les modèles fondés sur les établissements, l'unité d'analyse est constituée d'établissements plutôt que de délinquants. Ces modèles sont particuliers en ce sens que la taille de l'échantillon renvoie également au nombre d'établissements. Comme ils définissent et évaluent les concepts/caractéristiques des établissements, il importe que le nombre d'établissements soit

assez grand pour permettre une variation dans le concept en question. La confiance dans les résultats augmente si plus d'établissements font partie de l'échantillon.

Les Lignes directrices du CCDCR contiennent trois éléments servant à évaluer la taille de l'échantillon. Le premier sert à évaluer la confiance que la taille de l'échantillon du groupe expérimental est assez grande pour permettre de déterminer des différences préexistantes importantes qui pourraient influencer sur le résultat. La taille de l'échantillon a trait au nombre de participants pour lesquels des données de suivi ont été obtenues. Les cotes de confiance sont fondées sur la taille de l'échantillon nécessaire pour permettre une efficacité suffisante (c.-à-d. 70 %) pour déterminer les effets minimes (c.-à-d. $d = 0,20$), moyens (c.-à-d. $d = 0,50$) et importants (c.-à-d. $d = 0,80$) (Cohen, 1988). Des critères applicables aux études à affectation aléatoire et aux études à affectation non aléatoire sont présentés.

Le deuxième élément sert à évaluer la confiance que la taille de l'échantillon du groupe témoin est assez grande pour déterminer des différences préexistantes importantes qui pourraient influencer sur le résultat. Les critères des cotes de confiance sont identiques à ceux de l'élément précédent.

Le troisième élément s'applique aux modèles fondés sur les établissements et il n'est pas prévu de l'utiliser avec les autres modèles. Il sert à évaluer la confiance que la taille de l'échantillon d'établissements est assez grande pour permettre une variation du concept en question et de déterminer des différences importantes entre les établissements. Lorsqu'il y a moins de six établissements dans l'échantillon, le niveau de confiance est faible. Lorsqu'il y a plus de huit établissements, on a confiance que l'échantillon est assez nombreux pour permettre d'évaluer et d'examiner les différences dans les taux de récidive. La confiance augmente à mesure que s'accroissent le nombre d'établissements où le concept est présent et le nombre d'établissements où le concept est absent. Le niveau de confiance serait généralement plus élevé si la répartition était de 50-50 (4 contre 4) plutôt qu'extrême (1 contre 7).

Dans le cas des modèles fondés sur les établissements, l'efficacité statistique est encore déterminée par le nombre de délinquants, et non le nombre d'établissements. Par conséquent, il est utile d'évaluer la taille de l'échantillon du groupe expérimental et la taille de l'échantillon du groupe témoin ainsi que la taille de l'échantillon d'établissements.

6. Taille de l'échantillon du ou des groupes expérimentaux

Concept : Le concept général est une taille de l'échantillon suffisante pour permettre de déterminer les différences préexistantes importantes entre les groupes. Cet élément vise à évaluer la confiance du codeur que la taille de l'échantillon est suffisante pour permettre une efficacité de 70 % dans la détermination des effets minimes ($d = 0,20$), moyens ($d = 0,50$) et(ou) importants ($d = 0,80$) entre les groupes. La confiance augmente lorsqu'il y a une forte probabilité que la taille de l'échantillon soit assez grande pour permettre de déterminer ces différences (c.-à-d. efficacité). Pour les études à affectation aléatoire bien mises en œuvre, l'affectation des sujets réduit au minimum la probabilité de différences préexistantes entre les groupes. Par conséquent, les critères concernant les cotes de confiance pour ce genre d'étude sont les tailles d'échantillon nécessaires pour déterminer les effets moyens ($d = 0,50$) et importants ($d = 0,80$), avec une efficacité de 70 %. Pour toutes les autres études, il est crucial de démontrer l'équivalence préexistante des groupes. Par conséquent, les critères pour les cotes de confiance de ces études sont la taille des échantillons nécessaires pour déterminer les effets minimes ($d = 0,20$) et moyens ($d = 0,50$) avec une efficacité de 70 %. Les valeurs proviennent du tableau d'efficacité 2.3.5 (intervalles bilatéraux α de 0,05) de Cohen (1988).

Indicateurs : Pour évaluer cet élément, il faut définir le modèle et la taille de l'échantillon total de l'étude et il faut connaître la taille de l'échantillon du groupe expérimental ainsi que les données de suivi. Il y a une certaine souplesse dans les cotes, car la taille des échantillons décrits ci-dessous est fondée sur 50 % du nombre (N) total dans l'étude. Si le nombre (N) total est le double des critères ci-dessous et si le groupe expérimental constitue de 30 à 70 % du nombre (N) total, la cote de confiance est appropriée.

Cote	Description
0	<i>Pour les études à affectation aléatoire, la taille de l'échantillon expérimental est inférieure à 20 sujets. Pour toutes les autres études, la taille de l'échantillon expérimental est inférieure à 50.</i>
1	<i>Pour les études à affectation aléatoire, la taille de l'échantillon expérimental est supérieure à 20 sujets, mais inférieure à 50. Pour toutes les autres études, la taille de l'échantillon expérimental est supérieure à 50 sujets, mais inférieure à 300.</i>
2	<i>Pour les études à affectation aléatoire, la taille de l'échantillon expérimental est de 50 sujets ou plus. Pour toutes les autres études, la taille de l'échantillon expérimental est de 300 sujets ou plus.</i>

6. Taille de l'échantillon du ou des groupes expérimentaux

Cet élément porte sur la taille de l'échantillon du groupe expérimental et les données de suivi ainsi que le niveau de confiance que la taille de cet échantillon a une efficacité suffisante pour permettre de déterminer les différences. Les décisions concernant la suffisance de la taille des échantillons suffisante doivent être fondées sur la taille de l'échantillon avec données de suivi .

<i>Taille de l'échantillon total de l'étude :</i>		<i>N =</i>	
<i>Taille de l'échantillon du groupe expérimental avec données de suivi :</i>		<i>n =</i>	
<i>Pourcentage de sujets de l'étude dans le groupe expérimental :</i>		<i>%</i>	
Cote de confiance			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Peu de confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences</i>	<i>Certaine confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences</i>	<i>Grande confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : taille de l'échantillon du ou des groupes expérimentaux

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences

1

Certaine confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences

2

Grande confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences

Raison(s) de la cote :

7. Taille de l'échantillon du ou des groupes témoins

Concept : Le concept général est une taille de l'échantillon suffisante pour permettre de déterminer les différentes préexistantes importantes entre les groupes qui pourraient influencer sur les résultats. Cet élément vise à évaluer la confiance du codeur que la taille de l'échantillon est suffisante pour permettre une efficacité de 70 % dans la détermination des effets minimes ($d = 0,20$), moyens ($d = 0,50$) et(ou) importants ($d = 0,80$) entre les groupes. La confiance augmente lorsqu'il y a une forte probabilité que la taille de l'échantillon soit assez grande pour permettre de déterminer ces différences (c.-à-d. efficacité). Pour les études à affectation aléatoire bien mises en œuvre, l'affectation des sujets réduit au minimum la probabilité de différences préexistantes entre les groupes. Par conséquent, les critères concernant les cotes de confiance pour ce genre d'étude sont les tailles d'échantillon nécessaires pour déterminer les effets moyens ($d = 0,50$) et importants ($d = 0,80$), avec une efficacité de 70 %. Pour toutes les autres études, il est crucial de démontrer l'équivalence préexistante des groupes. Par conséquent, les critères pour les cotes de confiance de ces études sont la taille des échantillons nécessaires pour déterminer les effets minimes ($d = 0,20$) et moyens ($d = 0,50$) avec une efficacité de 70 %. Les valeurs proviennent du tableau d'efficacité 2.3.5 (intervalles bilatéraux α de 0,05) de Cohen (1988)

Indicateurs : Pour évaluer cet élément, il faut définir le modèle et la taille de l'échantillon total de l'étude et il faut connaître la taille de l'échantillon du groupe témoin ainsi que les données de suivi. Il y a une certaine souplesse dans les cotes, car la taille des échantillons décrits ci-dessous est fondée sur 50 % du nombre (N) total dans l'étude. Si le nombre (N) total est le double des critères ci-dessous et si le groupe témoin constitue de 30 à 70 % du nombre (N) total, la cote de confiance est appropriée.

Modèles de gammes de risques/normes : Le groupe témoin dans ces modèles est un échantillon normatif prélevé dans d'autres études. Si ces normes sont raisonnables, une cote de « 2 – grande confiance » convient.

Cote	Description
0	<i>Pour les études à affectation aléatoire, la taille de l'échantillon témoin est inférieure à 20 sujets.</i> <i>Pour toutes les autres études, la taille de l'échantillon témoin est inférieure à 50.</i>
1	<i>Pour les études à affectation aléatoire, la taille de l'échantillon témoin est supérieure à 20 sujets, mais inférieure à 50.</i> <i>Pour toutes les autres études, la taille de l'échantillon témoin est supérieure à 50 sujets, mais inférieure à 300.</i>
2	<i>Pour les études à affectation aléatoire, la taille de l'échantillon témoin est de 50 sujets ou plus.</i> <i>Pour toutes les autres études, la taille de l'échantillon témoin est de 300 sujets ou plus.</i>

7. Taille de l'échantillon du ou des groupes témoins

Cet élément porte sur la taille de l'échantillon du groupe témoin et les données de suivi ainsi que le niveau de confiance que la taille de cet échantillon a une efficacité suffisante pour permettre de déterminer les différences. Les décisions concernant la suffisance de la taille des échantillons doivent être fondées sur la taille de l'échantillon avec données de suivi.

<i>Taille de l'échantillon total de l'étude :</i>		<i>N =</i>	
<i>Taille de l'échantillon du groupe témoin avec données de suivi :</i>		<i>n =</i>	
<i>Pourcentage de sujets de l'étude dans le groupe témoin :</i>		<i>%</i>	
Cote de confiance			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Peu de confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences</i>	<i>Certaine confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences</i>	<i>Grande confiance que l'échantillon est assez nombreux pour permettre une efficacité suffisante pour déterminer les différences</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Taille de l'échantillon du ou des groupes témoins

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

*Peu de confiance que
l'échantillon est assez nombreux
pour permettre une efficacité
suffisante pour déterminer les
différences*

1

*Certaine confiance que
l'échantillon est assez nombreux
pour permettre une efficacité
suffisante pour déterminer les
différences*

2

*Grande confiance que
l'échantillon est assez nombreux
pour permettre une efficacité
suffisante pour déterminer les
différences*

Raison(s) de la cote :

8. Taille de l'échantillon d'établissements : modèles fondés sur les établissements

Concept : Comme l'unité d'analyse dans les modèles fondés sur les établissements est l'établissement plutôt que le délinquant, le nombre d'établissements de l'échantillon est essentiel. La taille de l'échantillon doit être suffisante pour que plus d'un établissement constitue une variation par rapport au concept en question.

Indicateurs : Lorsqu'on porte un jugement sur la taille de l'échantillon, il vaut la peine de tenir compte du nombre d'établissements et du nombre de sujets représentés par les établissements. Il faut plus d'établissements si le nombre provenant de chaque établissement est faible (voir les éléments précédents).

Cote	Description
0	<i>Des données ont été recueillies et l'analyse était fondée sur moins de six établissements ou unités opérationnelles.</i>
1	<i>Des données ont été recueillies et l'analyse était fondée sur de six à huit établissements ou unités opérationnelles.</i>
2	<i>Des données ont été recueillies et l'analyse était fondée sur plus de huit établissements ou unités opérationnelles.</i>

8. Taille de l'échantillon d'établissements : modèles fondés sur les établissements

Cet élément porte sur la taille de l'échantillon d'établissements et le niveau de confiance que cette taille assure une variabilité suffisante du concept en question pour permettre de déterminer les différences. Des données sur le nombre d'établissements de l'échantillon et la taille de l'échantillon pour chaque établissement doivent être enregistrées.

<i>Nombre total d'établissements de l'échantillon :</i>		<i>N =</i>	
<i>Taille de l'échantillon pour chaque établissement avec données de suivi :</i>		<i>n₁ =</i>	<i>n₆ =</i>
		<i>n₂ =</i>	<i>n₇ =</i>
		<i>n₃ =</i>	<i>n₈ =</i>
		<i>n₄ =</i>	<i>n₉ =</i>
		<i>n₅ =</i>	<i>n₁₀ =</i>
Cote de confiance			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Peu de confiance que la taille soit assez grande pour permettre une variation suffisante du concept en question</i>	<i>Certaine confiance que la taille soit assez grande pour permettre une variation suffisante du concept en question</i>	<i>Grande confiance que la taille soit assez grande pour permettre une variation suffisante du concept en question</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Taille de l'échantillon des établissements – modèles fondés sur les établissements

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance que la taille soit assez grande pour permettre une variation suffisante du concept en question

1

Certaine confiance que la taille soit assez grande pour permettre une variation suffisante du concept en question

2

Grande confiance que la taille soit assez grande pour permettre une variation suffisante du concept en question

Raison(s) de la cote :

IV. Attrition

Le départ (attrition) des participants est un facteur important à prendre en considération dans l'évaluation de la qualité d'une étude. Il faut tenir compte du nombre de participants perdus et de l'étape de l'étude où le départ s'est produit. À mesure que l'attrition augmente, la possibilité de biais s'accroît. Plus précisément, l'attrition se produit d'abord pendant le processus de sélection des sujets. À cette étape, les participants à l'étude, en particulier les participants éventuels du groupe expérimental, sont déterminés. Les chercheurs excluent certains délinquants (p. ex. les délinquants psychotiques). De plus, il y a des délinquants qui refusent de participer ou qui rejettent carrément le traitement (p. ex. qui refusent de participer parce qu'« ils ne sont pas des délinquants sexuels »).

Il est extrêmement difficile d'interpréter le taux de récurrence des « délinquants qui refusent de participer » par rapport au groupe expérimental. Si le taux de récurrence des « délinquants qui refusent de participer » est semblable à celui du groupe témoin et est plus élevé que celui du groupe expérimental, cela semblerait corroborer l'efficacité du traitement. Il semble que les « délinquants qui refusent de participer » auraient pu réduire leur potentiel de récurrence en participant au traitement. Cependant, si l'on incluait les « délinquants qui refusent de participer » dans le groupe expérimental dans le but de procéder à une analyse des sujets retenus au début de l'étude, cela réduirait l'effet apparent du traitement. Par contre, si les « délinquants qui refusent de participer » affichent un faible taux de récurrence (semblable à celui du groupe expérimental et inférieur à celui du groupe témoin), le traitement semblerait inutile et pourrait même avoir un effet négatif sur certains délinquants. Paradoxalement, le fait d'inclure les « délinquants qui refusent de participer » dans le groupe expérimental accroîtrait l'efficacité apparente du traitement. Même si les données empiriques donnent à penser que les taux de récurrence observés chez ceux qui refusent le traitement sont semblables à ceux qui commencent le traitement (Hanson et coll., 2002), les études les plus faciles à interpréter sont celles dans laquelle le traitement est accepté par la plupart des délinquants à qui il est offert et où il y a peu de « délinquants qui peuvent refuser de participer » dans le groupe témoin.

La population des délinquants « traitables » se limite aux délinquants dont la participation serait envisagée (c.-à-d. qui répondent à certains critères de sélection) et aux délinquants qui envisageraient de participer à l'étude ou au traitement. C'est ce sous-ensemble qui sert de base à l'évaluation de l'efficacité du traitement. Toutefois, des critères de sélection étroits peuvent entraîner la constitution d'un échantillon artificiellement homogène, ce qui limiterait la généralisabilité des résultats (Westen, Novotny et Thompson-Brenner, 2004). De plus, un critère de sélection très restrictif accroît la difficulté de créer un groupe témoin équivalent.

L'importance du biais se rapporte à l'équivalence des critères de sélection pour le groupe expérimental et le groupe témoin (p. ex. risque élevé, motivé, désireux d'envisager une participation, dont la période de la peine à purger est assez longue pour terminer le traitement). Des renseignements clairs et précis sur les critères de sélection pour les délinquants interrogés en

vue de leur participation éventuelle à l'étude (au sein de groupes expérimentaux et(ou) témoins, le cas échéant) permettent au lecteur d'estimer les effets éventuels des biais de sélection.

En plus du biais de sélection initial, les délinquants peuvent abandonner le traitement à divers moments. Par exemple, certains délinquants peuvent abandonner après avoir été acceptés, mais avant de commencer réellement le traitement, et d'autres peuvent abandonner après avoir assisté à une ou plusieurs séances de traitement. Nous désignons cette perte sous le nom d'attrition.

Il y a de solides preuves selon lesquelles les délinquants qui abandonnent le traitement risquent plus de récidiver que ceux qui le terminent (Hanson et coll., 2002). Par conséquent, les renseignements sur ceux qui abandonnent le traitement devraient être enregistrés régulièrement, et ceux-ci devraient être inclus avec le groupe expérimental (analyses des sujets retenus au début de l'étude). De plus, les chercheurs devraient consigner les raisons de l'attrition : a) les fournisseurs du traitement mettent fin au traitement en raison d'une mauvaise conduite au sein ou à l'extérieur du programme; b) retrait volontaire (p. ex. manque d'intérêt ou de motivation); c) raisons administratives (p. ex. annulation du programme). Au moment de la conception des études éventuelles, les chercheurs doivent envisager des méthodes pour limiter l'attrition pendant le traitement et être conscients du risque d'attrition différentielle. Par exemple, l'attrition différentielle peut se produire lorsque les participants à risque élevé et plus difficiles abandonnent un programme de traitement rigoureux alors que les participants à risque élevé du groupe témoin n'abandonnent pas parce qu'on exige peu d'eux. Même si une partie du biais introduit par l'attrition peut être prise en compte au moyen d'analyses des sujets retenus au début de l'étude, la meilleure façon de décrire des taux d'attrition élevés consiste à les considérer comme un échec de la mise en œuvre du programme plutôt que comme un test de l'efficacité du traitement. Des études prospectives devraient inclure des mesures visant à établir et à soutenir la participation des délinquants.

La perte de participants pendant le suivi est également une menace à la validité. Si seulement le casier judiciaire sert à détecter la récidive, il faudrait pouvoir repérer tous les participants avec la même efficacité. Un problème éventuel que peut poser l'utilisation de dossiers officiels est la suppression des dossiers anciens et inactifs (Hanson et Nicholaichuk, 2000). Lorsque les groupes expérimentaux et témoins sont choisis à partir de périodes différentes ou dans des administrations différentes, les méthodes de conservation des dossiers peuvent introduire un biais. Par exemple, dans les études rétrospectives, l'appariement inversé des dossiers existants pour lequel les groupes sont choisis à partir des dossiers existants peut introduire un biais du simple fait que seulement certains délinquants auraient encore un casier judiciaire (généralement, les jeunes délinquants criminellement actifs). Il y a un autre problème : il se peut que les commissions d'examen des établissements n'autorisent pas les chercheurs à recueillir ou à conserver des données sur les individus qui se retirent de l'étude. La probabilité de taux d'attrition différentielle pendant le suivi est plus forte lorsque les données sur le suivi sont recueillies directement auprès des participants. De plus, les chercheurs peuvent avoir établi des relations beaucoup plus étroites avec les participants au traitement actif qu'avec ceux du groupe témoin. De bonnes études prospectives doivent être fondées sur des méthodes explicites pour assurer le suivi des participants et les maintenir dans le protocole (Shadish, Cook et Campbell, 2002).

Les lignes directrices du CCDCR contiennent quatre éléments permettant d'évaluer le biais susceptible d'être introduit par l'attrition. Le premier élément évalue le biais potentiel des critères explicites utilisés pour la sélection des sujets dans les groupes expérimental et témoin. Les renseignements pertinents comprennent les critères utilisés pour définir le groupe de délinquants éventuels, y compris les délinquants qui refusent de participer. Si les critères utilisés créent une attente selon laquelle un ou plusieurs groupes seraient différents en ce qui concerne des variables importantes, cela introduit un biais. Enfin, si les critères de sélection sont stricts ou idiosyncrasiques, il sera difficile d'appliquer les résultats à d'autres populations. Les cotes de l'ampleur estimée et du sens du biais (c.-à-d. favorise l'efficacité du traitement ou diminue l'efficacité du traitement) sont nécessaires.

Le deuxième élément évalue le biais susceptible d'être introduit par l'abandon du programme. Cet élément porte sur les départs qui se produisent après que le groupe est assigné ou offert au délinquant, mais avant l'achèvement (c.-à-d. abandon). Les renseignements pertinents comprennent le nombre et les caractéristiques des délinquants qui ont abandonné rapidement et les raisons pour lesquelles ils l'ont fait (c.-à-d. après avoir accepté de participer, mais avant de commencer officiellement le programme) ou qui ont abandonné plus tard (c.-à-d. après avoir commencé officiellement le programme). Des évaluations de l'ampleur estimée et du sens du biais sont nécessaires. L'élément ne porte pas sur les procédures d'analyse utilisées (p. ex. analyse des sujets retenus au début de l'étude), mais seulement sur le nombre et les caractéristiques de ceux qui ont abandonné et les raisons de leur abandon. Si le nombre de délinquants qui ont abandonné est suffisamment important (supérieur à 50 %), la mise en oeuvre de l'étude est considérée comme un « échec ».

Le troisième élément, Analyse des sujets retenus au début de l'étude, évalue comment le ou les chercheurs traitent l'attrition susmentionnée lorsqu'ils estiment les effets du traitement. Un biais peut être introduit lorsque le taux d'abandon du programme est plus que minimal et lorsque ces participants sont exclus de l'analyse. Dans tous les cas, une analyse des sujets retenus au début de l'étude est l'approche la moins biaisée (c.-à-d. qu'elle tient compte des « abandons » dans le calcul du taux de récidive général du groupe expérimental). Il est important d'enregistrer tout renseignement fourni sur le groupe des délinquants qui ont abandonné (c.-à-d. risque, proportion de l'ensemble du groupe, taux de récidive, raisons de l'attrition) et de déterminer si ce groupe a été inclus ou non dans le groupe expérimental pour l'estimation de l'efficacité du traitement. Des évaluations de l'ampleur estimative et du sens du biais par rapport à l'analyse des sujets retenus au début de l'étude sont nécessaires.

Le quatrième et dernier élément de la section sur l'attrition est l'attrition pendant le suivi. Les renseignements pertinents sont le nombre de participants perdus en raison du manque de renseignements sur le suivi. Les biais attribuables à l'attrition pendant le suivi peuvent être introduits lorsqu'il y a des taux d'attrition différentielle entre les groupes ou lorsque les taux généraux sont élevés dans les deux groupes. Des évaluations de l'ampleur estimative et du sens des biais attribuables à l'attrition pendant le suivi sont nécessaires.

9. Sélection des sujets

Concept : Cet élément concerne les biais introduits par les critères servant à déterminer quels délinquants peuvent participer à un programme de traitement. Tous les programmes fixent des limites concernant l'acceptation (et[ou] le refus de l'acceptation) à un traitement en fonction de facteurs comme les besoins perçus (p. ex. préférences sexuelles déviantes) et la capacité d'en bénéficier (p. ex. la motivation, l'absence d'une psychose active). Ces limites peuvent introduire des biais et limiter la généralisabilité. Cet élément ne porte pas directement sur la conception de l'étude ou les procédures d'affectation des sujets à des groupes expérimentaux, car cet aspect est examiné sous l'élément 13 (c.-à-d. l'évaluation de l'équivalence des groupes a priori).

Les critères de sélection pour le traitement peuvent être généraux et vastes (p. ex. reconnu coupable d'une infraction sexuelle) ou spécifiques et précis (p. ex. pédophiles non diagnostiqués qui vont se faire traiter dans des cliniques privées). Plus les critères sont précis et spécifiques, plus il y a de possibilités que des biais soient introduits et que les résultats ne s'appliquent pas à la population générale des délinquants sexuels.

Il peut y avoir un lien étroit entre les critères de sélection (exclusion ou inclusion) relatifs à la participation ou non au traitement et le risque (p. ex. les préférences sexuelles déviantes, les délinquants à risque élevé, les psychopathes) ou un lien éloigné avec le risque (p. ex. temps à purger pendant la peine, psychose active, langue, alphabétisation). Plus le lien entre les critères de sélection et le risque est étroit, plus il est important de tenir compte de l'incidence des critères d'inclusion et d'exclusion sur les caractéristiques des groupes expérimentaux. Par exemple, lorsqu'il y a un lien évident entre les critères d'exclusion et le risque (p. ex. psychopathes), beaucoup de biais peuvent être introduits si des critères d'exclusion identiques ne sont pas utilisés pour le groupe témoin (p. ex. le groupe témoin peut compter des psychopathes).

Indicateurs : Pour évaluer cet élément, il faut déterminer les critères de sélection pour l'admission au traitement (critères d'inclusion et d'exclusion). Lorsque ces critères sont généraux et vastes, il y a plus de chances que le groupe expérimental et le groupe témoin soient comparables, en particulier lorsque des critères identiques sont utilisés pour les deux groupes. Lorsque les critères de sélection sont précis, en plus de limiter la généralisabilité, cela accroît l'importance d'utiliser des critères identiques pour le groupe témoin.

Un autre indicateur est le rapport entre les critères d'exclusion et d'inclusion et le risque. Parmi les exemples de lien clair entre les facteurs de sélection et le risque, il y a le fait de restreindre le traitement aux personnes ayant des préférences sexuelles déviantes, de choisir seulement les clients « faciles à gérer » ou d'exclure les délinquants ayant des troubles de la personnalité. Parmi les facteurs de sélection pour lesquels il y a un rapport faible ou indirect avec le risque, mentionnons la période de la peine qui reste à purger, la psychose active et les facteurs de réceptivité comme la motivation, la langue et l'alphabétisation. Lorsque la proportion de délinquants exclus est faible, il faut un lien clair avec le risque pour qu'un biais important soit introduit. Lorsque le taux d'exclusion est élevé (p. ex. seulement les délinquants analphabètes

sont admis au traitement), un biais important peut être introduit par des facteurs qui ont un rapport faible ou indirect avec le risque de récidive.

Enfin, dans le cas des études où l'expérimentateur n'a pas déterminé qui a reçu le traitement (p. ex. les études de cohortes pour lesquelles le traitement a été donné au fil du temps), il faut tenir compte du pourcentage du groupe témoin qui aurait été traité s'il avait eu accès au traitement (p. ex. en procédant à une estimation à partir du taux de refus ou d'abandon du groupe expérimental).

Cote de sélection des sujets

Cote	Description
0	<p><i>Les critères de sélection entre les deux groupes n'étaient pas identiques ET il y a un lien direct entre les critères d'exclusion du traitement et le risque.</i></p> <p style="text-align: center;">OU</p> <p><i>Les critères de sélection entre les deux groupes n'étaient pas identiques ET la proportion relative des délinquants exclus du traitement est importante par rapport à celle du groupe témoin.</i></p>
1	<p><i>Les critères de sélection entre les deux groupes n'étaient pas identiques. La proportion des délinquants exclus est faible, et l'on s'attend à ce que le rapport entre les critères d'exclusion du traitement et le risque soit faible.</i></p> <p style="text-align: center;">OU</p> <p><i>Les critères de sélection sont identiques, mais inhabituellement étroits, de telle sorte que le groupe expérimental et le groupe témoin ne sont pas représentatifs des autres échantillons auxquels l'intervention peut s'appliquer.</i></p>
2	<p><i>Les critères de sélection sont identiques. Les critères d'inclusion et d'exclusion sont généraux et vastes et s'appliquent aux groupes expérimental et témoin.</i></p>

9. Sélection des sujets

Cet élément porte sur les exigences relatives à l'« admission au traitement », y compris les critères d'inclusion et d'exclusion. Il y a de plus en plus de biais à mesure que ces critères deviennent plus restrictifs et il devient plus difficile d'utiliser ces critères pour choisir les sujets des groupes témoins. En outre, il y a plus de biais lorsque le nombre de délinquants pouvant refuser de participer qui font partie du groupe témoin augmente.

<i>Renseignements sur la sélection des sujets provenant de l'étude : (critères d'inclusion et d'exclusion)</i>			
		<i>Groupe expérimental</i>	<i>Groupe témoin</i>
<i>Nombre de tous les sujets possibles</i>		<i>N =</i>	<i>N =</i>
<i>Nombre de sujets prêts à participer à prendre en considération et qu'il conviendrait d'inclure</i>		<i>N =</i>	<i>N =</i>
Cotes des biais			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Sélection des sujets			
<i>Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?</i>			
<i>Si de nouveaux renseignements ont été obtenus, précisez.</i>			
	<i>Groupe expérimental</i>		<i>Groupe témoin</i>
<i>Nombre de tous les sujets possibles</i>	<i>N =</i>		<i>N =</i>
<i>Nombre de sujets prêts à participer à prendre en considération et qu'il conviendrait d'inclure</i>	<i>N =</i>		<i>N =</i>
Cotes révisées des biais			
0	1	2	
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>	
Sens révisé des biais			
?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

10. Abandon du programme

Concept : Le concept général est le biais introduit par les départs survenus entre le moment où un délinquant manifeste la volonté de participer au traitement (ou se voit affecter à un groupe précis de l'étude) et l'achèvement de ce traitement. Le départ des participants à l'étude a lieu lorsque, après avoir manifesté la volonté de participer et avoir été acceptés, les délinquants changent d'idée et refusent de participer à l'étude. Certains peuvent abandonner avant d'assister aux séances (p. ex. ils changent d'idée après avoir appris qu'ils doivent être transférés dans un établissement différent). Ils peuvent aussi cesser de participer au traitement (ou à l'étude) après avoir commencé le traitement pour diverses raisons (p. ex. cessation volontaire, décision des fournisseurs de traitement de les renvoyer, raisons administratives). Cet élément vise à évaluer l'ampleur et le sens des biais qui résultent de cette attrition. Le biais est proportionnel à l'augmentation du nombre d'« abandons » et(ou) à la différence entre les « abandons » d'un groupe et ceux de l'autre groupe. La mise en œuvre d'un programme doit être considérée comme un échec lorsque le pourcentage des « abandons » est trop élevé.

Indicateurs : Les indicateurs qui pourraient diminuer les biais seraient les cas où un pourcentage élevé de participants a terminé le traitement entrepris. Afin d'évaluer la proportion des abandons, il faut connaître le nombre de sujets qui ont envisagé de participer, le nombre de sujets qui ont décidé tôt d'abandonner (c.-à-d. des manifestations initiales d'intérêt aux premiers stades du traitement), les raisons du retrait et le nombre de sujets qui ont cessé de participer plus tard (c.-à-d. qui ont commencé le traitement ou la participation à l'étude, mais qui ne l'ont pas terminé). Lorsque le pourcentage de ceux qui commencent et terminent un traitement est élevé, il y a une diminution des biais.

Échec de la mise en œuvre : Les abandons sont le fait de ceux qui ont manifesté un intérêt au départ à l'égard du traitement ou de la participation à l'étude (p. ex. ils n'ont pas rejeté au début le traitement ou la participation), mais qui n'ont pas terminé le traitement ou leur participation à l'étude. Lorsque le taux d'abandon est de 50 % ou plus, il est considéré comme un échec de la mise en œuvre.

Cote	Description
0	<i>La mise en oeuvre du programme est un échec (c.-à-d. taux d'abandon de 50 % ou plus)</i> OU <i>De 51 % à 59 % des participants ont terminé le traitement. Le taux d'abandon oscille entre 41 % et 49 %¹.</i>
1	<i>Entre 60 % et 79 % ont terminé le traitement. Le taux d'abandon varie de 21 % à 40 %¹.</i>
2	<i>Au moins 80 % des sujets ont terminé le traitement ou la participation. Le taux d'abandon est de 20 % ou moins¹.</i>

¹ Ces critères proviennent de Thomas et coll. (2004).

10. Abandon du programme

Cet élément porte sur l'attrition, c'est-à-dire après que les participants ont indiqué qu'ils aimeraient participer (au traitement ou à une étude), avant la fin du traitement ou de l'expérience (c.-à-d. abandon). Les biais augmentent à mesure que le pourcentage des abandons s'accroît. Si le taux d'abandon est de 50 % ou plus, le biais dans les résultats, quelle que soit la façon dont le chercheur traite l'attrition lorsqu'il estime l'efficacité du traitement, est trop important, et la mise en œuvre du programme est considérée comme un échec.

<i>Renseignements sur l'abandon du programme provenant de l'étude</i>		<i>Groupe expérimental</i>	<i>Groupe témoin</i>
<i>Nombre de sujets prêts à participer à prendre en considération et qu'il conviendrait d'inclure (voir l'élément précédent)</i>		<i>N =</i>	<i>N =</i>
<i>Nombre de sujets qui ont cessé de participer avant la fin du premier mois de traitement</i>		<i>N =</i>	<i>N =</i>
<i>Nombre de sujets qui ont cessé de participer pendant le traitement</i>		<i>N =</i>	<i>N =</i>
<i>Nombre de sujets qui ont terminé le traitement</i>		<i>N =</i>	<i>N =</i>
TAUX D'ABANDON			
<i>Échec de la mise en œuvre? (c.-à-d. taux d'abandon de 50 % ou plus). Donner les raisons.</i>		<i>Non</i>	<i>OUI</i>
<i>Cotes des biais = 0</i>			
Cotes des biais			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Beaucoup de biais : Échec de la mise en œuvre</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Abandon du programme

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cotes révisées des biais

0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
--	---	---

Sens révisé des biais

? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
---	--	--------------------------------------	--

Raison(s) de la cote :

11. Analyse des sujets retenus au début de l'étude

Concept : Le concept général consiste à déterminer comment le ou les chercheurs ont traité l'abandon du programme lorsqu'ils ont estimé l'efficacité du traitement. Dans toutes les études, les chercheurs devraient utiliser une analyse des sujets retenus au début de l'étude pour calculer les ampleurs de l'effet. Toutefois, à mesure que les taux d'attrition augmentent, le nombre de biais de l'analyse des sujets retenus au début de l'étude s'accroît également. L'objectif de cet élément consiste à évaluer l'ampleur et le sens des biais qui résultent de la façon dont le chercheur traite l'attrition lorsqu'il calcule les effets du traitement.

Indicateurs : Il y a deux facteurs importants à prendre en considération. Un indicateur est l'importance de l'attrition évaluée dans l'élément précédent. Le deuxième indicateur consiste à déterminer si le ou les chercheurs ont utilisé une analyse des sujets retenus au début de l'étude ou si les abandons ont été traités d'une autre façon (p. ex. exclus de l'analyse ou inclus dans le groupe témoin). Des biais peuvent être introduits lorsqu'il y a plus qu'un nombre minimal d'« abandons » et que les « abandons » sont exclus de l'analyse. Les biais sont réduits au minimum lorsqu'il y a un petit nombre d'abandons et qu'une analyse des sujets retenus au début de l'étude est effectuée.

Cote	Description
0	<i>Cote de 0 attribuée à l'abandon du programme (c.-à-d. que le taux d'abandon dépasse 41 %), qu'une analyse des sujets retenus au début de l'étude soit utilisée ou non.</i> OU <i>Cote de 1 attribuée à l'abandon du programme (c.-à-d. que le taux d'abandon varie de 21 % à 40 %), et les « abandons » sont exclus de l'estimation de l'ampleur de l'effet.</i>
1	<i>Cote de 1 attribuée à l'abandon du programme (c.-à-d. que le taux d'abandon varie de 21 % à 40 %), et les « abandons » sont inclus (analyse des sujets retenus au début de l'étude) dans l'analyse.</i> OU <i>Cote de 2 attribuée à l'abandon du programme (c.-à-d. que le taux d'abandon est de 20 % ou moins, et les « abandons » sont exclus de l'ampleur de l'effet.</i>
2	<i>Cote de 2 attribuée à l'abandon du programme (c.-à-d. que le taux d'abandon est de 20 % ou moins), et l'ampleur de l'effet est calculée au moyen d'une analyse des sujets retenus au début de l'étude.</i>

11. Analyse des sujets retenus au début de l'étude

Cet élément porte sur la façon dont le ou les chercheurs ont traité l'abandon du programme lorsqu'ils ont estimé l'efficacité du traitement. Des biais peuvent être introduits lorsqu'il y a plus qu'un nombre minimal d'« abandons » et que les « abandons » sont exclus de l'analyse. Il est important d'enregistrer les renseignements sur l'abandon du programme (p. ex. risque, proportion par rapport à l'ensemble du groupe, taux de récurrence, raisons de l'abandon) et de déterminer si ces « abandons » ont été inclus ou non dans le groupe expérimental dans l'estimation de l'efficacité du traitement.

<i>Information sur l'analyse des sujets retenus au début de l'étude provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Analyse des sujets retenus au début de l'étude

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cotes révisées des biais

0	1	2
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>

Sens révisé des biais

?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>

Raison(s) de la cote :

12. Attrition pendant le suivi

Concept : Le concept général est celui des biais introduits par l'abandon des sujets pendant le suivi. En général, les sujets sont perdus à cause de données manquantes ou insuffisantes sur le suivi. Dans les études de cohortes et les études rétrospectives, il est essentiel d'être conscient de l'attrition éventuelle pendant le suivi en raison de la suppression des dossiers plus vieux et inactifs, des méthodes différentes de conservation des dossiers et de l'appariement inversé des dossiers existants. Des biais systématiques et différentiels peuvent être introduits simplement à cause de la disponibilité différente de dossiers entre les groupes, les périodes et les administrations. L'objectif de cet élément est d'évaluer l'ampleur et le sens des biais qui résultent de l'attrition pendant le suivi.

Indicateurs : Les indicateurs éventuels de biais réduits seraient les cas où il y aurait des renseignements adéquats sur presque tous les participants, y compris ceux qui ont abandonné pendant le suivi. Les indicateurs qui pourraient accroître les biais sont les taux élevés d'attrition dans les renseignements sur le suivi, les données manquantes ou inadéquates sur le suivi pour les abandons ou si les sujets ont été choisis pour le groupe expérimental ou témoin en fonction de la disponibilité des dossiers.

Cote	Description
0	<i>Les renseignements sur le suivi sont disponibles pour moins de 70 % des sujets de l'un ou l'autre groupe.</i> OU <i>La différence dans le taux d'attrition entre les groupes est supérieure à 10 %.</i>
1	<i>Les renseignements sur le suivi sont disponibles pour au moins 70 % des sujets de chaque groupe.</i> ET <i>La différence dans le taux d'attrition entre les groupes est de 10 % ou moins.</i>
2	<i>Des renseignements sur le suivi sont disponibles pour au moins 90 % des sujets de chaque groupe.</i>

12. Attrition pendant le suivi

Cet élément porte sur l'attrition pendant le suivi. Il y a deux sources possibles de biais qui peuvent être introduits. Premièrement, il y a les renseignements manquants sur le suivi en raison de dossiers inadéquats ou de difficultés à trouver les dossiers des participants. Deuxièmement (problème plus grave), il y a les pratiques ou politiques différentes de conservation des dossiers (p. ex. les anciens dossiers sont détruits dans certains cas) et/ou processus d'extraction (p. ex. données non disponibles sur les participants qui se sont retirés de l'étude).

<i>Information sur les abandons pendant le suivi provenant de l'étude :</i>			
<i>Pourcentage des sujets du groupe expérimental avec suivi :</i>		<i>Pourcentage des sujets du groupe témoin avec suivi :</i>	
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : attrition pendant le suivi

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cotes révisées des biais

0	1	2
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>

Sens révisé des biais

?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>

Raison(s) de la cote :

V. Équivalence des groupes

Étant donné que la principale menace à la validité est le manque d'équivalence entre les groupes, il importe pour les chercheurs d'examiner l'ampleur des différences préexistantes entre le groupe expérimental et le groupe témoin. Ce ne sont pas toutes les différences qui sont importantes; les différences qui comptent sont celles qui ont trait à la possibilité de récidive. Le résultat qui nous intéresse en fin de compte est la récidive, mais il est impossible de mesurer la « récidive » au moment de l'évaluation initiale. Il s'agit d'un problème spécial de la recherche sur les résultats du traitement des délinquants sexuels. Pour bien d'autres troubles (p. ex. dépression, déficience physique), il est possible d'apparier de façon convaincante les groupes au cours du traitement préliminaire. Sans accès direct au problème en question, toutefois, les chercheurs sont obligés de s'en remettre à des « variables substitutives » pour estimer l'équivalence (Shadish et coll., 2002).

Les chercheurs ne peuvent jamais prouver l'équivalence préexistante entre le groupe expérimental et le groupe témoin lorsqu'il y a une corrélation imparfaite entre les variables substitutives et le résultat. Comme nous ne pouvons pas prédire la récidive avec une exactitude parfaite, il serait toujours possible d'attribuer les différences observées dans la récidive à des différences non observées dans le potentiel de risque au moment de l'évaluation initiale. La confiance dans l'équivalence des groupes augmente lorsque a) les méthodes d'affectation des sujets réduisent au minimum la probabilité de différences systématiques; b) les variables de contrôle expliquent une grande partie de la variance du résultat.

Pour procéder à un examen minutieux de l'équivalence avant le traitement, il faut effectuer une évaluation complète du risque de récidive de chaque délinquant dans le groupe expérimental et le groupe témoin. Ces évaluations du risque sont plus crédibles lorsqu'elles sont effectuées par des évaluateurs compétents (qui ne connaissent pas le résultat) au moyen des meilleures approches disponibles. Même s'il y a un certain débat sur la meilleure façon d'évaluer le risque de récidive, la plupart des évaluateurs utilisent un ou plusieurs outils d'évaluation du risque actuariels tout en prenant en considération les facteurs extérieurs au plan actuariel. Une autre méthode consiste à tenir compte d'une liste de facteurs de risque validée empiriquement. Quelle que soit la méthode utilisée, il importe de montrer que la méthode d'évaluation du risque que pose le délinquant est valide dans l'échantillon de recherche. Plus l'association avec la récidive est étroite, plus on peut avoir confiance que la plupart des facteurs de risque pertinents ont été pris en considération. Les mesures de contrôle optimales varieront selon les critères des résultats utilisés. Les facteurs associés à la récidive sexuelle ne sont pas identiques aux facteurs associés à la récidive violente ou criminelle. Les chercheurs doivent insister sur les variables de contrôle associées le plus au résultat en question.

Même s'il y a eu des progrès considérables dans l'évaluation du risque que présentent les délinquants sexuels adultes, on en sait beaucoup moins sur l'évaluation du risque que posent les délinquantes sexuelles ou les jeunes délinquants sexuels. Par conséquent, les chercheurs qui effectuent des études sur les résultats du traitement auprès de ces populations auraient peu confiance dans leur choix des variables de contrôle.

Comme l'équivalence des groupes est essentielle à l'obtention d'une estimation non biaisée de l'efficacité du traitement, les Lignes directrices du CCDCR comptent trois éléments pour évaluer ce facteur. Le premier porte sur l'équivalence des groupes a priori, fondée sur la conception de l'étude et les procédures d'affectation des sujets, et sert à déterminer si celle-ci introduit un biais. Un examen minutieux des procédures utilisées pour recruter les participants et les retenir dans les deux groupes peut permettre d'établir des attentes quant à l'équivalence des groupes. Par exemple, dans le cas de procédures d'affectation aléatoire, nous pouvons nous attendre à ce qu'il n'y ait à peu près pas de différences a priori entre les groupes. Tout comme pour tous les éléments qui évaluent les biais, il y a une évaluation de l'ampleur et du sens des biais. Étant donné l'importance de déterminer une équivalence des groupes a priori, d'autres questions sont présentées pour aider à évaluer les modèles d'affectation aléatoire, de cohorte et de gammes de risques.

Le deuxième élément porte sur le caractère adéquat de la recherche des différences préexistantes entre les groupes et évalue la confiance du codeur que la recherche était suffisante pour permettre de déterminer les différences préexistantes importantes. Les renseignements pertinents comprennent les mesures utilisées et les autres renseignements recueillis pour évaluer l'équivalence des groupes. Cet élément ne tient pas compte des résultats de la comparaison; il évalue plutôt les outils et les méthodes utilisés pour comparer les groupes. Par exemple, quels que soient les résultats, on peut faire davantage confiance à la comparaison des groupes fondée sur des outils d'évaluation du risque validés qu'aux mesures de la personnalité qui n'ont à peu près pas de rapport avec le risque de récurrence sexuelle.

Le troisième élément, Constatations sur l'équivalence des groupes, porte sur les résultats de cette recherche des différences préexistantes et évalue les biais potentiels de la similarité ou des différences entre les groupes. Pour cet élément, le codeur examine les données qui ont trait à la possibilité de récurrence qui ont été présentées pour démontrer l'équivalence des groupes. Lorsqu'il y a une différence entre les groupes, c'est cette différence qui peut introduire un biais et menacer la validité des résultats. Cet élément sert à évaluer les données présentées et non la ou les procédures statistiques utilisées pour tenir compte des différences (un élément distinct évalue l'utilisation des procédures de contrôle statistique). Les codeurs doivent évaluer l'ampleur et le sens des biais d'après les données présentées sur l'équivalence des groupes.

13. Équivalence des groupes a priori

Concept : Le concept général est la mesure dans laquelle on s'attendrait à ce que le groupe expérimental et le groupe témoin soient équivalents d'après le modèle de l'étude et les procédures utilisées pour affecter les participants à chaque groupe. La preuve de l'équivalence des groupes n'est pas évaluée dans cet élément (il s'agit d'un élément distinct). Les facteurs à prendre en considération sont le genre de modèle et les procédures d'affectation (p. ex. affectation aléatoire). Idéalement, les procédures d'affectation permettent de s'attendre à ce que les deux groupes soient équivalents. Cela se produit dans les modèles d'affectation aléatoire bien mis en oeuvre où le processus de randomisation est un processus où a) la séquence d'affectation est imprévisible; b) ceux qui procèdent à l'affectation réelle ne connaissent pas les affectations effectuées, connues sous le nom de dissimulation de l'affectation (Altman et coll., 2001). Selon la méta-analyse de Schultz, Chalmers, Hayes et Altman (1995) sur les essais cliniques randomisés, les études où la dissimulation de l'affectation est inadéquate révèlent des effets du traitement beaucoup plus importants que celles où la dissimulation est adéquate. D'autre part, beaucoup d'études n'utilisent pas l'affectation aléatoire. Dans ces études, il importe de comparer les critères de sélection pour le groupe expérimental et les critères d'inclusion dans le groupe témoin. De cette manière, on peut évaluer les différences potentielles entre les groupes fondées strictement sur les procédures d'assignation ou d'affectation de l'étude.

Indicateurs : Les indicateurs qui pourraient réduire les biais sont l'affectation aléatoire, la dissimulation de l'affectation et la probabilité égale que tout sujet soit affecté à l'un ou l'autre groupe. Les indicateurs qui pourraient accroître les biais sont les cas où les sujets choisis pour un groupe n'ont à peu près pas de chance d'être dans l'autre groupe.

Cote	Description
0	<i>Les procédures de conception et d'allocation sont susceptibles de se traduire par le placement systématique des délinquants présentant des caractéristiques différentes sur le plan du risque dans des groupes différents.</i>
1	<p data-bbox="316 1371 1443 1493"><i>Les procédures de conception et d'allocation sont susceptibles de se traduire par des différences systématiques entre les groupes, mais les caractéristiques pour lesquelles ils diffèrent ne sont pas liées directement au risque (p. ex. langue, lieu).</i></p> <p data-bbox="852 1493 911 1528" style="text-align: center;">OU</p> <p data-bbox="316 1539 1443 1661"><i>Il est possible que les participants qui présentent certaines caractéristiques pertinentes au risque (p. ex. risque, besoin) soient plus susceptible d'être dans un groupe que dans l'autre.</i></p>
2	<i>En raison des procédures de conception et d'affectation, pour tous les participants, la probabilité de faire partie du groupe expérimental ou du groupe témoin est égale.</i>

Pour aider à déterminer les différences a priori, des listes de contrôle distinctes sont fournies pour trois genres d'études : a) affectation aléatoire; b) modèles de gammes de risques/normes; c) modèles de cohorte. Il n'est pas prévu de faire de ces listes de contrôle des éléments distincts; elles font plutôt ressortir les préoccupations particulières associées aux différentes formes d'affectation des sujets. S'ils évaluent l'un de ces trois modèles, les évaluateurs doivent remplir la liste de contrôle appropriée avant d'établir les cotes d'équivalence des groupes a priori.

Modèles d'essais randomisés – procédures de randomisation

Concept : L'affectation aléatoire bien exécutée réduit la possibilité de différences préexistantes entre les groupes. Dans la mesure où le processus de randomisation échoue, la probabilité de biais s'accroît. Par conséquent, il importe de prendre en considération l'intégrité du processus d'affectation aléatoire.

Facteurs à prendre en considération : Il y a trois éléments précis à examiner. Premièrement, le processus de randomisation doit être réellement aléatoire, chaque sujet ayant une chance égale d'être affecté à chaque groupe. L'affectation à tour de rôle, l'affectation fondée sur le jour de la semaine, la date de naissance ou le numéro d'identification du dossier sont des processus qui ne sont pas vraiment aléatoires. Ces procédures sont assujetties à l'influence de l'expérimentateur et peuvent modifier l'équivalence des groupes. Deuxièmement, l'expérimentateur ne doit pas connaître l'affectation à un groupe au moment de l'affectation; autrement l'affectation peut être biaisée. Parmi les exemples de procédures de dissimulation de l'affectation figurent l'utilisation d'enveloppes opaques scellées numérotées séquentiellement ou le recours à un bureau central qui ne connaît pas les caractéristiques de l'affectation des sujets aux groupes. Troisièmement, les procédures de randomisation peuvent tomber en panne ou changer pendant l'étude par suite d'événements imprévus (p. ex. une modification des politiques, des pressions systémiques exercées pour qu'on traite certains individus), qui entraînent des différences préexistantes entre les groupes.

<i>L'affectation est aléatoire?</i>	
<i>Les procédures d'affectation ont abouti à une affectation prévisible ou prévue à un groupe, comme l'affectation à tour de rôle ou selon le jour de la semaine. La probabilité d'affectation des sujets à chaque groupe n'était pas aléatoire.</i>	0
<i>Les procédures d'affectation décrites ont abouti à une affectation réellement aléatoire, comme l'utilisation de numéros au hasard.</i>	1
<i>Dissimulation de l'affectation?</i>	
<i>L'affectation n'était pas dissimulée. L'expérimentateur connaissait l'affectation aux groupes ou pouvait l'avoir prévue et des biais inconscients peuvent avoir influé sur l'affectation des sujets, leur admissibilité ou les efforts visant à faire participer les sujets.</i>	0
<i>Les procédures ont permis de dissimuler l'affectation à l'expérimentateur grâce à l'utilisation d'enveloppes opaques scellées numérotées séquentiellement, ou un bureau central a procédé à l'affectation sans connaître les caractéristiques des sujets.</i>	1
<i>Procédures d'affectation uniformes?</i>	
<i>Les procédures d'affectation ont été modifiées pendant l'expérience.</i>	0
<i>Les procédures d'affectation ont été suivies de manière uniforme sans changement pendant la durée de l'étude.</i>	1

N. B. : Pour les modèles d'affectation aléatoire, les facteurs susmentionnés aideront à évaluer l'équivalence des groupes a priori.

Modèles de gammes de risques/normes – validité des normes

Concept : Le concept général est la mesure dans laquelle il y a des différences préexistantes entre le groupe expérimental et les normes ou gammes de risques auxquelles il est comparé. Dans le cas de ces modèles, il y a toujours une attente a priori de certaines différences préexistantes entre les groupes et l'on s'attend donc à un certain biais. Il est impossible de démontrer l'équivalence des groupes pour tous les facteurs.

Facteurs à prendre en considération : Il y a deux questions à prendre en considération. La première est l'exactitude de la mesure du risque. Une bonne mesure devrait avoir une exactitude prédictive élevée, et il ne devrait à peu près pas manquer de données. Il est essentiel, lorsqu'on établit une comparaison avec les gammes de risques ou les normes, que l'évaluation du risque pour le groupe expérimental soit aussi complète que possible. Deuxièmement, l'équivalence de la population normative avec l'échantillon expérimental doit être examinée pour permettre de déterminer les effets possibles des cohortes ou de l'administration. Par exemple, les taux de récidive du groupe normatif (c.-à-d. groupe témoin) ou du groupe expérimental peuvent être influencés par les facteurs relatifs à la cohorte ou à l'administration (p. ex. comparer des normes récentes à une cohorte de délinquants traités des années 1980).

<i>Précision de l'évaluation?</i>	
<i>La fiabilité de l'évaluation n'a pas été évaluée (ou est inadéquate) et(ou) les évaluations étaient incomplètes en raison de données manquantes.</i>	0
<i>Ainsi concernant l'évaluation du risque actuarielle du groupe expérimental, elle ne manque pas de données et une fiabilité élevée des évaluateurs a été établie.</i>	1
<i>Exactitude des normes?</i>	
<i>Les normes (gammes de risques) sont fondées sur certaines données empiriques (p. ex. l'échantillon normatif est fondé sur moins de 1 000 délinquants et(ou) l'échantillon normatif ne compte pas de délinquants de ce pays), mais la confiance dans les normes n'est pas très grande.</i>	0
<i>Les normes (gammes de risques) sont fondées sur des données empiriques solides, elles ont fait l'objet d'une validation croisée et les caractéristiques du groupe normatif sont semblables à celles du groupe expérimental.</i>	1
<i>Effets de la cohorte ou de l'administration?</i>	
<i>Il y a des différences de cohorte et(ou) des différences possibles sur le plan de l'administration entre le groupe normatif et le groupe expérimental.</i>	0
<i>Le groupe expérimental est susceptible d'être représentatif du groupe normatif sur le plan de la cohorte (temps) et de l'administration (emplacement géographique).</i>	1

N. B. : Pour les modèles de gammes de risques/normes, les facteurs susmentionnés aideront à évaluer l'équivalence des groupes a priori.

Modèles des cohortes – effets de cohorte

Concept : Le concept général est l'influence de la cohorte sur les résultats de l'étude. Comme il s'agit d'échantillons de délinquants de différentes périodes, il pourrait s'être produit des changements systémiques au fil du temps et(ou) des changements dans les caractéristiques des délinquants qui influent sur les taux de récidive. Compte tenu de ces modèles, les chercheurs doivent accorder une attention particulière aux effets possibles de cohorte.

Facteurs à prendre en considération : Il y a deux facteurs à prendre en considération. Premièrement, il faut vérifier les effets systémiques de cohorte sur la récidive. En déterminant si la variation de la date de mise en liberté est associée à des différences de la récidive (au moyen de l'échantillon de l'étude et de la population dans laquelle il a été tiré), le chercheur peut évaluer les effets systémiques de cohorte. Deuxièmement, il faut aussi examiner les effets possibles de cohorte sur l'évolution du profil des délinquants. Par exemple, il se peut que les caractéristiques d'une cohorte soient différentes de celles d'une autre. Comme il est difficile d'évaluer tous les facteurs possibles associés à la récidive et que les populations et les systèmes changent et évoluent au fil du temps, il est essentiel que les modèles de cohorte portent sur les changements possibles des populations à l'étude et, dans la mesure du possible, qu'ils tiennent compte de cette variation systémique.

<i>Effets de cohorte – Changements systémiques?</i>	
<i>La recherche n'a pas porté sur l'évaluation des changements systémiques dans les cohortes.</i> OU <i>Des effets de cohorte ont été observés après l'évaluation des changements systémiques dans l'échantillon de l'étude et(ou) la population où il a été tiré.</i>	0
<i>Le chercheur a évalué les changements systémiques dans les cohortes en examinant l'échantillon de l'étude et la population où il a été tiré. Des effets de cohorte n'ont pas été observés après une recherche approfondie.</i>	1
<i>Effets de cohorte – Changements des délinquants?</i>	
<i>La recherche n'a pas porté sur l'évaluation des changements dans les caractéristiques des délinquants.</i> OU <i>Des effets de cohorte ont été observés après l'évaluation des caractéristiques des délinquants dans l'échantillon de l'étude et(ou) la population où il a été tiré.</i>	0
<i>Le chercheur a évalué les changements dans les caractéristiques des délinquants dans l'échantillon de l'étude et la population où il a été tiré. Des effets de cohorte n'ont pas été observés après une recherche approfondie.</i>	1

N. B. : Pour les modèles de cohorte, les facteurs susmentionnés aideront à évaluer l'équivalence des groupes a priori. Il faut examiner la façon dont le chercheur a tenu compte des effets de cohorte dans l'élément Efficacité des procédures statistiques de contrôle des biais.

13. Équivalence des groupes a priori

Cet élément porte sur les attentes a priori à l'égard de l'équivalence des groupes fondée sur les caractéristiques de l'étude et les critères de sélection des sujets. Par exemple, on ne s'attend pas à beaucoup de différences a priori dans les grandes études à affectation aléatoire. Toutefois, si les critères d'inclusion dans un groupe sont différents de ceux de l'autre ou des autres groupes, il y a certaines attentes a priori que les groupes sont différents en ce qui concerne les variables relatives au résultat. Pour cet élément, ce sont les attentes concernant l'équivalence des groupes, et non les preuves empiriques, qui sont évaluées.

<i>Information sur l'équivalence des groupes a priori provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : Équivalence des groupes a priori

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cotes révisées des biais

0	1	2
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>

Sens révisé des biais

?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>

Raison(s) de la cote :

14. Qualité de la recherche de différences préexistantes

Concept : Le concept général consiste à déterminer comment et avec quelles mesures le ou les chercheurs tentent d'évaluer l'équivalence des groupes avant le traitement. En d'autres termes, quelle est la qualité du travail qu'ont effectué le ou les chercheurs pour examiner la mesure dans laquelle les différences préexistantes entre les groupes pourraient expliquer les différences observées dans le résultat. Cet élément vise à déterminer la confiance du codeur que l'approche, la méthode, les outils et(ou) les renseignements qu'ont utilisés l'auteur ou les auteurs pour comparer les différences préexistantes étaient suffisants pour permettre de déterminer les différences possibles et importantes.

Indicateurs : Les indicateurs qui pourraient accroître la confiance que la recherche était suffisante seraient l'utilisation d'outils d'évaluation du risque validés, les renseignements recueillis sur d'autres facteurs relatifs à la récidive, la fiabilité de l'information et la quantité de renseignements manquants. Cet élément évalue la confiance dans les renseignements recueillis pour examiner les différences préexistantes entre les groupes, et non les résultats réels de cette comparaison (ce qui est fait au prochain élément).

Cote	Description
0	<i>Aucun outil d'évaluation du risque validé n'a été utilisé. Les groupes peuvent avoir été évalués en fonction de nombreux facteurs, mais peu, le cas échéant, avaient trait à la possibilité de récidive. Il manquait beaucoup de renseignements.</i>
1	<i>Les groupes ont été évalués en fonction d'une mesure du risque validée ou d'un nombre considérable de facteurs liés à la possibilité de récidive qui ont indiqué une exactitude prédictive au moins modérée.</i>
2	<i>Les groupes ont été évalués en fonction d'au moins une mesure du risque validée, de renseignements supplémentaires recueillis sur d'autres facteurs liés à la possibilité de récidive extérieure aux mesures du risque, et la fiabilité des renseignements est démontrée. Il manque peu de renseignements, s'il en manque.</i>

14. Qualité de la recherche de différences préexistantes

Cet élément sert à examiner comment l'auteur ou les auteurs ont cherché à évaluer les différences préexistantes entre les groupes. L'évaluateur tente de déterminer le niveau de confiance que ces mesures, renseignements et(ou) facteurs évalués pourront détecter avec précision les différences préexistantes possibles. La fiabilité et la quantité de renseignements manquants influent sur la confiance. Cet élément évalue la confiance dans les renseignements recueillis pour examiner les différences préexistantes entre les groupes et non les résultats réels de cette comparaison (qui sont évalués par le prochain élément).

<i>Information sur la recherche des différences préexistantes provenant de l'étude :</i>			
Cote de confiance			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Peu de confiance que les procédures, les mesures et(ou) les renseignements utilisés sont liés empiriquement à la variable du résultat</i>	1 <i>Certaine confiance que la recherche des différences préexistantes était exhaustive, fiable et valide</i>	2 <i>Grande confiance que la recherche des différences préexistantes était exhaustive, fiable et valide</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : qualité de la recherche de différences préexistantes

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance que les procédures, les mesures et(ou) les renseignements utilisés sont liés empiriquement à la variable du résultat

1

Certaine confiance que la recherche des différences préexistantes était exhaustive, fiable et valide

2

Grande confiance que la recherche des différences préexistantes était exhaustive, fiable et valide

Raison(s) de la cote :

15. Constatations sur l'équivalence des groupes

Concept : Le concept général est la mesure dans laquelle les données corroborent l'équivalence du groupe expérimental et du groupe témoin. Cet élément vise à évaluer l'ampleur et le sens des biais d'après les caractéristiques préexistantes des groupes. Dans cet élément, ce sont les données qui sont évaluées, quelle que soit la façon dont les chercheurs ont analysé leurs données. Pour permettre d'évaluer cet élément, l'étude doit avoir présenté des données sur les deux groupes qui sont associés à la possibilité de récidive. Si les renseignements sont insuffisants pour permettre de déterminer l'étendue de l'équivalence entre les groupes, cet élément n'est pas évalué. En général, une *cote de confiance* de zéro pour l'élément précédent suppose que cet élément ne peut pas être évalué. Toutefois, s'il y a des différences claires concernant les variables du risque, cet élément peut être évalué lorsque la recherche des différences préexistantes était minimale.

Indicateurs : Les indicateurs sont fondés sur l'ampleur des différences entre le groupe expérimental et le groupe témoin en ce qui concerne le potentiel de risque. Les groupes sont considérés comme « essentiellement égaux » lorsque le niveau de probabilité des différences observées est supérieur à 0,40. Les groupes sont « très différents » lorsque la probabilité des différences observées est inférieure à 0,05 ou que l'ampleur de l'effet entre les groupes est assez importante (c.-à-d. que $d \geq 0,50$). Lorsque les deux groupes sont essentiellement égaux en ce qui concerne les mesures du risque et les autres renseignements relatifs à la possibilité de récidive, on s'attend à ce que les différences préexistantes entraînent un biais négligeable de l'ampleur de l'effet. D'autre part, on s'attendrait à un biais lorsque les groupes sont très différents en ce qui concerne la ou les mesures du risque ou s'il y a des différences statistiquement significatives pour les variables prédictives de la récidive.

Cote	Description
0	<p><i>Les groupes sont très différents en ce qui concerne la ou les mesures du risque ($p < 0,05$)</i></p> <p style="text-align: center;">OU</p> <p><i>Il n'y a aucune mesure du risque et il y a des différences statistiquement significatives concernant certaines des variables relatives à la possibilité de récidive.</i></p>
1	<p><i>Les groupes ne sont pas très différents en ce qui concerne la ou les mesures, et il y a des différences entre les groupes en ce qui concerne les variables pour lesquelles le lien avec le risque est faible ou indirect.</i></p> <p style="text-align: center;">OU</p> <p><i>Les groupes ne sont pas très différents en ce qui concerne la ou les mesures, et il y a une certaine incertitude dans les résultats en raison de la petite taille de l'échantillon ou de la signification statistique marginale ($0,05 < p < 0,40$).</i></p> <p style="text-align: center;">OU</p> <p><i>Une mesure du risque n'a pas été utilisée, et il n'y a pas différences significatives pour une gamme de facteurs de risque validés.</i></p>
2	<p><i>Les groupes sont essentiellement égaux en ce qui concerne la ou les mesures du risque ($p \geq 0,40$).</i></p> <p style="text-align: center;">ET</p> <p><i>Les groupes sont essentiellement égaux pour presque toutes les autres variables externes à la ou aux mesures du risque.</i></p>

15. Constatations sur l'équivalence des groupes

Cet élément porte sur les données présentées dans l'étude qui visent à démontrer l'équivalence des groupes. L'évaluateur doit évaluer ces résultats pour déterminer l'équivalence des groupes et la mesure dans laquelle les caractéristiques préexistantes des groupes peuvent introduire des biais dans l'estimation de l'efficacité du traitement.

<i>Information et résultats concernant l'équivalence des groupes provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : données sur l'équivalence des groupes			
<i>Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?</i>			
<i>Si de nouveaux renseignements ont été obtenus, précisez.</i>			
Cotes révisées des biais			
0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>	
Sens révisé des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

VI. Variables des résultats

Les variables des résultats revêtent une importance cruciale dans l'évaluation de l'efficacité du traitement. Toutes les études sur les résultats du traitement des délinquants sexuels devraient évaluer la récidive sexuelle, mais il est également utile d'examiner la catégorie générale de la récidive « grave », qui comprend les infractions sexuelles et les infractions avec violence. Les infractions avec violence suscitent de vives préoccupations chez le public, et il n'est pas inhabituel que les infractions sexuelles entraînent des condamnations pour infractions avec violence non sexuelles (p. ex. séquestration). Nous recommandons aussi que les évaluations mesurent la récidive générale (tous les genres).

Il y a trois aspects des renseignements sur la récidive à prendre en considération : a) la suffisance de la durée du suivi pour la détection de la récidive; b) les sources d'information qui fournissent des données sur la ou les variables des résultats; c) l'équivalence des renseignements sur le suivi entre le groupe expérimental et le groupe témoin. La durée du suivi et les sources d'information sur la récidive jouent un rôle clé dans le degré de confiance qu'on peut accorder aux résultats. Les différences dans les renseignements sur la récidive introduisent des biais.

Une raison explique la difficulté de réaliser des études sur les résultats du traitement des délinquants sexuels : le résultat en question ultime – la récidive sexuelle – est un événement peu fréquent. Les taux de récidive sexuelle observés sont d'environ 10 à 15 % après cinq ans et de 20 % après 10 ans (Harris et Hanson, 2004). De plus, il y a souvent un décalage important entre la perpétration de l'infraction et la découverte de l'infraction (si jamais elle l'est) et l'enregistrement de celle-ci dans les bases de données accessibles aux chercheurs. La plupart des chercheurs proposent de laisser passer des périodes de suivi minimales avant de chercher à évaluer les taux de récidive sexuelle. La récidive observée pendant les périodes de suivi très courtes peut résulter davantage des caractéristiques arbitraires du système de justice pénale que des caractéristiques des délinquants ou du traitement qu'ils ont reçu. On ne s'entend pas sur la période de suivi minimale nécessaire pour qu'une étude soit crédible, mais on a proposé des périodes de trois à cinq ans.

Des taux de base faibles constituent moins un problème lorsque les chercheurs s'intéressent aux critères plus généraux des résultats qu'à la seule récidive sexuelle. Si la récidive était le principal critère des résultats, il serait possible d'obtenir un échantillon adéquat de récidivistes au moyen de tailles plus petites et de périodes plus courtes que celles proposées dans les présentes lignes directrices.

La validité et la fiabilité des données sur la récidive jouent également un rôle dans le niveau de confiance dans les résultats de l'étude. On peut accorder une grande confiance aux résultats lorsque la source d'information est crédible (p. ex. dossiers officiels), que la recherche est approfondie (p. ex. il ne manque pas de cas pendant la collecte des données) et que l'étude est fondée sur de multiples sources d'information (p. ex. casiers judiciaires nationaux et des États),

rapports de police et dossiers de probation). Plus la ou les sources d'information sont valides et fiables, plus on peut avoir confiance dans les résultats.

Il est important que la période de suivi soit la même pour le groupe expérimental et le groupe témoin. On pourrait s'attendre à de telles différences dans les modèles de cohorte, mais toutes les études doivent tenir compte de l'équivalence des périodes de suivi. Les procédures de contrôle statistique utilisées en raison de l'inégalité de la période de suivi ont une utilité limitée. Si les différences dans les périodes de suivi sont importantes (p. ex. $p < 0,05$), les chercheurs ont un problème évident; toutefois, lorsque la taille des échantillons est petite, les différences devraient être considérables pour être détectées. Il est prudent, pour toutes les études, de tenir compte des différences possibles dans les périodes de suivi au moyen a) de périodes de suivi fixes (p ex. cinq ans), ou b) d'une analyse de survie. Les chercheurs doivent cependant noter qu'une analyse de survie pourrait encore introduire un biais lorsque l'hypothèse des hasards proportionnels n'est pas vérifiée ou que la forme de la courbe de survie change au fil du temps.

Les Lignes directrices du CCDCR évaluent trois aspects des données sur les résultats. Un élément évalue la confiance dans la durée du suivi pour détecter la récidive sexuelle. Les renseignements pertinents comprennent la durée et l'étendue du suivi pour chaque groupe. La confiance a trait à la suffisance de la durée du suivi pour la détection de la récidive.

Le deuxième élément évalue la confiance dans la validité et la fiabilité des renseignements sur la récidive. Les renseignements pertinents comprennent la description du caractère exhaustif des dossiers, la possibilité que de multiples sources d'information aient servi à détecter la récidive et la crédibilité des sources d'information. Il est également souhaitable de vérifier le nombre de délinquants qui n'ont pas pu faire l'objet d'un suivi en raison de leur incarcération prolongée, de leur expulsion ou de leur décès. La confiance a trait à la validité et à la fiabilité des données sur la récidive.

Le troisième et dernier élément de la section évalue les biais qui peuvent être introduits par l'équivalence du suivi. Les renseignements pertinents comprennent toute différence dans la durée et l'étendue du suivi, les sources d'information sur la récidive ou les autres différences importantes entre les groupes dans la collecte de renseignements sur la récidive. Il faut établir des cotes de l'ampleur et du sens des biais.

16. Durée du suivi

Concept : Le concept général consiste à déterminer si la durée du suivi pour l'échantillon de l'étude est suffisamment longue pour permettre de détecter la récurrence. Cet élément vise à évaluer la confiance que la durée du suivi est suffisante pour fournir des renseignements exacts sur la récurrence.

Indicateurs : Pour cet élément, on a besoin de données sur la période de suivi moyenne pour l'échantillon de l'étude. La confiance s'accroît à mesure que la période de suivi moyenne s'allonge. Les périodes de suivi étaient fondées sur l'utilisation de la récurrence sexuelle ou la récurrence violente comme critère de résultat. Si des formes plus fréquentes de récurrence étaient utilisées (p. ex. toute nouvelle condamnation, manquement aux conditions de la libération conditionnelle), il serait possible d'obtenir un nombre suffisant de récidivistes dont la période de suivi serait plus courte; par conséquent, les études faisant appel aux formes les plus courantes de récurrence peuvent recevoir une cote de « 2 – grande confiance », d'après les périodes de suivi de moins de cinq ans.

Cote	Description
0	<i>Période de suivi moyenne de moins de 3 ans (36 mois) pour l'échantillon de l'étude.</i>
1	<i>Période de suivi moyenne de 3 ans (36 mois), mais de moins de 5 ans (60 mois).</i>
2	<i>Période de suivi moyenne de 5 ans (60 mois) ou plus.</i>

16. Durée du suivi

Cet élément porte sur la durée du suivi et le niveau de confiance suffisant pour détecter la récurrence. Les renseignements sur la durée du suivi de tout l'échantillon de l'étude et chaque groupe doivent être enregistrés.

<i>Information sur la durée du suivi provenant de l'étude :</i>			
			<i>Suivi moyen (gamme)</i>
<i>Groupe expérimental</i>			
<i>Groupe témoin</i>			
<i>Échantillon de l'étude</i>			
Cote de confiance			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Peu de confiance que la durée du suivi est assez longue pour permettre de détecter la récurrence</i>	<i>Certaine confiance que la durée du suivi est assez longue pour permettre de détecter la récurrence</i>	<i>Grande confiance que la durée du suivi est assez longue pour permettre de détecter la récurrence</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : durée du suivi

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance que la durée du suivi est assez longue pour permettre de détecter la récidive

1

Certaine confiance que la durée du suivi est assez longue pour permettre de détecter la récidive

2

Grande confiance que la durée du suivi est assez longue pour permettre de détecter la récidive

Raison(s) de la cote :

17. Validité et fiabilité des renseignements sur la récidive

Concept : Les renseignements pertinents comprennent l'exhaustivité des dossiers, la possibilité que des sources d'information multiples aient été utilisées pour détecter la récidive et la crédibilité des sources d'information. La confiance a trait à la validité et à la fiabilité des données.

Indicateurs : La confiance est plus grande lorsque les renseignements sur la récidive sont recueillis à partir des dossiers officiels (nationaux ou locaux). Lorsque les sources officielles sont complétées par des renseignements d'autres sources crédibles, la confiance s'en trouve accrue. Des exemples de sources d'information crédibles sur la récidive seraient la police et les organismes de protection et de bien-être de l'enfance.

Cote	Description
0	<i>Les renseignements sur la récidive proviennent d'auto-évaluations seulement ou des dossiers des établissements et s'appliquent à un petit échantillon de récidivistes (p. ex. réadmission dans un foyer de groupe ou une maison de transition).</i>
1	<i>Une seule source d'information sur la récidive, des dossiers nationaux ou locaux ou des renseignements d'autres sources crédibles (p. ex. organismes de bien-être de l'enfance, enquêtes de la police).</i>
2	<i>Sources d'information multiples sur la récidive. Une source doit être des dossiers officiels sur les antécédents criminels (nationaux ou locaux). L'autre source peut être une autre source crédible (p. ex. organismes de bien-être de l'enfance, enquêtes de la police).</i>

17. Validité et la fiabilité des renseignements sur la récidive

Cet élément porte sur la qualité des renseignements sur la récidive et le niveau de confiance que ces renseignements sont exacts. Enregistrer toutes les sources d'information sur la récidive.

<i>Sources d'information sur la récidive provenant de l'étude :</i>			
Cote de confiance			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Peu de confiance que les sources d'information sur le suivi sont des indicateurs valides et fiables de la récidive</i>	1 <i>Certaine confiance que les sources d'information sur le suivi sont des indicateurs valides et fiables de la récidive</i>	2 <i>Grande confiance que les sources d'information sur le suivi sont des indicateurs valides et fiables de la récidive</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : validité et la fiabilité des renseignements sur la récidive

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0	1	2
<i>Peu de confiance que les sources d'information sur le suivi sont des indicateurs valides et fiables de la récidive</i>	<i>Certaine confiance que les sources d'information sur le suivi sont des indicateurs valides et fiables de la récidive</i>	<i>Grande confiance que les sources d'information sur le suivi sont des indicateurs valides et fiables de la récidive</i>

Raison(s) de la cote :

18. Équivalence du suivi

Concept : Le concept général est la mesure dans laquelle les récidivistes ont une chance égale d'être détectés dans le groupe expérimental et le groupe témoin. Les facteurs à prendre en considération sont les sources d'information sur la récidive et la durée du suivi. Des différences dans les sources d'information ou la période de suivi introduiraient des biais.

Indicateurs : Les indicateurs qui pourraient accroître les biais sont les périodes de suivi très différentes entre les groupes ou les cas où les sources d'information sur la récidive sont différentes. Les biais sont moins nombreux lorsque les renseignements sur la récidive ont été obtenus à partir des mêmes sources et que la durée de la période de suivi était la même pour chaque groupe (sur le plan de la tendance centrale et de la variabilité). Les biais sont les moins nombreux lorsque la période de suivi est fixe pour tous les participants.

Modèles de gammes de risques/normes : Ces modèles recevraient normalement une cote de « 0 – biais considérable » parce que les critères des résultats utilisés pour créer les normes seraient généralement différents des critères des résultats utilisés pour détecter la récidive dans le groupe expérimental. Elle pourrait être autre que zéro si l'étude et les normes utilisées étaient fondées sur les mêmes critères des résultats dans la même administration.

Cote	Description
0	<i>Les sources d'information pour chaque groupe sont différentes.</i> OU <i>La durée du suivi entre les groupes est différente et l'on s'attendrait à ce que ces différences aient une incidence considérable sur les taux de récidive globaux (p. ex. 3 ans au lieu de 5 ans) et une analyse de survie n'a pas été utilisée.</i>
1	<i>Les sources d'information pour chaque groupe sont identiques.</i> ET <i>La durée du suivi entre les groupes est différente, mais <u>soit</u></i> <i>a) on ne s'attendrait pas à ce que ces différences aient une incidence sur les taux de récidive globaux (p. ex. 21 ans contre 23 ans);</i> OU <i>b) la durée du suivi entre les groupes est très différente, et une analyse de survie a été utilisée.</i>
2	<i>La durée du suivi pour chaque groupe est fixe et équivalente.</i> ET <i>Les sources d'information sur la récidive pour chaque groupe sont identiques.</i>

18. Équivalence du suivi

Cet élément porte sur les biais introduits par le manque d'équivalence des renseignements sur la récurrence (durée, sources d'information). Si, par exemple, la période de suivi du groupe expérimental est plus courte que celle du groupe témoin, cela introduit un biais favorisant le groupe expérimental. Un autre exemple est le cas où des renseignements sur le suivi ont été recueillis à partir de dossiers de l'État pour le groupe témoin et où des auto-évaluations ont été utilisées pour le groupe expérimental. Ces différences dans les sources de renseignements sur le suivi peuvent entraîner des biais.

<i>Information et résultats concernant l'équivalence des renseignements sur le suivi provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : équivalence du suivi

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cotes révisées des biais

0	1	2
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>

Sens révisé des biais

?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>

Raison(s) de la cote :

VII. Comparaisons correctes établies

Dans toutes les études, les chercheurs doivent prendre un certain nombre de décisions lorsqu'ils évaluent l'effet général du traitement. Par exemple, le chercheur doit choisir parmi de multiples critères des résultats celui qu'il utilisera (p. ex. arrestations ou condamnations, durée du suivi) et les sujets de l'analyse. Il doit également choisir les variables à utiliser s'il a recours à des procédures statistiques pour tenir compte des différences collectives ou individuelles dans la possibilité de récidive ainsi que les options employées (p. ex. utiliser des valeurs moyennes ou supprimer les cas contenant des valeurs manquantes par paire ou en fonction d'une liste). Souvent il est impossible de déterminer entièrement d'avance ces décisions en raison de facteurs imprévisibles comme l'attrition et les renseignements manquants; néanmoins, ces décisions peuvent influencer sur la confiance dans les résultats et introduire des biais autres que ceux introduits par les facteurs décrits plus haut.

Comme les Lignes directrices du CCDCR définissent les études de qualité supérieure comme celles ayant le moins de biais et le degré de confiance le plus élevé, il y a trois éléments qui portent sur l'évaluation de l'effet du traitement. Le premier élément évalue la recherche d'une association à tout prix, qui a trait aux analyses exploratoires d'un ensemble de données avec peu de justification a priori. Ces analyses exploratoires a posteriori augmentent la probabilité d'une erreur de type 1 (conclusion erronée qu'un effet existe). L'ampleur de l'effet fondée sur cette « recherche d'une association à tout prix » influe sur la confiance qu'on peut avoir dans l'ampleur de l'effet indiquée.

Le deuxième élément évalue l'efficacité des procédures statistiques de contrôle des biais. Comme les différences entre le groupe expérimental et le groupe témoin sont courantes, même dans les modèles qui visent à réduire au minimum la possibilité de ces différences (p. ex. affectation aléatoire), les chercheurs tentent de tenir compte des différences observées dans l'évaluation de l'ampleur de l'effet au moyen de procédures statistiques. Bon nombre de ces méthodes sont statistiquement complexes et exigent de prendre des décisions concernant certaines options (p. ex. traitement des données manquantes). Il ne semble pas y avoir de consensus sur la meilleure approche ni sur l'efficacité avec laquelle ces approches doivent limiter les biais dans les échantillons où il y a des différences préexistantes considérables. Cet élément évalue la confiance dans ces procédures au moment de l'estimation de l'efficacité du traitement. C'est-à-dire que cet élément évalue le niveau de confiance dans les procédures statistiques utilisées pour réduire au minimum tout biais inhérent à l'étude (p. ex. faiblesses de la conception, facteurs de sélection des sujets, différences préexistantes entre les groupes et différences dans le suivi).

Enfin, le troisième élément, appelé évaluation de la comparaison qui introduit le moins de biais, porte sur les biais possibles dans l'ampleur de l'effet générale calculée et présentée par le ou les chercheurs. L'ampleur de l'effet du traitement est déterminée par la comparaison du groupe des délinquants « traités » et du groupe « témoin ». Dans chaque étude, il est possible de calculer un

effet du traitement d'après les sujets qui introduirait le moins de biais possibles avec un degré de confiance raisonnable (p. ex. taille de l'échantillon). À cette fin, il faut tenir compte d'un certain nombre de facteurs qui ont été décrits plus haut comme l'attrition (p. ex. les sujets qui ne terminent pas le traitement, qui refusent de participer ou qui abandonnent), la taille de l'échantillon et la validité et la fiabilité des renseignements sur le suivi. La comparaison affichant le moins de biais et le degré de confiance le plus élevé comprendrait généralement une comparaison entre le plus grand nombre possible de sujets affectés à chaque groupe et les renseignements les plus détaillés, valides et fiables possibles.

19. Recherche d'une association à tout prix

Concept : Le concept général consiste à déterminer si les comparaisons indiquées sont fondées sur un raisonnement théorique corroboré par la recherche et si l'étude visait précisément à évaluer ou mesurer ces comparaisons. Un autre facteur à prendre en considération est le nombre général de comparaisons, en particulier le nombre de comparaisons a posteriori. La probabilité qu'une erreur de type 1 augmente avec l'accroissement du nombre de comparaisons. Il est impossible de traiter les biais introduits par les comparaisons multiples en utilisant des valeurs p prudentes (p. ex. $p < 0,001$) parce que les données en question sont les différences entre les groupes, et non la signification statistique de ces différences. Cet élément vise à évaluer la confiance du codeur que les comparaisons ont été prévues a priori et ne résultent pas d'une « recherche d'une association à tout prix ».

Indicateurs : Les indicateurs qui pourraient diminuer le degré de confiance sont un grand nombre de comparaisons, des analyses a posteriori ou les cas où les justifications a priori sont faibles. Les indicateurs qui pourraient accroître le degré de confiance sont les études où peu de comparaisons ont été établies, les cas où des analyses ont été déterminées a priori et où ces comparaisons se reflètent dans d'autres caractéristiques inhérentes de l'étude, comme les objectifs du traitement, les méthodes de traitement ou les critères de sélection des sujets.

Cote	Description
0	<i>L'ampleur de l'effet était une des nombreuses comparaisons a posteriori, et ces comparaisons n'étaient pas prévues.</i>
1	<i>On ne sait pas clairement si les comparaisons primaires étaient planifiées a priori ou a posteriori. Ces comparaisons semblent raisonnables, mais il est possible qu'elles aient été établies a posteriori.</i>
2	<i>Les comparaisons (ampleur de l'effet) ont été planifiées a priori. La justification de ces comparaisons se reflète dans la conception de l'étude (p. ex. facteurs de sélection des sujets ou objectifs du traitement). Le nombre de comparaisons nécessaires pour tester les hypothèses était minime.</i>

19. Recherche d'une association à tout prix

Cet élément porte sur la planification a priori et le nombre de comparaisons établies pour mesurer l'efficacité du traitement et le niveau de confiance que l'estimation de l'efficacité du traitement n'est pas attribuable à la chance. Il faut enregistrer toutes les comparaisons et l'ampleur de l'effet.

<i>Estimation de l'efficacité du traitement et des médiateurs provenant de l'étude :</i>			
Cote de confiance			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Peu de confiance dans les résultats, car les comparaisons ont été établies a posteriori</i>	1 <i>Certaine confiance dans les résultats, car les comparaisons peuvent avoir été établies a posteriori</i>	2 <i>Grande confiance dans les résultats, car les comparaisons ont été planifiées a priori</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : recherche d'une association à tout prix

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance dans les résultats, car les comparaisons ont été établies a posteriori

1

Certaine confiance dans les résultats, car les comparaisons peuvent avoir été établies a posteriori

2

Grande confiance dans les résultats, car les comparaisons ont été planifiées a priori

Raison(s) de la cote :

20. Efficacité des procédures statistiques de contrôle des biais

Concept : Le concept général consiste à déterminer si les chercheurs ont utilisé des statistiques pour tenir compte des différences préexistantes entre les groupes et si ces procédures seraient efficaces. Il y a trois facteurs à prendre en considération. Un facteur est l'équivalence des groupes. Lorsque les différences préexistantes sont importantes (prévues ou démontrées), la confiance dans l'efficacité des procédures de contrôle statistique des biais sont considérablement réduites. Lorsque les différences sont minimales, la confiance que les procédures de contrôle statistique peuvent permettre un contrôle efficace des biais augmente. Le deuxième facteur à prendre en considération est la qualité des variables de contrôle statistique. La qualité des variables de contrôle statistique a trait à leur fiabilité et aux relations prédictives avec le résultat. Lorsque les variables de contrôle ne s'appliquent pas au risque, la confiance dans les analyses de contrôle statistique des biais est faible. D'autre part, lorsque les variables de contrôle sont fiables et qu'il existe un lien étroit entre elles et le résultat, la confiance augmente. Le troisième facteur à prendre en considération est la crédibilité des procédures statistiques utilisées. Il y a un certain nombre d'analyses statistiques différentes qui pourraient être utilisées, et le chercheur doit en choisir une. Par exemple, certaines procédures statistiques exigent une répartition normale des scores, mais il peut être clair que les scores ne sont pas répartis normalement.

Indicateurs : Trois éléments évalués précédemment fournissent des renseignements qui peuvent aider à évaluer cet élément. Lorsqu'il étudie les différences entre les groupes, le codeur peut examiner les scores concernant l'équivalence prévue a priori des groupes (élément 13) et les constatations sur l'équivalence des groupes (élément 15). Pour la qualité des variables de contrôle utilisées dans les analyses, l'élément 14 (qualité de la recherche de différences préexistantes) peut fournir des renseignements utiles. Dans cet élément, le codeur doit tenir compte de ces facteurs ainsi que des procédures statistiques et des variables de contrôle utilisées. Il y a peu de confiance lorsque aucune procédure de contrôle statistique n'a été utilisée, lorsqu'il existe de grandes différences entre les groupes, prévues a priori ou démontrées statistiquement, ou lorsque les variables de contrôle ne s'appliquaient pas au risque (p. ex. groupe ethnique, langue, lieu). Il convient cependant de noter que dans les modèles de traitement de rechange, les chercheurs voudraient également tenir compte des variables qui s'appliqueraient à la capacité des délinquants de bénéficier des interventions (c.-à-d. variables de réceptivité), même lorsque ces variables peuvent n'avoir aucun lien direct avec le risque (p. ex. la langue).

Cote	Description
0	<p><i>Il n'y avait aucun contrôle du risque dans le modèle initial (p. ex. appariement fondé sur le risque, affectation aléatoire), et des contrôles statistiques a posteriori n'ont pas été utilisés.</i></p> <p style="text-align: center;">OU</p> <p><i>Des procédures de contrôle statistique ont été utilisées, mais elles étaient probablement insuffisantes pour contrôler les biais en raison des grandes différences entre les groupes, ou les variables de contrôle statistique étaient inadéquates.</i></p>
1	<p><i>Les contrôles du risque étaient inhérents dans le modèle initial (p. ex. appariement fondé sur le risque, affectation aléatoire), et des contrôles statistiques a posteriori n'ont pas été utilisés.</i></p> <p style="text-align: center;">OU</p> <p><i>Des procédures de contrôle statistique ont été utilisées et étaient probablement efficaces, car il y avait une certaine incertitude au sujet de leur capacité de contrôler les biais.</i></p>
2	<p><i>Des procédures de contrôle statistique ont été utilisées et étaient efficaces, car il y avait peu de différences entre les groupes, et les variables de contrôle utilisées étaient adéquates pour contrôler les biais possibles.</i></p>

20. Efficacité des procédures statistiques de contrôle des biais

Cet élément porte sur les méthodes statistiques utilisées pour tenir compte des différences préexistantes au moment de l'estimation de l'efficacité du traitement et du niveau de confiance dans ces procédures statistiques. Il convient de noter les méthodes statistiques et les variables dont tient compte l'analyse des variables.

<i>Procédures statistiques utilisées dans l'étude pour estimer les effets du traitement</i>			
Cote de confiance			
—	0	1	2
<i>Renseignements insuffisants pour permettre une évaluation</i>	<i>Peu de confiance que les procédures statistiques étaient suffisantes pour tenir compte des différences entre les groupes</i>	<i>Certaine confiance que les procédures statistiques pouvaient être suffisantes pour tenir compte des différences entre les groupes</i>	<i>Grande confiance que les procédures statistiques étaient suffisantes pour tenir compte des différences entre les groupes</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : efficacité des procédures statistiques de contrôle des biais

Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?

Si de nouveaux renseignements ont été obtenus, précisez.

Cote de confiance révisée

0

Peu de confiance que les procédures statistiques étaient suffisantes pour tenir compte des différences entre les groupes

1

Certaine confiance que les procédures statistiques pouvaient être suffisantes pour tenir compte des différences entre les groupes

2

Grande confiance que les procédures statistiques étaient suffisantes pour tenir compte des différences entre les groupes

Raison(s) de la cote :

21. Évaluation de la comparaison qui introduit le moins de biais

Concept : Le concept général consiste à déterminer si le chercheur choisit de baser l'effet du traitement sur les sujets et la ou les variables des résultats qui seraient les moins susceptibles d'introduire des biais et donneraient le niveau de confiance le plus élevé. Souvent, lorsqu'il évalue l'effet du traitement, un chercheur doit prendre un certain nombre de décisions concernant les sujets à inclure dans l'étude et choisir parmi les multiples résultats à utiliser. Ces décisions devraient tenir compte de tous les facteurs précédents qui peuvent introduire des biais (p. ex. attrition, équivalence des groupes, mesures des résultats) et modifier la confiance dans les résultats (p. ex. taille de l'échantillon). L'objectif du chercheur est d'estimer l'efficacité du traitement contenant le moins de biais possible et produisant le degré de confiance le plus élevé possible.

Indicateurs : Les indicateurs potentiels sont tous les éléments précédents comme les données manquantes, la taille de l'échantillon, l'attrition, etc., ainsi que d'autres facteurs pertinents comme la composition du groupe témoin, la façon dont les chercheurs ont examiné les données manquantes et les options utilisées dans les analyses statistiques.

Cote	Description
0	<i>Les chercheurs ont inclus certains sujets dans les analyses et(ou) les en ont exclus, qui introduisaient des biais dans l'estimation de l'efficacité du traitement.</i> OU <i>Les chercheurs ont choisi une mesure du résultat dont on s'attend raisonnablement qu'elle introduise des biais.</i>
1	<i>La comparaison choisie était plausible, mais il y avait une certaine incertitude quant à savoir si la comparaison ou l'analyse statistique optimale a été choisie.</i>
2	<i>Utilisant la mesure du résultat la plus complète, la plus valide et la plus fiable possible, les chercheurs ont affecté le plus de sujets possible à chaque groupe lorsqu'ils ont évalué les effets du traitement.</i>

21. Évaluation de la comparaison qui introduit le moins de biais

Cet élément porte sur les groupes (et les sujets de chaque groupe) et les mesures utilisées pour déterminer l'efficacité du traitement et les biais potentiels que cela introduit dans les résultats. Dans la plupart des cas, il y a un certain nombre de comparaisons possibles. Il faut déterminer la comparaison utilisée par les chercheurs et évaluer la comparaison la moins biaisée et la plus fiable disponible.

<i>Information et résultats concernant l'efficacité du traitement provenant de l'étude :</i>			
Cotes des biais			
— <i>Renseignements insuffisants pour permettre une évaluation</i>	0 <i>Introduit un nombre considérable de biais</i>	1 <i>Certains biais susceptibles d'être introduits dans les résultats</i>	2 <i>Attente de biais négligeables dans les résultats</i>
Sens des biais			
? <i>Impossible d'évaluer le sens des biais</i>	+1 <i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	0 <i>Aucun biais prévu</i>	-1 <i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

Vaut-il la peine de chercher des renseignements supplémentaires ou de nouvelles analyses? Dans l'affirmative, remplir la page suivante.

Renseignements supplémentaires : évaluation de la comparaison qui introduit le moins de biais			
<i>Quels sont les renseignements supplémentaires recherchés et pourquoi le sont-ils?</i>			
<i>Si de nouveaux renseignements ont été obtenus, précisez.</i>			
Cotes révisées des biais			
0	1	2	
<i>Introduit un nombre considérable de biais</i>	<i>Certains biais susceptibles d'être introduits dans les résultats</i>	<i>Attente de biais négligeables dans les résultats</i>	
Sens révisé des biais			
?	+1	0	-1
<i>Impossible d'évaluer le sens des biais</i>	<i>Les biais accroissent probablement le degré d'efficacité du traitement</i>	<i>Aucun biais prévu</i>	<i>Les biais diminuent probablement le degré d'efficacité du traitement</i>
<i>Raison(s) de la cote :</i>			

VIII. Évaluation globale

L'évaluation globale est un sommaire général de la qualité de l'étude. Il s'agit d'une forme de jugement structuré, fondé sur chacun des éléments et les cotes sommaires de la confiance et des biais. Pour l'évaluation globale, il est utile de définir clairement les principales conclusions de l'étude (ampleurs de l'effet, tailles des échantillons). Nous recommandons aussi que les codeurs examinent les raisons de leurs décisions lorsqu'ils transcrivent les scores qu'ils ont attribués à chacun des éléments sur la feuille sommaire. Les codeurs évaluent ensuite leur niveau de confiance dans l'étude d'après la qualité de l'évaluation et de l'examen des menaces potentielles à la validité interne. Ensuite, ils examinent l'ampleur et le sens des biais qui se reflètent dans les principales conclusions (ampleurs de l'effet). Même s'il devrait y avoir une corrélation entre les évaluations sommaires de la confiance et des biais et les évaluations de chacun des éléments, nous n'avons pas proposé un algorithme précis pour traduire les évaluations individuelles dans les évaluations sommaires. Après que les évaluations sommaires des biais et de la confiance sont terminées, toutefois, des directives explicites sont fournies pour la production d'une évaluation globale de la qualité de l'étude selon quatre catégories : rejetée, faible, bonne et excellente.

- a) **EXCELLENTE.** Grande confiance que l'étude contient un biais minime dans l'estimation de l'efficacité du traitement des délinquants sexuels. Elle est bien conçue et bien exécutée et donne des résultats convaincants. L'étude peut présenter des problèmes mineurs, mais ces problèmes sont peu susceptibles d'influer sur les principales conclusions ou de modifier le sens des effets observés.
- b) **BONNE.** Grande confiance que l'étude ne contient pas plus que quelques biais (évaluation intermédiaire). Des efforts raisonnables ont été faits pour corriger les menaces à la validité, mais il reste beaucoup d'inconnues.
- c) **FAIBLE.** Certaine confiance que l'étude ne contient pas plus que quelques biais. L'étude présente des lacunes graves, mais elle peut être pertinente à la question de l'efficacité du traitement. Données faibles au mieux.
- d) **REJETÉE.** Faible confiance dans les résultats, *ou* biais considérables. L'étude présente de multiples lacunes importantes. On s'attendrait à ce que les procédures utilisées introduisent des biais considérables, ou l'étude manque de renseignements importants nécessaires pour éliminer les explications de rechange plausibles des conclusions.

ÉVALUATION GLOBALE

AMPLEUR DE L'EFFET			
0			
<i>Impossible d'évaluer l'ampleur de l'effet : aucune autre évaluation requise.</i>	1		
<i>Ampleur de l'effet évaluée</i>			
<i>Estimation de l'ampleur de l'effet introduisant le moins de biais provenant de l'étude.</i>			
<i>Détails sur l'ampleur de l'effet (c.-à-d. description des sujets et du traitement, données utilisées pour l'évaluation).</i>			
Confiance			
<i>Évaluer la confiance dans la validité interne de l'étude. La confiance a trait à la validité et à la fiabilité de l'ampleur de l'effet comme mesure dans cette étude de l'effet du traitement (mis en œuvre et donné à l'échantillon de sujets selon la description) sur la mesure du résultat utilisée.</i>			
0	1	2	
<i>Peu ou pas de confiance</i>	<i>Certaine confiance</i>	<i>Grande confiance</i>	
<i>Raison(s) de la cote :</i>			
Biais			
<i>Évaluation de la quantité et du sens des biais inhérents à l'ampleur de l'effet évaluée ci-dessus. Il faut tenir compte du modèle et des facteurs spécifiques de l'étude et donner les raisons de la cote.</i>			
0	1	2	
<i>Biais considérable</i>	<i>Certains biais</i>	<i>Aucun biais prévu (sens : 0)</i>	
-	-		
<i>Effet décroissant du traitement</i>	<i>Effet décroissant du traitement</i>		
+	+		
<i>Effet croissant du traitement</i>	<i>Effet croissant du traitement</i>		
?	?		
<i>Répercussion inconnue sur les résultats</i>	<i>Répercussion inconnue sur les résultats</i>		
<i>Raison(s) de la cote :</i>			
ÉVALUATION GLOBALE DE LA QUALITÉ DE L'ÉTUDE			
0	1	2	3
<i>Rejetée</i>	<i>Faible</i>	<i>Bonne</i>	<i>Excellente</i>
<i>Cote de confiance de 0</i>	<i>Cote de confiance de 1</i>	<i>Confiance de 1 et biais de 2</i>	<i>Cote de confiance de 2</i>
<i>OU</i>	<i>ET</i>	<i>OU</i>	<i>ET</i>
<i>Cotes des biais de -0 ou +0 ou ?0</i>	<i>Cotes des biais de -1 ou +1 ou ?1</i>	<i>Confiance de 2 et biais de -1 ou +1 ou ?1</i>	<i>Cotes des biais de 2</i>

Feuille sommaire du guide d'évaluation de la qualité de l'étude (1 de 2)

<i>Confiance</i>	<i>Peu de confiance</i> 0	<i>Certaine confiance</i> 1	<i>Grande confiance</i> 2
<i>Biais</i>	<i>Biais considérable</i>	<i>Certain biais</i>	<i>Biais négligeable</i>
<i>Sens des biais</i>	+ <i>Accroît le traitement</i>	- <i>Diminue le traitement</i>	? <i>Sens inconnu</i>

Contrôle administratif des variables indépendantes

1. Définition du traitement			
	<i>Confiance</i>		
2. Définition de la comparaison			
	<i>Confiance</i>		
3. Divers facteurs secondaires			
		<i>Biais</i>	<i>Sens</i>

Attentes de l'expérimentateur

4. Participation de l'expérimentateur			
		<i>Biais</i>	<i>Sens</i>
5. Évaluation à l'aveugle de la gestion des données			
		<i>Biais</i>	<i>Sens</i>

Taille de l'échantillon

6. Taille de l'échantillon du ou des groupes expérimentaux			
	<i>Confiance</i>		
7. Taille de l'échantillon du ou des groupes témoins			
	<i>Confiance</i>		
8. Taille de l'échantillon d'établissements : modèles fondés sur les établissements			
	<i>Confiance</i>		

Attrition

9. Sélection des sujets			
		<i>Biais</i>	<i>Sens</i>
10. Abandon du programme			
		<i>Biais</i>	<i>Sens</i>
11. Analyse des sujets retenus au début de l'étude			
		<i>Biais</i>	<i>Sens</i>
12. Attrition pendant le suivi			
		<i>Biais</i>	<i>Sens</i>

Feuille sommaire du guide d'évaluation de la qualité de l'étude (2 de 2)

Équivalence des groupes

13. Équivalence des groupes a prior			
		<i>Biais</i>	<i>Sens</i>
14. Qualité de la recherche de différences préexistantes	<i>Confiance</i>		
15. Constatations sur l'équivalence des groupes		<i>Biais</i>	<i>Sens</i>

Variables des résultats

16. Durée du suivi			
	<i>Confiance</i>		
17. Validité et fiabilité des renseignements sur la récidence	<i>Confiance</i>		
18. Équivalence du suivi		<i>Biais</i>	<i>Sens</i>

Comparaisons correctes établies

19. Recherche d'une association à tout prix			
	<i>Confiance</i>		
20. Efficacité des contrôles statistiques	<i>Confiance</i>		
21. Calcul de la comparaison qui introduit le moins de biais		<i>Biais</i>	<i>Sens</i>

ÉVALUATION GLOBALE

Ampleur de l'effet et nombre	
Confiance globale (0 = peu/aucune, 1 = moyenne, 2 = Grande)	
Quantité globale de biais (0 = considérable, 1 = moyenne, 2 = négligeable)	
Sens global des biais (? = inconnu, + = accroissement - = décroissement)	
Évaluation globale (0 = Rejetée, 1 = Faible, 2 = Bonne, 3 = Excellente)	

Bibliographie

- ALTMAN, D. G., K. F. SCHULTZ, D. MOHER, M. EGGER, F. DAVIDOFF, D. ELBOURNE, P. C. GÖTZCHE et T. LANG pour le CONSORT Group. « The revised CONSORT statement for reporting randomized trials: Explanation and elaboration », *Annals of Internal Medicine*, vol. 134, p. 663-694, 2001.
- COHEN, J. *Statistical Power analysis for the behavioural sciences*, 2^e éd., Hillsdale (New Jersey), Lawrence Erlbaum, 1988.
- COMITÉ DE COLLABORATION SUR LES DONNÉES COLLECTIVES RELATIVES AUX RÉSULTATS. *Lignes directrices sur l'évaluation des recherches sur l'efficacité des programmes de traitement pour délinquants (Lignes directrices du CCDCR), Partie 1 : Introduction et aperçu*, Rapport pour spécialistes 2007-02, Ottawa, Sécurité Canada, 2007.
- GENDREAU, P., C. GOGGIN et P. SMITH. « The forgotten issue in effective correctional treatment : Program implementation », *International Journal of Offender Therapy and Comparative Criminology*, vol. 43, p. 180-187, 1999.
- HANSON, R. K., A. GORDON, A. J. R. HARRIS, J. K. MARQUES, W. MURPHY, V. L. QUINSEY et M. C. SETO. « First report of the Collaborative Outcome Data Project on the effectiveness of psychological treatment of sex offenders », *Sexual Abuse: A Journal of Research and Treatment*, vol. 14, p. 169-194, 2002.
- HANSON, R. K. et T. NICHOLAICHUK. « A cautionary note regarding Nicholaichuk et al. », *Sexual Abuse: Journal of Research and Treatment*, vol. 12, p. 289-293, 2000.
- HARRIS, A. J. R. et R. K. HANSON. *La récidive sexuelle : d'une simplicité trompeuse*, Rapport pour spécialistes 2004-03, Sécurité publique et Protection civile Canada, 2004.
- JUNI, P., A. WITSCHI, R. BLOCH et M. EGGER. « The hazards of scoring the quality of clinical trials for meta-analysis », *Journal of the American Medical Association*, vol. 282, p. 1054-1060, 1999.
- SCHULTZ, K. F., I. CHALMERS, R. J. HAYES et D. J. ALTMAN. « Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials », *Journal of the American Medical Association*, vol. 273, p. 408-412, 1995.
- SHADISH, W. R., T. D. COOK et D. T. CAMPBELL. *Experimental and quasi-experimental designs for generalized causal inference*, Boston, Houghton Mifflin, 2002.

THOMAS, H., D. CILISKA, M. DOBBINS et S. MICUCCI. A process for systematically reviewing the literature: Providing the research evidence for public health nursing interventions, *Worldviews on Evidence-Based Nursing*, vol. 2, p. 91-99, 2004.

WESTEN, D., C. M. NOVOTNY et H. THOMPSON-BRENNER. « The empirical status of empirically supported psychotherapies: Assumptions, findings and reporting in controlled trials », *Psychological Bulletin*, vol. 130, p. 631-663, 2004.

ZAZA, S., L. K. WRIGHT-DE AGÜERO, P. A. BRISS, B. I. TRUMAN, D. P. HOPKINS, M. H. HENNESSY et coll. « Data collection instrument and procedure for systematic reviews in the *Guide to Community Preventive Services* », *American Journal of Preventive Medicine*, vol. 18, p. 44-74, 2000.

Membres du Comité

Anthony R. Beech, Ph. D., est professeur de psychologie criminologique au Centre for Forensic and Family Psychology, School of Psychology, University of Birmingham, Royaume-Uni, et Fellow de la British Psychological Society. a.r.beech@bham.ac.uk

Guy Bourgon, Ph. D., est agent de recherche à Sécurité publique Canada et professeur auxiliaire au Département de psychologie de l'Université Carleton, Ottawa, Canada.
Guy.Bourgon@ps.gc.ca

R. Karl Hanson, Ph. D., est agent de recherche principal à Sécurité publique Canada et professeur auxiliaire au Département de psychologie de l'Université Carleton, Ottawa, Canada. Karl.Hanson@ps.gc.ca

Andrew J. R. Harris, Ph. D., est gestionnaire de recherche principal, Direction de la recherche, Service correctionnel du Canada. Il n'enseigne pas. HarrisAJ@csc-scc.gc.ca

Calvin M. Langton, Ph. D., est chargé d'enseignement au Département de psychiatrie, Université de Toronto, Canada, et chercheur honoraire à la School of Community Health Sciences, University of Nottingham, Royaume-Uni. calvin.langton@utoronto.ca

Janice Marques, Ph. D., est psychologue consultante; elle a pris sa retraite récemment après avoir travaillé pour le Department of Mental Health de la Californie. Elle était présidente de l'Association for the Treatment of Sexual Abusers (ATSA) lorsque ce projet de collaboration a été lancé en 1998. jkmarques@sbcglobal.net

Michael H. Miner, Ph. D., est professeur agrégé au Program in Human Sexuality, Department of Family Medicine and Community Health, University of Minnesota, Minneapolis (Minnesota). miner001@umn.edu

William Murphy, Ph. D., est professeur au Département de psychiatrie, University of Tennessee Health Science Center, Memphis (Tennessee). wmurphy@utmem.edu

Michael Seto, Ph. D., est psychologue au Law and Mental Health Program, Centre for Addiction and Mental Health, et professeur agrégé au Département de psychiatrie et au Centre de criminologie de l'Université de Toronto. Michael_Seto@camh.net

Vernon Quinsey est professeur et chef de la Psychologie à l'Université Queen's, Kingston (Ontario). vernon.quinsey@queensu.ca

David Thornton, Ph. D., est directeur des traitements au Sand Ridge Secure Treatment Center, Mauston (Wisconsin). thorndm@dhfs.state.wi.us

Pamela M. Yates, Ph. D., est psychologue au Service correctionnel du Canada et spécialiste dans le traitement des délinquants sexuels. YatesPM@csc-scc.gc.ca