



Estimation de la prévalence —
Méthodes indirectes pour
estimer l'ampleur du problème
de la drogue

Estimation de la prévalence — Méthodes indirectes pour estimer l'ampleur du problème de la drogue

2



OFFICE DES NATIONS UNIES CONTRE LA DROGUE ET LE CRIME
Vienne

Estimation de la prévalence – Méthodes indirectes pour estimer l'ampleur du problème de la drogue

Programme mondial d'évaluation
de l'abus de drogues

Module 2 du référentiel



NATIONS UNIES
New York, 2003

Le module 2 du référentiel GAP a été établi par l'Office des Nations Unies contre la drogue et le crime au titre des activités menées dans le cadre du Programme mondial d'évaluation de l'abus de drogues (GAP).

Pour de plus amples renseignements, consulter le site Web du GAP (www.unodc.org), envoyer un message électronique à l'adresse: gap@unodc.org ou contacter la Section de la réduction de la demande, Office des Nations Unies contre la drogue et le crime, B.P. 500, A-1400 Vienne (Autriche).

Office des Nations Unies contre la drogue et le crime
Imprimé en Autriche, 2003



Préface

Le module 2 du référentiel Gap intitulé “Estimation de la prévalence – Méthodes indirectes pour estimer l’ampleur du problème de la drogue” a été élaboré avec l’aide de l’Office des Nations Unies contre la drogue et le crime (ONUDC), dans le cadre des activités du Programme mondial d’évaluation de l’abus de drogues (GAP). Le principal objectif du GAP est d’aider les pays à recueillir des données fiables et comparables à l’échelon international sur l’abus de drogues, à mettre en place des moyens au niveau local pour recueillir des données permettant d’orienter les activités de réduction de la demande et à améliorer la communication d’informations aux échelons transnational, régional et mondial sur les tendances en matière de drogues.

Le référentiel GAP a été élaboré afin d’aider les États Membres de l’Organisation des Nations Unies à mettre au point des systèmes de collecte d’informations sur les drogues qui soient culturellement adaptés à leur contexte national, ainsi que pour faciliter la mise en conformité des systèmes existants d’information sur les drogues avec des normes de bonne pratique internationalement reconnues, et en particulier l’harmonisation des indicateurs d’abus de drogues.

Le module 2 du référentiel GAP fait partie d’un ensemble de guides méthodologiques destinés à faciliter les activités de collecte de données. Les autres modules actuellement en cours d’élaboration concernent les domaines suivants: mise en place d’un système intégré d’information sur les drogues, enquêtes en milieu scolaire, interprétation et gestion des données pour l’élaboration des politiques et initiation à l’exploitation des données au moyen d’un progiciel statistique pour sciences sociales (SPSS).

Le GAP fournit également un appui technique et financier pour la mise en place de systèmes d’information sur les drogues, et soutient et coordonne des activités de collecte de données à l’échelle mondiale. Pour en savoir plus sur les modules du référentiel GAP, veuillez envoyer un message électronique à l’adresse: gap@unodc.org ou consulter le site Web du GAP (www.unodc.org).

La raison d’être du référentiel est d’offrir un guide pratique et accessible pour la réalisation des activités de collecte de données dans les domaines les plus importants. Les modules du référentiel n’ont pas la prétention d’apporter toutes les réponses, mais visent plutôt à fournir un point de départ pour la réalisation d’activités précises en renvoyant le lecteur à des sources d’information plus détaillées sur telle ou telle question. Les principes sur lesquels ils s’appuient pour la collecte des données ont été adoptés par un groupe international d’experts et approuvés par les États Membres de l’Organisation des Nations Unies. Bien que les modèles qui y sont présentés soient fondés sur des modèles existants qui se sont révélés efficaces dans la pratique, il est essentiel que les démarches soient adaptées aux nécessités et aux conditions locales. Le présent module donne des exemples précis indiquant au

lecteur comment ces principes et ces modèles généraux peuvent être adaptés à des contextes particuliers, mais il ne cherche pas à rendre compte de l'éventail complet des systèmes d'information et des méthodes de collecte de données sur les drogues qui existent actuellement ni de leur diversité.

Pour en savoir plus, veuillez consulter le site Web du GAP (www.unodc.org), envoyer un message électronique à l'adresse: gap@unodc.org ou vous mettre en rapport avec la Section de la réduction de la demande, Office des Nations Unies contre la drogue et le crime, B.P. 500, A-1400 Vienne (Autriche).

Remerciements

L'Office des Nations Unies contre la drogue et le crime tient à remercier les nombreux homologues nationaux qui ont communiqué des observations au sujet de la version préliminaire du référentiel et aidé à l'expérimenter, ainsi que les organisations et les personnes qui ont fourni des exemples de formulaires, d'outils et d'autres documents connexes utilisés pour la collecte des données. En particulier, il convient de remercier le Community Epidemiology Work Group, le Groupe européen de coopération en matière de lutte contre l'abus et le trafic illicite des stupéfiants (Groupe Pompidou), l'Observatoire européen des drogues et des toxicomanies, le Réseau d'épidémiologie sur l'abus des drogues de la Communauté de développement de l'Afrique australe, le Système d'information sur les drogues pour l'Afrique orientale et le Réseau d'informations sur les drogues pour les Caraïbes.

Le présent module a été établi par un groupe d'experts avec l'aide de l'Office des Nations Unies contre la drogue et le crime dans le cadre de l'exécution du Programme mondial d'évaluation de l'abus de drogues (GAP). Il convient de remercier, en particulier, Colin Taylor, qui en a coordonné l'élaboration et en a été le rédacteur, Mathew Hickman, qui a fait fonction de corédacteur, et Rebecca McKetin, qui est intervenue comme coordonnatrice et conseillère au cours des phases finales du projet. Le groupe d'experts était composé comme suit: Colin Taylor du National Addiction Centre (Londres); Mathew Hickman de l'Imperial College (Londres); Michael Lynskey du National Drug and Alcohol Research Centre, Nouvelle-Galles du Sud (Australie); Lucas Wiessing de l'Observatoire européen des drogues et des toxicomanies; Paul Griffiths, Rebecca McKetin et Kamran Niaz de l'Office des Nations Unies contre la drogue et le crime; et Anindya Chatterjee du Programme commun des Nations Unies sur le VIH/sida. Le groupe a bénéficié de l'assistance de Mathew Warner-Smith, Office des Nations Unies contre la drogue et le crime, Afrique australe.

Le module est le fruit de la collaboration des membres du groupe, lequel s'est réuni en tant que conseil de rédaction informel chargé d'aider le rédacteur. Son rôle a consisté à donner des conseils concernant la teneur et la structure du module et à indiquer qui était susceptible de contribuer à l'élaboration de ses différentes parties.

À sa première réunion, le groupe a apporté une contribution majeure en déterminant la portée et la structure du module. Son concours au projet a été inestimable.

Table des matières

Préface	III
Remerciements	V
INTRODUCTION	1
Renseignements généraux	1
Le rôle de l'estimation de la prévalence	3
Le projet GAP	4
CHAPITRE PREMIER. ÉVALUATION DE LA PRÉVALENCE DE L'USAGE DE DROGUES	9
Évaluation de la prévalence et nécessité d'évaluer le dommage	9
Estimations de prévalence et rôle des enquêtes sur la population générale ..	10
Prévalence de l'usage de drogues et autres méthodes épidémiologiques ..	12
Méthodes indirectes d'estimation de la prévalence	13
Enquêtes de prévalence nationales et locales	14
CHAPITRE II. MÉTHODES SPÉCIFIQUES D'ESTIMATION DE LA PRÉVALENCE ...	15
Aperçu général	15
Les méthodes multiplicateur/référence	20
Méthodes de saisie avec recoupement (capture-recapture)	36
Utilisation de techniques de modélisation évoluées	52
Estimations nationales de la prévalence établies par extrapolation à partir d'estimations locales	55
Exactitude des estimations de la prévalence	67
CHAPITRE III. PRINCIPES DIRECTEURS POUR L'ÉTABLISSEMENT D'ESTIMATIONS FONDÉES SUR LES RÉSULTATS DES ACTIVITÉS DE RECHERCHE	73
Principes généraux	73
Principes techniques	78
Annexe. Ressources générales utiles pour les études de prévalence	89

Tableaux

1. Sources de données susceptibles de servir de référence pour estimer la prévalence de l'usage problématique de drogues au moyen de méthodes de multiplicateur/référence	21
2. Estimation du nombre d'usagers injecteurs à partir du nombre de tests de dépiçage du VIH	22
3. Utilisation du nombre de décès par surdose pour estimer le nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud sur la base d'un taux de mortalité annuel de 1 %	23
4. Estimation du nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud à partir du nombre de patients suivant un traitement d'entretien à la méthadone	26
5. Estimation du nombre d'usagers d'opiacés à Bangkok, en 1991	38

6. Données de capture-recapture multisources tirées d'une étude sur les usagers de drogues par injection à Glasgow (Écosse), en 1990	42
7. Estimation de l'usage de drogues par injection à Glasgow – Nombre d'usagers et prévalence	42
8. Critères d'appariement utilisés pour les estimations de la prévalence dans l'étude de Bangkok	45
9. Critères d'appariement utilisés pour les estimations de la prévalence dans l'étude de Glasgow	46
10. Structure des données aux fins de l'estimation de la prévalence par extrapolation	59
11. Données relatives à l'abus de drogues concernant 11 régions du Royaume-Uni, y compris quatre points d'ancrage	63
12. Données et estimations de la prévalence au moyen d'indicateurs multiples concernant le Royaume-Uni (1996)	64

Études de cas

1. Étude réalisée à Toronto au moyen de la méthode du multiplicateur à partir de tests de dépistage du VIH	21
2. Décès par surdose en Nouvelle-Galles du Sud	23
3. Calcul du multiplicateur à partir de données relatives aux traitements et aux arrestations en Nouvelle-Galles du Sud	25
4. Évaluation nationale au Pakistan en 2000	28
5. Exemple concret tiré d'une étude menée à Bangkok	37
6. Étude sur les usagers de drogues par injection à Glasgow (Écosse)	41
7. Estimation de la prévalence des prostitué(e)s à Dhaka (Bangladesh)	46
8. Étude de Glasgow sur les prostituées	47
9. L'étude de Jersey	48
10. Extrapolation à partir de l'étude sur la Nouvelle-Galles du Sud	57
11. Estimation de la prévalence par extrapolation à l'aide d'indicateurs multiples (au Royaume-Uni)	61

Introduction

Renseignements généraux

Bien que certains États aient obtenu des succès en matière de lutte contre la demande de drogues illicites, l'abus des drogues continue d'augmenter dans le monde. En particulier, l'abus de drogues illicites dans certains pays en développement s'est accru de façon spectaculaire. Toutefois, l'ampleur de l'utilisation des drogues illicites reste mal connue, de même que les constantes et les tendances en la matière.

Pour mettre en place des politiques efficaces de réduction de l'abus des drogues, les gouvernements ont besoin de savoir quand, où et pourquoi des drogues illicites sont utilisées. Étant donné que les utilisateurs dans le monde entier ont accès à une plus grande variété de drogues et que les tendances sociales, en particulier parmi les jeunes, se répandent plus rapidement qu'auparavant grâce à l'amélioration des communications, les modes d'utilisation des drogues se propagent au-delà des frontières nationales. La mondialisation de l'abus des drogues fait que les mesures de réduction de la demande doivent être mondiales elles aussi, de même que le système d'information sur lequel elles reposent.

En 1998, à sa vingtième session extraordinaire consacrée à la lutte commune contre le problème mondial de la drogue, l'Assemblée générale des Nations Unies a adopté une déclaration politique (résolution S-20/2, annexe) fixant comme objectif d'éliminer ou de réduire sensiblement l'offre et la demande de drogues illicites d'ici à l'année 2008. C'était la première fois que la communauté internationale se mettait d'accord sur de tels objectifs précis en matière de contrôle des drogues. Toutefois, on ne dispose pas encore des données systématiques qui sont nécessaires pour suivre et évaluer les progrès dans la réalisation de ces objectifs. C'est pourquoi l'Assemblée générale a demandé à l'Office des Nations Unies contre la drogue et le crime (ONUDC) de fournir aux États Membres l'assistance nécessaire pour rassembler des données comparables. L'Office a été prié de recueillir et d'analyser ces données et de les présenter à la Commission des

stupéfiants. Pour répondre à ces demandes, l'Office a lancé le Programme mondial d'évaluation de l'abus de drogues (GAP). Ce programme a pour objet:

- a) D'aider les États Membres à mettre en place les dispositifs nécessaires pour recueillir des données fiables destinées à faciliter l'élaboration et la mise en œuvre de leurs politiques;
- b) D'encourager des partenariats régionaux pour mettre en commun les expériences et les progrès techniques;
- c) De contribuer à une meilleure compréhension des tendances mondiales en matière d'abus des drogues en encourageant l'adoption de méthodes rationnelles permettant de recueillir des données comparables.

Ces objectifs sont l'expression des recommandations formulées en ces termes dans la Déclaration sur les principes fondamentaux de la réduction de la demande de drogues que l'Assemblée générale a adoptée à sa vingtième session extraordinaire (résolution S-20/3, annexe):

“Les programmes de réduction de la demande devraient être fondés sur une évaluation régulière de la nature et de l'ampleur de la consommation et de l'abus des drogues, ainsi que des problèmes y afférents dans la population... Ce sont les États qui devraient s'acquitter de cette tâche, d'une manière complète, systématique et périodique en se fondant sur les résultats des études sur cette question, tout en tenant compte des facteurs géographiques et en utilisant des définitions, des indicateurs et des procédures analogues pour l'évaluation de la situation concernant les drogues.”

Le principal objectif du GAP est d'aider les États Membres à se doter des moyens de recueillir des données pertinentes pouvant être comparées à l'échelon international et d'évaluer l'ampleur et les tendances de l'abus des drogues aux niveaux national, régional et mondial. La mise en place de ces systèmes d'information nationaux et régionaux devrait non seulement contribuer à la création de capacités locales de collecte de données susceptibles d'orienter les activités de réduction de la demande, mais aussi améliorer la communication d'informations aux niveaux transnational, régional et mondial sur les tendances en matière de drogues. Le présent module a été élaboré à l'appui de ce processus pour aider les États Membres de l'Organisation des Nations Unies à mettre au point des systèmes de collecte d'informations sur les drogues qui soient culturellement adaptés à leur contexte national, ainsi que pour faciliter la mise en conformité des systèmes existants d'information sur les drogues avec des normes de bonne pratique internationalement reconnues, et en particulier à l'harmonisation des indicateurs d'abus de drogues. La normalisation des indicateurs et la généralisation de méthodes rationnelles de collecte de données permettront de mieux analyser les tendances en matière d'abus des drogues, tant dans les pays développés que dans les pays en développement. Pour en savoir plus sur le GAP, veuillez consulter son site Web (www.unodc.org), envoyer un message électronique à l'adresse: gap@unodc.org ou vous mettre en rapport avec la Section de la réduction de la demande, Office des Nations Unies contre la drogue et le crime, B.P. 500, A-1400 Vienne (Autriche).

Le rôle de l'estimation de la prévalence

Les deux principales questions que les décideurs posent aux responsables de la collecte de données sur l'abus des drogues sont les suivantes: "Quel est le nombre d'usagers de drogues dans la population du pays et ce nombre varie-t-il?". Connaître le nombre de toxicomanes est utile pour évaluer l'impact probable de l'abus des drogues sur la société et l'ampleur de mesures qui peuvent être nécessaires pour y faire face. Par exemple, si l'on sait combien de personnes s'injectent des drogues, on pourra calculer le niveau des services à assurer pour réduire les comportements liés à l'infection par le virus de l'immunodéficience humaine (VIH) et déterminer si le nombre de places disponibles dans les centres de traitement est suffisant. Comprendre la dynamique du problème de drogues permet non seulement d'en évaluer l'impact probable, mais aussi d'avertir les décideurs d'une détérioration de la situation ou au contraire d'apporter la preuve que la prévention ou les autres mesures qui ont été prises sont efficaces. Dans de nombreux pays, et notamment dans ceux où la drogue est un problème relativement nouveau, il est particulièrement utile de disposer d'une estimation de l'ampleur de ce problème pour attirer l'attention des décideurs et du public sur la nécessité de prendre des mesures et d'investir des ressources pour y faire face.

Pour les épidémiologistes, ces deux questions ont trait à l'estimation de la prévalence et de l'incidence. Ceux-ci emploient, s'agissant d'une maladie ou d'un autre problème de santé, le terme "incidence" pour désigner le nombre de nouveaux cas diagnostiqués, et le terme "prévalence" pour désigner le nombre total de cas actuels (ce nombre comprenant à la fois les cas nouvellement diagnostiqués et ceux qui l'avaient déjà été auparavant). Pour calculer les taux d'incidence et de prévalence, on utilise comme dénominateur une population définie par une aire géographique, une durée ou d'autres caractéristiques précises telles que l'âge, la race ou le sexe.

Bien que la nécessité de disposer d'informations sur l'ampleur du problème de drogues soit évidente, celles-ci sont extrêmement difficiles à obtenir dans la pratique. Bien des pays ne sont pas en mesure d'estimer le nombre de toxicomanes, et, lorsque de telles estimations existent, leur exactitude est souvent douteuse. Diverses raisons expliquent qu'ils se heurtent à des difficultés dans ce domaine et qu'il soit plus difficile d'estimer la prévalence de l'usage de drogues que l'ampleur de tout autre phénomène médical. L'usage de substances psychoactives interdites n'est pas réductible à un modèle de maladie. Bien que les diagnostics de dépendance et d'abus de drogues problématique soient subordonnés à des critères très rigoureux, ceux-ci sont souvent difficiles à appliquer dans la pratique lors d'enquêtes sociales. En outre, un grand nombre d'usagers de drogues, auxquels les décideurs s'intéressent pourtant, n'entrent pas dans ces catégories de diagnostic. La diversité des substances dont il est fait abus, les différentes voies d'administration utilisées et les différentes combinaisons observées pour ce qui est des doses et de la fréquence des prises viennent encore compliquer les choses. En outre, la nature illicite de la toxicomanie et le fait que celle-ci est socialement réprouvée mettent le chercheur en présence de difficultés particulières, qui sont inconnues dans la plupart des autres domaines de l'épidémiologie.

Étant donné la diversité des comportements de consommation de drogues, il faut, lorsqu'on étudie la prévalence de l'usage de drogues, commencer par définir exactement le comportement que l'on veut mesurer. La solution la plus commode consiste habituellement à utiliser des mesures diachroniques de la consommation classées par types de drogues. Des informations supplémentaires sur la consommation de drogues par injection sont aussi nécessaires en règle générale. Les enquêtes sur l'ensemble de la population sont un des moyens couramment utilisés pour recueillir ces informations. Cependant, pour un certain nombre de raisons, ces enquêtes peuvent être très inefficaces pour évaluer certains types d'abus de drogues. En particulier, elles ne permettent pas, bien souvent, d'estimer avec exactitude les comportements rares mais particulièrement néfastes comme l'héroïnomanie, la cocaïnomanie ou la consommation de drogues par injection, et la sous-déclaration peut être un problème. En outre, ces enquêtes sont des opérations techniquement complexes qui exigent beaucoup de ressources et qui sont de toute façon difficilement réalisables dans de nombreux pays en développement. Mais on peut aussi évaluer le nombre de toxicomanes dans chacune de ces catégories en ayant recours aux méthodes d'estimation indirectes, qui constituent le thème principal du présent module du référentiel GAP. Ses auteurs l'ont rédigé en ayant présents à l'esprit les besoins des pays en développement ou en transition, mais il est maintenant largement admis que les méthodes de ce type sont très efficaces pour mieux estimer l'étendue du problème de drogues, même dans les pays où des enquêtes de grande ampleur sur les ménages sont également réalisées.

Le projet GAP

Buts du référentiel GAP

Le présent référentiel a pour but d'offrir un point de départ pratique aux personnes qui souhaitent utiliser des méthodes indirectes pour estimer la prévalence de l'abus des drogues. Ces méthodes, très fréquemment utilisées, sont idéales pour estimer le nombre d'usagers de drogues invétérés ou problématiques, tels que les héroïnomanes ou les usagers par injection qui se droguent sur la voie publique. Les méthodes décrites ici, en effet, s'appuient souvent sur le fait que l'usager de drogues figure – ou a des chances de figurer – dans un registre quelconque: répertoire de police, rapport médical, voire médico-légal. Le référentiel a pour vocation non pas de former un manuel technique complet, mais de poser les principaux problèmes méthodologiques et pratiques qu'il faudra résoudre pour réussir une estimation. Des indications renvoient, pour chaque thème, à des documents techniques plus approfondis. Certaines techniques nécessiteront l'intervention d'un statisticien qualifié, auquel cas le lecteur sera informé des contraintes techniques que présente la méthode choisie. L'une des raisons d'être du référentiel est de donner une idée exacte des moyens techniques qui sont nécessaires pour réaliser une estimation et d'en faciliter ainsi la planification et l'exécution.

Documents complémentaires

Le présent référentiel n'est pas le seul outil de ce type, pas plus qu'il n'a pour vocation d'être la seule source d'information technique utilisée pour planifier une estimation. Lors de son élaboration, on s'est attaché à établir des liens avec certains des documents connexes existants et à les compléter. On trouvera, tout au long du texte, des références à des documents techniques; pour tout renseignement général, le lecteur est invité à se reporter:

- a) Au guide de l'Organisation mondiale de la santé intitulé "*Guide to Drug Abuse Epidemiology*" (2000) (WHO/MSD/MSB 00.3), qui présente un aperçu général de l'épidémiologie des drogues. Ce guide peut être téléchargé à l'adresse Internet suivante: www.who.int/substance_abuse/PDFfiles/EPI_GUIDE_A.pdf;
- b) À la monographie scientifique n° 1 de l'Observatoire européen des drogues et des toxicomanies (OEDT) intitulée "Estimations de la prévalence de la consommation problématique de drogues en Europe" (Luxembourg, Office des publications officielles des Communautés européennes, 1997), qui donne une description plus détaillée des questions relatives à l'estimation de la prévalence;
- c) Aux instruments techniques de l'OEDT qui présentent de manière plus approfondie les méthodes décrites ici. Le rapport d'un groupe de travail de l'Union européenne sur l'estimation de la prévalence, contenant un projet de directives sur l'estimation de la prévalence de la consommation problématique de drogues au niveau national, peut être consulté à l'adresse Internet: www.emcdda.org/situation/themes/problem_drug_use.shtml ou obtenu en écrivant à l'Observatoire européen des drogues et des toxicomanies, Rua da Cruz de Santa Apolónia, 23-25, P-1100 Lisbonne (Portugal).

D'autres documents de recherche et sites Web sont indiqués dans l'annexe.

Questionnaire destiné aux rapports annuels

Questionnaire destiné aux rapports annuels, deuxième partie

Lors du choix des critères à utiliser pour définir la population présentant un intérêt aux fins d'une estimation de la prévalence, il conviendrait d'adopter, chaque fois que possible, des catégories couramment utilisées. Il existe, au niveau mondial, un mécanisme qui permet d'obtenir une vue d'ensemble de la situation en matière d'abus de drogues: le questionnaire destiné aux rapports annuels. Ce questionnaire est le moyen utilisé par les États Membres pour communiquer à la Commission des stupéfiants, conformément aux traités internationaux relatifs au contrôle des drogues, des informations sur divers aspects du problème des drogues illicites. Pour de plus amples informations, se reporter à l'adresse Internet suivante: www.unodc.org.

La deuxième partie du questionnaire est celle qui est la plus utile pour la surveillance des caractéristiques et des tendances de la consommation de drogues à l'échelle

mondiale. Une version révisée de cette partie a été adoptée en janvier 2002 pour tenir compte des indicateurs clefs de la consommation de drogues qui ont été convenus (voir la présentation générale de ces indicateurs dans la section intitulée "Principes applicables à la collecte des données" du site Internet www.unodc.org/pdf/drug_demand_gap_lisbon_consensus.pdf) et faire en sorte que le questionnaire soit suffisamment souple pour que des pays, dont les capacités de collecte de données ne sont pas comparables, puissent y répondre. En particulier, le questionnaire révisé permet de recueillir, à l'échelle mondiale, des données concernant un ensemble convenu d'indicateurs clefs de la consommation de drogues en s'appuyant sur trois niveaux d'information: avis généraux d'experts, données non normalisées ou partiellement chiffrées et données chiffrées normalisées. Bien qu'il soit destiné à ne recueillir que des données récapitulatives, le questionnaire est un bon moyen d'encourager la collecte de données multisources sur un ensemble harmonisé d'indicateurs clefs. Il ne prétend pas répondre à tous les besoins des décideurs, mais peut servir de point de départ à la collecte de données. En adoptant les principaux paramètres du questionnaire, les pays contribuent également à faire en sorte que les activités de collecte de données produisent des informations compatibles avec les normes internationales.

À l'heure actuelle, le bilan de la situation mondiale en matière de drogues est établi sur la base des données recueillies au moyen du questionnaire, que viennent compléter d'autres documents publiés concernant la consommation de drogues. Il fait aussi largement appel aux données fournies aux niveaux national et régional par les systèmes d'information sur les drogues. Le texte du dernier rapport sur la situation mondiale en matière d'abus des drogues peut être consulté à l'adresse Internet suivante: www.unodc.org/cnd_session_45.html.

Catégories courantes d'informations communiquées

Les catégories privilégiées par le questionnaire pour la fourniture d'estimations quantitatives de la consommation de drogues sont indiquées ci-après.

Périodes. Le questionnaire demande, pour chaque drogue répertoriée, des renseignements concernant sa consommation "au moins une fois au cours de la vie" (parfois appelée consommation au cours de la vie) et "au cours des 12 derniers mois". Les autres mesures généralement utilisées sont la consommation "au cours du mois précédent" (parfois appelée consommation actuelle) et la "consommation quotidienne" au cours du mois écoulé, qui correspond généralement, sur le plan pratique, à la "consommation d'une drogue 20 fois ou plus au cours du mois précédant l'entretien".

Types de drogues. Il importe de préciser, dans toute estimation de la prévalence, le type de drogue. Le choix des drogues sur lesquelles on se concentrera dépendra de la nature du problème. La deuxième partie du questionnaire utilise la liste ci-dessous, qui englobe les principaux types de substances dont il est fait abus, mais n'est pas exhaustive. Aussi des cases supplémentaires sont-elles prévues pour mentionner d'autres substances. Les types de drogues pris en considération sont les suivants:

- a) Type cannabis: marijuana (feuille) et hachisch (résine);
- b) Opioïdes: héroïne, opium et autres opioïdes (par exemple morphine, fentanyl et buprénorphine);
- c) Cocaïne: poudre (sels), crack et autres formes de cocaïne;
- d) Type amphétamine: amphétamine, méthamphétamine et amphétamines de type ecstasy;
- e) Sédatifs et tranquillisants: barbituriques et benzodiazépines;
- f) Hallucinogènes: LSD et autres hallucinogènes;
- g) Solvants et inhalants: cette catégorie comprend diverses substances volatiles, telles que le pétrole, l'essence, les adhésifs, les produits aérosol (peintures à pulvériser, désodorisants, analgésiques à pulvériser, par exemple), anesthésiants (protoxyde d'azote, par exemple), produits de nettoyage, solvants et désodorisants d'intérieur (nitrite d'amyle, nitrite de butyle, par exemple);
- h) Autres drogues: toutes substances illicites qui n'entrent pas dans les catégories susmentionnées, comme le gamma-hydroxybutyrate, les stéroïdes anabolisants-androgéniques et le khat.

Voie d'administration. Les conséquences néfastes de l'abus des drogues varient en fonction non seulement du type de drogue consommée, mais aussi de la façon dont cette drogue est consommée (voie d'administration). L'injection de drogues, en particulier, peut avoir des conséquences extrêmement graves, telles que la transmission du VIH et un risque accru de surdose et de dépendance. Les voies d'administration les plus courantes sont: la voie orale, la voie nasale, l'inhalation sous forme de fumée ou de vapeurs ("chasser le dragon") et l'injection.

La plupart des systèmes d'information tentent de différencier les données sur les toxicomanes selon la voie d'administration, en particulier pour les drogues injectables. Ce point étant souvent jugé particulièrement important, des questions distinctes seront posées au répondant pour lui demander s'il a pratiqué l'injection au cours de sa vie, s'il la pratique actuellement et, le cas échéant, s'il partage le matériel d'injection avec d'autres personnes.

Consommation problématique de drogues. Les méthodes décrites dans le présent référentiel ont été conçues principalement pour tenter de mieux estimer les modes "cachés" et problématiques de consommation de drogues qu'il est particulièrement difficile d'observer par d'autres moyens. Il s'agit notamment de la consommation régulière d'héroïne ou de crack et de l'injection de drogues. On utilise parfois, pour estimer la prévalence, une définition composite du mode de consommation combinant en une seule catégorie plusieurs comportements. L'Union européenne, par exemple, produit des estimations de "la consommation problématique de drogues", qui se définit comme "l'usage de drogues par injection ou l'usage de longue durée ou régulier d'opiacés, de cocaïne et/ou d'amphétamines" (pour de plus amples renseignements, consulter l'adresse Internet suivante: www.emcdda.org).

Élaboration d'outils d'évaluation

On espère que le référentiel aidera les chercheurs, lorsqu'ils planifieront leurs estimations de la prévalence, à appliquer des normes reconnues et internationalement acceptées. La structure proposée pourra aider à positionner ces estimations dans différents contextes nationaux, culturels, sociaux et géographiques. Les méthodes et définitions devraient pouvoir convenir à des structures administratives et à des systèmes de collecte de données plus ou moins développés.

Le référentiel a pour but non seulement d'aider à tirer le meilleur parti des données disponibles dans différents pays, mais aussi d'encourager la présentation systématique de ces données sous des formes qui faciliteront le suivi de la prévalence de la consommation de drogues. Tout aussi important: on espère qu'il encouragera la réalisation de nouvelles enquêtes et la mise en place de nouveaux systèmes d'enregistrement de données, qui permettront d'obtenir des résultats coordonnés sur les types et le niveau de consommation de drogues. Ce sont l'élaboration de ces systèmes d'enregistrement à long terme et l'aptitude à réaliser des évaluations ponctuelles ou répétées qui forment, au bout du compte, l'objectif le plus important du projet GAP et du présent référentiel.

Nécessité de principes directeurs et de bonnes pratiques

L'acquisition de l'aptitude à appliquer des méthodes d'estimation de la prévalence de l'abus de drogues est une priorité constante du projet et est le thème récurrent du référentiel. Pour effectuer des analyses techniques, on a souvent besoin de conseils d'experts – dans le cas présent, de statisticiens expérimentés. Le référentiel met en avant la logique et les difficultés de ces méthodes pour permettre au chercheur de déterminer quand il aura besoin de conseils et d'un soutien.

Il existe, cependant, en matière de recherche, une compétence plus générale aussi importante que la maîtrise technique de certaines analyses épidémiologiques – c'est celle relative à la structuration des enquêtes épidémiologiques et des informations requises. C'est là que des principes directeurs reconnus peuvent jouer un rôle utile. Ces principes revêtent la forme de conseils pratiques, non statistiques. Le présent référentiel recommande d'en faire la base de la démarche scientifique et de toute enquête épidémiologique qui pourrait être organisée dans un pays. Le principal objectif est de mettre sur pied, dans les pays mêmes, des mécanismes de soutien entre le personnel et les groupes participant au suivi de l'abus de drogues. Le but est également d'encourager, à ce propos, la consultation internationale dans le cadre d'une structure commune d'idées et de principes. De ce point de vue, le développement de mécanismes de consultation dans le cadre d'un ensemble convenu de normes devrait être considéré comme aussi important que l'élaboration d'un langage scientifique commun aux différents groupes participant à ce travail.

Les questions susmentionnées sont traitées dans une publication de l'Observatoire européen des drogues et des toxicomanies disponible sur Internet (www.emcdda.org/multimedia/project_reports/situation/guidelines_prevalence_pdu_mb_09-01.pdf), à laquelle les auteurs du présent référentiel doivent beaucoup.

Évaluation de la prévalence de l'usage de drogues

Chapitre premier

Le présent module offre un aperçu de certaines des méthodes mises au point pour surmonter la difficulté d'estimer la prévalence de drogues par des techniques d'enquête standard. Ces méthodes, très générales, peuvent servir à estimer la prévalence de l'usage de n'importe quelle drogue, mais le contexte décrit ici est celui de l'abus chronique de drogues, tel l'usage illicite d'opiacés. Dans ce contexte, leur principale justification tient au fait que l'usage illicite d'opiacés est relativement rare si l'on considère le nombre d'utilisateurs en pourcentage de la population, qu'il s'agit d'un comportement réprouvé et qu'en raison de sa nature illicite il ne peut exister, pour ce dernier, de système national d'enregistrement.

Les techniques d'enquête standard appliquées à la population générale conviennent généralement mal à ces situations. Des techniques telles que les méthodes multiplicateur/référence et par capture-recapture peuvent permettre d'estimer la prévalence de l'usage de drogues lorsque les méthodes standard échouent; elles sont proposées ici comme solution de substitution aux enquêtes sur l'ensemble de la population.

Évaluation de la prévalence et nécessité d'évaluer le dommage

Il existe, de la part des responsables politiques, un besoin et une demande croissants d'estimations de la prévalence leur permettant d'orienter leur action (voir la monographie scientifique n° 1 de l'OEDT, "Estimations de la prévalence de la consommation problématique de drogues en Europe"). Il en a résulté la publication d'un nombre croissant de manuels et d'analyses de méthodes (OEDT 1997 et 2000; Hickman *et al.*, 2002; International Working Group for Disease Monitoring and Forecasting, 1995; Frischer *et al.*, "Mortality among injecting drug users...", 1993; Hser *et al.*, 1992; et Reuter, 1993). Ce qui différencie le présent module des autres, c'est qu'il se focalise sur les pays en développement et qu'il forme un manuel rassemblant des exemples et des méthodes applicables à ces pays.

Des estimations de prévalence (concernant, par exemple, l'usage d'héroïne et l'abus de drogues par injection) sont nécessaires aux fins de l'élaboration des politiques, notamment pour les tâches suivantes:

- a) Planification et affectation des moyens de réduction, de traitement et de prévention de la consommation problématique de drogues et de ses conséquences. Il faut souvent, pour justifier une augmentation de crédits, produire une estimation de la prévalence;
- b) Suivi d'objectifs clefs de la politique en matière de drogue, par exemple en surveillant le pourcentage d'utilisateurs problématiques de drogues suivant un traitement ou le degré de couverture d'une activité de réduction des dommages. Une façon de mesurer ces objectifs est d'évaluer le nombre de personnes en contact avec les services de santé par rapport, notamment, à la population toxicomane totale estimée;
- c) Interprétation des principaux problèmes liés à l'usage de drogues. La prévalence des infections à VIH et au virus de l'hépatite C (VHC), des décès par surdose et des délits associés à l'usage de drogues dans la population générale est en rapport avec les risques encourus et les comportements adoptés par les utilisateurs problématiques de drogues, ainsi qu'avec leur nombre.

Pour les responsables politiques, en supposant qu'ils suivent de près les comportements qui nuisent gravement à la santé publique, l'essentiel est de trouver une méthode qui permette d'estimer le nombre d'héroïnomanes, d'utilisateurs de drogues par injection ou de consommateurs de crack. Tel est l'objet du présent référentiel.

Celui-ci comporte des omissions notables et délibérées. L'intention est de regrouper des méthodes indirectes d'estimation de la consommation problématique de drogues et des problèmes connexes. Sont exclues, en particulier, les enquêtes scolaires, qui évaluent directement le problème au début de la vie. Les enquêtes spécialisées de ce type, qui portent sur certains sous-ensembles de la population, font l'objet de différents modules GAP.

Estimations de prévalence et rôle des enquêtes sur la population générale

Il est largement admis que les estimations de prévalence établies à partir d'enquêtes sur la population générale, c'est-à-dire, en règle générale, d'informations sur les habitudes de consommation de drogues obtenues par autodéclaration auprès d'un échantillon de ménages, sont inadaptées pour évaluer l'ampleur de l'usage illicite de drogues, en particulier d'héroïne. Les méthodes d'enquête classiques conviennent généralement à un nombre restreint de questions, pour lesquelles on peut considérer que les biais et erreurs de réponse et les fausses déclarations n'influent que faiblement le taux de prévalence global. Pour ce qui est des drogues, cela veut dire

que des questions concernant la prévalence, par exemple, de l'usage de cannabis, du tabagisme, voire de la consommation d'alcool au cours de la vie peuvent être envisagées dans le cadre d'enquêtes sur la population générale. La prévalence de ces comportements s'estime en ajoutant – superposant – des questions aux questionnaires déjà utilisés dans les enquêtes sur la population générale. Ces questions ajoutées visent souvent à déterminer, par exemple, des taux de prévalence du tabagisme ou de la consommation fréquente d'alcool sur une année.

L'inadéquation des méthodes d'enquête standard pour estimer la prévalence de l'abus chronique de drogues tient principalement à deux éléments:

- a) Sous-couverture: les toxicomanes chroniques sont nombreux hors des ménages; ne sont donc pas couverts certains groupes clefs de toxicomanes non inclus dans les cadres d'échantillonnage des enquêtes: sans-abri, personnes arrêtées et, éventuellement, personnes traitées en institution;
- b) Sous-déclaration: des usagers de drogues peuvent simplement refuser de répondre ou répondre par la négative comme l'ont montré, par exemple, de récents travaux comprenant une autodéclaration et des analyses simultanées de cheveux; plus les comportements sont répréhensibles, plus le problème est exacerbé: on pense, par exemple, que la consommation d'héroïne est plus souvent sous-déclarée que celle de marijuana.

Ces facteurs réduisent le taux, déjà faible en termes relatifs, d'abus chronique de drogues, ce dont témoignent les faibles chiffres de prévalence donnés, par exemple, dans les enquêtes sur la criminalité et la morbidité psychiatrique du Royaume-Uni de Grande-Bretagne et d'Irlande du Nord et par l'enquête nationale auprès des ménages sur l'abus de drogues aux États-Unis d'Amérique. Il est certes possible d'appliquer des facteurs de correction en utilisant des techniques d'estimation par quotient, telles celles décrites dans certains rapports du National Institute on Drug Abuse et examinées ci-après. Généralement, cependant, une estimation nécessite, pour être efficace, des procédures plus spécifiquement adaptées au contexte de l'usage de drogues.

Méthodes de correction de la sous-estimation dans les enquêtes sur la population générale

Il a été noté que des facteurs liés à l'usage de drogues, tels que les données d'arrestation et de traitement de la toxicomanie, étaient également sous-déclarés ou insuffisamment couverts dans les enquêtes sur la population générale. Lorsqu'on dispose de données nationales concernant ces facteurs, on peut s'en servir pour corriger les chiffres des enquêtes en pondérant le nombre de répondants signalant une arrestation ou un traitement pour atteindre ce que l'on sait être le chiffre correct. Cette utilisation des chiffres nationaux comme référence pour corriger une sous-estimation dans l'analyse de données d'enquêtes nationales est un exemple type d'application de l'estimation statistique dite "par quotient".

L'intérêt de cette pondération visant à tenir compte des comportements en matière de drogues est qu'elle pondère aussi automatiquement la prévalence concomitante de l'abus de drogues. Cette procédure est à appliquer avec prudence, car l'ajustement des chiffres de la prévalence de drogues découle d'une méthode conçue pour ajuster les facteurs connexes, non les chiffres de la prévalence eux-mêmes. Les résultats doivent être considérés davantage comme une amélioration que comme une correction des estimations originales.

Prévalence de l'usage de drogues et autres méthodes épidémiologiques

Ce qui caractérise les populations d'usagers de drogues, c'est leur difficulté d'accès. Ce sont des populations qui se "cachent" des institutions officielles et échappent à leurs registres; en raison du risque de stigmatisation et de l'illégalité, de surcroît, elles se "cachent" également dans l'inexactitude des réponses qu'elles peuvent donner. Habituellement, l'échantillonnage classique consiste à utiliser des registres de la population générale pour former le cadre d'échantillonnage dont on extraira un échantillon. Lorsqu'on abandonne, en raison de son inadéquation, l'outil du statisticien qu'est le cadre d'échantillonnage, il ne reste, pour évaluer la prévalence, que quelques options. Nous examinons brièvement, ci-dessous, certaines d'entre elles.

Sondage aréolaire

Certains sondages de population ne nécessitent pas de cadre d'échantillonnage. C'est le cas du sondage aréolaire, dans lequel on divise de façon conceptuelle une région ou un pays en zones opérationnelles, généralement de même taille, dont on extrait un échantillon. Pour estimer la taille de la population globale ou celle d'un sous-ensemble défini tel que les usagers problématiques de drogues, les chercheurs comptent alors physiquement, dans chaque zone, les "cas" répondant aux critères définis (dans certaines variantes de cette méthode, on n'utilise pas des "zones", mais des "transects linéaires").

Dans certaines situations, c'est souvent la seule méthode utilisable; parfois, elle peut, en outre, être efficace. Pour réaliser un comptage physique, il faut pouvoir aisément identifier les cas: on ne peut se permettre, dans la pratique, de recourir à des procédures longues et compliquées pour identifier un cas répondant aux critères définis. Dans une enquête sur les usagers problématiques de drogues, cela n'est pas une mince affaire. Il faut, en outre, que les structures sociales du pays conviennent à cette méthode. On peut citer, comme exemple, les enquêtes nationales sur les toxicomanes réalisées autrefois au Pakistan, dans lesquelles il était demandé aux mollahs ou chefs de village d'identifier les toxicomanes locaux.

Populations spécifiques

Des chercheurs ont tenté de résoudre certains des problèmes de sous-couverture inhérents au calcul de la prévalence de la consommation de drogues en utilisant, lors de la conception de sondages, des cadres d'échantillonnage spécialisés tels que les registres de services d'urgence ou de police. Généralement, cependant, ceux-ci ont une couverture trop étroite pour que l'on puisse en tirer des conclusions générales concernant la consommation problématique de drogues.

Il est une exception notable, à savoir lorsqu'un groupe fait lui-même l'objet d'une enquête particulière, tels les écoliers. Des enquêtes nationales – et, en fait, internationales –, telles que le Projet européen d'enquêtes scolaires sur l'alcool et d'autres drogues (ESPAD), peuvent être considérées comme couvrant raisonnablement la population cible, surtout lorsque des dispositions sont prises pour englober les écoliers absentéistes.

Il convient de noter que des répertoires évidents d'usagers de drogues tels que les registres des centres de traitement, des programmes d'échange d'aiguilles ou autres, qui ne peuvent servir de cadre d'échantillonnage pour une enquête de prévalence car ils ne couvrent que des usagers de drogues, sont des sources d'information que les méthodes d'estimation indirecte ont vocation à utiliser. Des techniques spéciales sont généralement requises pour en tirer le meilleur parti.

Méthodes indirectes d'estimation de la prévalence

C'est précisément pour tirer parti des listes existantes d'usagers de drogues et de comportements liés à la drogue que l'on a recours aux méthodes indirectes d'association. Ces méthodes consistent à reconnaître d'abord les insuffisances des registres et autres sources de données disponibles. Elles tiennent compte du fait que les populations visées peuvent n'être que partiellement observées, et qu'aucun de ces registres ou de ces sources ne recouvre complètement la population des usagers problématiques de drogues.

Néanmoins, les méthodes indirectes permettent d'entreprendre de compter les usagers problématiques de drogues à partir de ces listes incomplètes. On utilise parfois, parallèlement à la collecte de données primaires, des études complémentaires de la population visée, en tentant de tenir compte des difficultés que pose l'échantillonnage des usagers de drogues eux-mêmes.

Les méthodes d'analyse décrites dans le présent module sont des méthodes à simple multiplicateur et par capture-recapture. Il est en outre donné une description des méthodes un peu plus complexes que sont les modèles fondés sur le taux d'occurrence, ainsi qu'un bref exposé sur l'établissement par extrapolation, à partir d'études de prévalence locales, d'estimations de la prévalence au niveau national.

Fiabilité et recours à plusieurs méthodes

Les méthodes d'estimation indirectes peuvent être très peu fiables. De surcroît, les estimations obtenues reposent sur des hypothèses souvent invérifiables qui, si elles sont erronées, peuvent produire des biais tout aussi graves que dans des enquêtes de population.

Certaines des méthodes d'enquête standard reposant sur l'estimation par intervalle de confiance peuvent être utilisées lorsque les hypothèses des méthodes indirectes sont valables, mais il ne peut pas être remédié au manque de solidité de ces méthodes par les méthodes habituelles d'estimation par intervalle de confiance.

Il est courant, par conséquent, de rechercher les concordances et convergences qui se dégagent d'estimations obtenues par différents moyens indirects. Pour une bonne part, la fiabilité de ces méthodes ne se juge que par la mesure dans laquelle elles convergent vers une estimation commune plausible.

Enquêtes de prévalence nationales et locales

Des méthodes indirectes peuvent théoriquement être utilisées au niveau national pour déterminer une prévalence dans la population générale; les méthodes multiplicateur/référence, en particulier, sont parfois utilisées de cette façon. Plus généralement, cependant, elles sont utilisées à une échelle géographique plus restreinte où elles sont plus faciles à mettre en œuvre pour tirer parti des données disponibles localement et tenir compte des variations locales.

On ne saurait souligner assez la nécessité d'estimer, au niveau local, la prévalence de drogues et les nuisances connexes, les variations locales en la matière étant très importantes. Cela dit, il demeure souvent nécessaire d'effectuer des estimations nationales, ce qui peut se faire par extrapolation à partir d'études de prévalence locales. Ce faisant, on tente de prédire la prévalence dans des zones où il n'a été réalisé aucune enquête locale en les comparant à des zones où la prévalence est connue – ou, plutôt, a été estimée. Pour effectuer ces comparaisons, il faut donc disposer, en l'absence de chiffres réels concernant la prévalence de l'abus de drogues, de données apparentées que l'on appelle "indicateurs" de l'abus de drogues.

Si l'on peut organiser ces "indicateurs" de façon qu'ils couvrent l'ensemble d'un pays tout en continuant d'être disponibles au niveau local, on peut alors extrapoler à l'ensemble du pays, zone par zone, les estimations de prévalence locales. Les prévisions qui peuvent être faites pour d'autres localités sont en elles-mêmes utiles, mais il est aisé de les combiner pour estimer également la prévalence au niveau national.

Ces méthodes d'extrapolation, qui reposent sur des techniques de régression statistique, sont parfois dites "méthodes à indicateurs multiples" ou "estimation par synthèse".

Méthodes spécifiques d'estimation de la prévalence

Chapitre II

Aperçu général

Les principales méthodes sont esquissées ici pour permettre au lecteur de repérer les chapitres du module qui l'intéressent et d'ignorer le reste. Ces méthodes d'analyse sont toutes destinées à produire des estimations de prévalence concernant des populations essentiellement cachées, du moins pour ce qui est de figurer dans un cadre d'échantillonnage. La présente section décrit ainsi brièvement la façon dont les méthodes diffèrent et les données dont il faut disposer pour pouvoir les appliquer. L'attention du lecteur est appelée sur les principes directeurs de l'Observatoire européen des drogues et des toxicomanies (OEDT) relatifs aux méthodes indirectes (www.emcdda.org/multimedia/project_reports/situation/guidelines_prevalence_pdu_mb_09-01.pdf).

Méthodes multiplicateur/référence

Ces méthodes consistent à utiliser des données préexistantes – généralement nationales – concernant un comportement ou un événement commun à la population d'usagers problématiques de drogues: données d'arrestation pour usage ou possession de drogues, statistiques d'accidents ou de services d'urgence et, plus directement, données sur les traitements ou les décès liés à la drogue. Ces informations préexistantes, qui peuvent être simplement un décompte anonyme du comportement clef sur une période de temps donnée, sont appelées informations de référence. Parallèlement à ces données nationales, il faut disposer d'une estimation du pourcentage de la population cible ayant connu cet événement, c'est-à-dire ayant été arrêtée, étant décédée, etc.; l'inverse de ce pourcentage est appelé le multiplicateur. Pour estimer le multiplicateur associé, il faut généralement réaliser une petite sous-enquête distincte; là encore, des registres anonymes suffisent le plus souvent.

Un document publié par Hartnoll *et al.* ("Estimating the prevalence of opioid dependence", *Lancet*, vol. 338 (1985), p. 203 à 205) décrit

l'application de la technique la plus simple, qui utilise le nombre de décès recensé chez les usagers de drogues. Pour estimer le nombre d'usagers de drogues sur une année donnée, l'auteur utilise deux éléments:

- a) Le nombre de décès d'usagers de drogues recensés cette année-là, par exemple 3 000, qui constitue la "référence";
- b) Le taux de mortalité recensé cette année-là chez les usagers de drogues, par exemple 2 % sur l'année ou une personne sur cinquante décédant dans l'année, qui constitue le multiplicateur.

À partir de ces deux chiffres, on estime que le nombre d'usagers de drogues au cours de cette année-là est la population requise pour qu'un taux de mortalité de 2 % entraîne 3 000 décès. Si une personne sur cinquante meurt, la population totale doit alors avoir été de $3\ 000 \times 50$, soit 150 000. Ce calcul est remarquable par sa simplicité et par son caractère direct.

Le pourcentage de la population cible dans la référence peut être obtenu séparément et indépendamment en réalisant des entretiens/interrogatoires ou d'autres études spécifiques. Parfois, il est possible d'utiliser des chiffres provenant de données déjà publiées s'ils correspondent à la population cible, voire d'une enquête sur la population générale si elle rassemble un grand nombre d'usagers de drogues appartenant à la population cible. Il existe différents types d'enquêtes par multiplicateur: enquêtes par désignation, à multiplicateur de mortalité, à multiplicateur de traitement, etc.

Enquêtes par capture-recapture

Les enquêtes par capture-recapture utilisent également des listes préexistantes, en particulier des listes de la population cible d'usagers de drogues et des listes dans lesquelles les individus peuvent être identifiés (par nom, numéro d'identité ou autrement). Ces listes, bien entendu, sont des listes incomplètes de la population cible, ce que la méthode compense en utilisant plusieurs listes de ce type, par exemple une liste d'usagers de drogues arrêtés et une liste d'usagers de drogues traités.

Les méthodes par capture-recapture ont été utilisées au départ pour estimer la taille de populations animales; la terminologie a persisté, de façon abusive, tandis que les méthodes ont migré vers le contexte sociologique, en particulier vers l'estimation de la prévalence de l'usage de drogues. Le terme "capture" désigne le fait que l'usager de drogues est inscrit dans un centre de traitement ou figure dans des registres d'arrestation, par exemple, ce qui permet d'établir une "liste de capture" d'individus identifiés. Le terme "recapture" correspond alors au fait d'apparaître sur deux listes de capture (ou plus) lorsqu'on recoupe les registres. Il importe, à ce propos, d'admettre que les listes elles-mêmes ne recouvrent presque jamais l'ensemble de la population cible; le calcul capture-recapture donne le nombre d'usagers de drogues absents de l'une ou de plusieurs des listes.

L'un des premiers exemples de cette méthode a été fourni par Ghodse (1980), lequel utilisait deux sources de données officielles britanniques. L'index des toxicomanes du Ministère de l'intérieur britannique était un registre d'usagers de drogues de 1^{re} catégorie, qui avaient, d'une manière ou d'une autre, attiré l'attention des autorités médicales; il était reproché à ce registre de ne susciter qu'une réponse médiocre. À partir du registre des décès britannique, il a été possible de recenser les décès liés à la toxicomanie, présumés être des décès de toxicomanes. La première de ces listes de capture était elle-même considérée comme représentant de façon incomplète le nombre total de toxicomanes "notifiés à l'index des toxicomanes du Ministère de l'intérieur britannique". La seconde liste, par recoupement avec la première, détermine, parmi les décès liés à la drogue, le pourcentage de toxicomanes notifiés à l'index du Ministère de l'intérieur. En supposant que ce taux de notification soit le même chez les toxicomanes vivants, il indique le pourcentage requis pour déterminer la mesure dans laquelle la liste du Ministère de l'intérieur sous-estime le nombre total de toxicomanes britanniques.

Cette méthode est généralement étendue pour englober deux listes (ou plus) d'usagers de drogues déduites ou construites indépendamment, qui peuvent ne pas toujours être ordonnées chronologiquement comme elles l'étaient dans les études animales initiales. Le principe général est que les données officielles – les listes systématiques d'usagers de drogues – recouvrent toujours de façon incomplète cette population. Le but est de calculer la mesure dans laquelle une liste sous-estime la population d'usagers de drogues.

Il faut alors, pour recouper les listes utilisées, que les cas soient identifiables, de façon à pouvoir déterminer la proportion de concordances directement à partir des sources de données utilisées dans l'enquête. Cette méthode ne nécessite aucun entretien ni aucune étude spécialisée de la population cible, et les listes de capture utilisées peuvent couvrir cette population de façon incomplète.

Extensions et modèles avancés fondés sur le taux d'occurrence

Les méthodes fondées sur le taux d'occurrence d'événements reposent initialement sur un ensemble de méthodes analogues à celles du multiplicateur, et les "événements" sont généralement les mêmes contacts avec l'institution que ceux utilisés dans les enquêtes multiplicateur/référence ou par capture-recapture. Les deux types d'information qu'il faut obtenir sont:

- a) La fréquence globale, parmi tous les usagers de drogues (autant ceux qui sont en contact que ceux qui ne le sont pas), à laquelle des contacts se produisent;
- b) Le nombre de contacts pris avec l'institution.

Ces informations permettent de déduire le nombre total d'usagers de drogues qui étaient actifs pendant la période de collecte de données. Ce nombre est simplement la somme d'événements d'un type donné divisée par la fréquence à laquelle ces événements se produisent.

Estimer la fréquence des contacts n'est pas aisé, car ce que l'on veut savoir, en particulier, c'est le pourcentage d'usagers de drogues qui ne prendront pas contact. L'astuce, bien entendu, consiste à parvenir à calculer la fréquence en n'interrogeant que des usagers de drogues qui sont en contact. S'il est possible d'émettre des hypothèses solides concernant la fréquence relative à laquelle un usager de drogues noue de multiples contacts, il est possible d'estimer, à partir du schéma de contacts répétés, quel pourcentage d'usagers de drogues ne prend jamais contact.

Dans la pratique, la procédure d'estimation est plus compliquée que la formule à simple multiplicateur, car le schéma de contacts multiples étudié peut être complexe et parce que ces méthodes prévoient généralement, pour différents groupes d'usagers de drogues, différentes fréquences de contact. L'intérêt de cette méthode, cependant, est qu'elle permet d'effectuer des déductions quant à la taille de la population totale d'usagers de drogues en interrogeant des personnes rencontrées à quelques endroits soigneusement sélectionnés.

Difficultés liées au choix de la méthode

Outre les données spécifiques dont il faut disposer pour chacune de ces méthodes, d'autres considérations pratiques influencent le choix du type d'enquête. La première a trait à l'échelle – ou portée – de l'enquête. Bien que toute méthode puisse s'appliquer à une population locale ou nationale d'usagers de drogues, les enquêtes par capture-recapture sont souvent plus faciles à mettre sur pied à l'échelon local, ne serait-ce que pour confronter les identificateurs. Une autre considération a trait au fait de savoir si l'on dispose de ressources suffisantes pour réaliser une enquête spécialisée afin d'estimer le multiplicateur. En fonction des données officielles dont on dispose, des enquêtes de mortalité peuvent parfois durer longtemps; s'il existe des sources convenables contenant déjà l'information, celles-ci offrent, faute de mieux, la meilleure solution. Un troisième facteur est le fait de savoir s'il est possible de réaliser des entretiens – généralement nécessaires à l'obtention de renseignements précis – ou si les registres existants suffisent. Pour ce qui est de l'exactitude et de la fiabilité, toutes ces méthodes d'estimation ont un haut degré d'incertitude; il est préférable de disposer d'une vaste enquête dans laquelle le nombre de toxicomanes observés est le plus important possible et celui de ceux ignorés aussi faible que possible. Il est généralement préférable de posséder, sur les usagers de drogues, trop d'informations que pas assez: des informations appropriées concernant, notamment, le sexe et l'âge peuvent permettre d'améliorer les estimations dans tous les cas.

Extrapolation et estimation par synthèse

Les sections décrivant certaines méthodes comprennent une section sur l'extrapolation à certaines régions de la prévalence de l'usage de drogues connue dans d'autres

régions. Généralement, cette méthode sert à estimer, par généralisation, la prévalence nationale à partir d'enquêtes locales.

Les méthodes d'extrapolation ne sont pas véritablement, en soi, une méthode d'estimation de la prévalence; en principe, cependant, elles remplissent la même fonction lorsqu'on dispose de données concernant certaines régions. En règle générale, ces méthodes relèvent de l'"estimation par synthèse", même s'il ne s'agit, techniquement, de rien d'autre que de procédures de régression statistique standard. La méthode parfois connue sous le nom d'estimation "à indicateurs multiples" suit aussi, fondamentalement, le même principe et est utilisée dans des circonstances analogues.

Le principe essentiel de l'estimation par synthèse et de toute autre méthode d'extrapolation consiste à utiliser les chiffres de prévalence connus de certaines régions pour estimer ceux d'autres régions. Pour ce faire, les régions dites "cibles" doivent, même si elles ne disposent pas, bien entendu, de chiffres de prévalence proprement dits, disposer de données qui soient identiques (ou très semblables) à celles des régions dites "d'ancrage", pour lesquelles il existe effectivement des estimations de prévalence. Ces données sont appelées indicateurs de drogues: ce sont des mesures qui sont liées à la prévalence de l'abus de drogues mais qui ne peuvent, en elles-mêmes, fournir à ce sujet de chiffre satisfaisant. C'est la comparaison des régions cibles et d'ancrage sur le plan de ces variables qui permet de généraliser les prévalences connues.

Il s'agit donc essentiellement d'une procédure à forte intensité de données qui utilise des données disponibles dans une subdivision géographique du pays.

Le rôle des études de cas dans les sections ci-après

Le référentiel dans son ensemble et les principes directeurs qu'il présente s'appuient sur des exemples concrets provenant de la littérature scientifique existante. Bien qu'il soit difficile d'isoler certaines méthodes, les exemples ont été répartis dans les différents chapitres, de façon qu'ils se trouvent là où ils sont le plus pertinents. Ces exemples et commentaires constituent près de la moitié du document; le reste est un important travail regroupant les divers éléments des méthodes d'estimation indirecte et soulignant les liens qui existent entre eux.

Les méthodes décrites sont toutes illustrées, dans chaque section, par des études de cas. Ces études forment une partie essentielle de la description des méthodes et donnent de nombreux renseignements pratiques ainsi qu'un aperçu théorique de leur mise en œuvre. Il a été choisi, autant que possible, des études réalisées hors du cadre européen, dans des régions où les données et les installations de recherche sont loin d'être parfaites. Elles ont également été choisies pour illustrer l'application de méthodes d'estimation dans des circonstances difficiles; elles signalent

donc les faiblesses et les approximations des méthodes et ne doivent pas être considérées elles-mêmes comme parfaites.

Les méthodes multiplicateur/référence

De toutes les méthodes d'estimation indirecte, la méthode multiplicateur/référence est probablement la plus simple à appliquer et celle utilisée depuis le plus longtemps en épidémiologie des drogues. Sa souplesse d'application la rend utile dans bien des circonstances. Elle consiste normalement, à partir d'informations concernant la taille connue d'un sous-groupe bien défini de la population d'usagers de drogues visée, à établir, par généralisation, une estimation de l'ensemble de cette population en utilisant un coefficient multiplicateur.

Après avoir présenté d'abord un exemple simple d'application de cette méthode, on examinera les atouts et les inconvénients qu'elle présente dans un contexte plus large.

Technique de multiplication simple

La méthode multiplicateur/référence suppose que l'on dispose de certaines informations concernant un sous-groupe de la population cible – généralement le nombre de toxicomanes qui prennent contact avec un organisme particulier –, à partir duquel on essaye d'estimer le nombre de toxicomanes supplémentaires dans l'ensemble de la population cible. Si, par exemple, on sait combien de toxicomanes suivaient un traitement en 2001 et que l'on sait qu'environ un toxicomane sur dix a suivi un traitement en 2001, on peut alors “multiplier” le nombre de toxicomanes traités par dix pour obtenir une estimation du nombre total de toxicomanes. Ces deux éléments – le chiffre connu de toxicomanes en traitement (la référence) et la proportion estimative de toxicomanes en traitement (permettant de calculer le multiplicateur) – ont donné leur nom à cette méthode.

Les difficultés que comporte cette méthode et leurs conséquences sont examinées en détail. Son avantage réside dans la grande simplicité du principe de calcul et dans le fait qu'elle peut être appliquée à un large éventail de données. On peut, bien sûr, utiliser d'autres groupes de référence que celui des “personnes en traitement”, si l'on connaît le chiffre de référence et le multiplicateur correspondant. On trouvera, au tableau 1, une liste de sources de données susceptibles de servir de référence pour estimer la prévalence.

L'étude de cas ci-après illustre les calculs de base nécessaires pour estimer le nombre d'usagers par injection à Toronto (Canada) en 1996 au moyen de la méthode multiplicateur/référence.

Tableau 1. Sources de données susceptibles de servir de référence pour estimer la prévalence de l'usage problématique de drogues au moyen de méthodes de multiplicateur/référence

Source de données	Exemple
Centres de traitement spécialisés	Usagers de drogues sous méthadone recevant des soins en ambulatoire ou en établissement
Services à bas seuil	Usagers de drogues fréquentant des centres d'accueil "portes ouvertes" ou contactés par des agents de proximité
Programme d'échange de seringues	Usagers de drogues inscrits dans des programmes d'échange de seringues
Services des urgences	Usagers de drogues admis aux urgences à la suite d'une surdose
Laboratoires	Usagers de drogues ayant subi des tests de dépistage du VIH, du VHC ou du virus de l'hépatite B (VHB)
Police/Prisons	Usagers de drogues arrêtés ou incarcérés pour faits de drogues ou pour d'autres infractions
Sursis probatoire	Usagers de drogues en sursis probatoire
Évaluations des services sociaux	Usagers de drogues ayant fait l'objet d'une évaluation par les services sociaux locaux
Foyers pour usagers de drogues	Usagers de drogues vivant dans des foyers
Registres de toxicomanes	Usagers de drogues inscrits dans un registre central
Enquêtes sur les usagers de drogues à problème	Enquêtes locales sur les usagers de drogues
Décès par surdose	Nombre de décès dus à une surdose d'opiacés

Source: Hickman *et al.*, "Estimating the prevalence of problematic drug use: a review of methods and their application", *Bulletin on Narcotics* (publication des Nations Unies, à paraître).

Étude de cas n° 1. Étude réalisée à Toronto au moyen de la méthode du multiplicateur à partir de tests de dépistage du VIH (Calcul multiplicateur/référence simple)

Archibald *et al.* (2001) ont présenté une méthode multiplicateur/référence, permettant d'estimer la prévalence de l'usage de drogues par injection à partir d'informations fournies par des laboratoires sur le nombre de tests de dépistage du VIH par usager injecteur, ainsi que des données d'enquête concernant la proportion de ces usagers ayant subi un test de dépistage du VIH pendant une année donnée. Les résultats concernant Toronto pour l'année 1996 sont examinés plus loin. L'emploi de cette méthode suppose que l'on dispose de deux éléments, le premier étant un chiffre de référence connu, en l'occurrence le nombre de tests de dépistage du VIH réalisés en 1996 sur des usagers injecteurs à Toronto. Ce nombre, qui a été enregistré lors d'activités systématiques de collecte d'informations, s'établit à 4 050. Il représente la partie connue de la population d'usagers.

Pour déterminer le nombre total d'usagers injecteurs, il faut savoir quelle est la proportion de ces usagers qui ne figurent pas dans les registres de dépistage du VIH. Le second élément nécessaire pour appliquer cette méthode est donc un multiplicateur indiquant combien d'usagers supplémentaires à Toronto n'ont pas subi de tests de dépistage du VIH en 1996. Ce chiffre peut être simple à calculer si l'on détermine la proportion d'usagers de drogues qui ont fait ces tests de dépistage du VIH au cours de cette période. Dans cet exemple, on savait, d'après d'autres études, que 25 % des usagers – soit un sur quatre – avaient fait des tests. Comme l'indique le tableau 2 ci-dessous, le calcul est simple: si un usager sur quatre a fait le test, le nombre total d'usagers doit être de $4 \times 4\,050$, soit 16 200.

La méthode suppose que l'on dispose d'une estimation non biaisée du multiplicateur. L'idéal serait que cette estimation soit obtenue à partir d'un échantillon

représentatif d'usagers de drogues à problème, dans un intervalle de temps et à un emplacement géographique correspondant exactement à ceux du chiffre de référence qui sera utilisé. En pratique, cela arrive rarement. Dans cette étude de cas sur Toronto, les auteurs ont utilisé un multiplicateur provenant d'une étude effectuée dans une autre ville et ont supposé que ce multiplicateur serait identique à Toronto en 1996.

Tableau 2. Estimation du nombre d'usagers injecteurs à partir du nombre de tests de dépistage du VIH

Élément	Valeurs utilisées	Estimations
Référence (R)	Nombre de tests de dépistage subis par des usagers injecteurs en 1996 ^a	4 050
	Proportion d'usagers injecteurs déclarant avoir subi un test de dépistage du VIH au cours de l'année écoulée ^b	25 %
Multiplicateur (M)	Multiplicateur calculé; établi à: 1,0 / 0,25 (soit un sur quatre)	4,0
Estimation de la population Laboratoires	Référence x multiplicateur (R * M)	16 200
Police/Prisons	Usagers de drogues arrêtés ou incarcérés pour faits de drogues ou pour d'autres infractions	
Sursis probatoire	Usagers de drogues en sursis probatoire	
Évaluations des services sociaux	Usagers de drogues ayant fait l'objet d'une évaluation par les services sociaux locaux	
Foyers pour usagers de drogues	Usagers de drogues vivant dans des foyers	
Registres de toxicomanes	Usagers de drogues inscrits dans un registre central	
Enquêtes sur les usagers de drogues à problème	Enquêtes locales sur les usagers de drogues	
Décès par surdose	Nombre de décès dus à une surdose d'opiacés	

Source: Archibald *et al.* (2001).

Note: Pour des raisons de présentation et de clarté, les chiffres donnés dans le tableau à titre d'exemple ont été légèrement modifiés par rapport à ceux de la publication dont ils sont tirés.

^aDéterminé à partir de rapports de laboratoires.

^bDéterminée à partir d'enquêtes locales sur les usagers de drogues par injection.

Il est intéressant de noter que l'étude de cas n° 1 utilise des données standard, relevées de façon systématique – en l'occurrence, le nombre d'usagers injecteurs ayant subi des tests de dépistage du VIH au cours de l'année écoulée, communiqué par des sources officielles –, comme chiffre de référence pour la partie "officiellement visible" de la population d'usagers injecteurs, et que, pour déterminer la population totale à partir de cette partie officiellement visible, elle se fonde sur des informations qui figurent dans d'autres études publiées. Dans cet exemple, il n'a donc pas été nécessaire d'effectuer de nouvelles recherches pour réunir des informations, puisque celles-ci étaient déjà disponibles sous une forme ou une autre.

Bien sûr, comme indiqué dans l'étude de cas, on a beaucoup transigé sur la précision en supposant que le coefficient multiplicateur de 4 pouvait s'appliquer à Toronto, alors qu'il avait été établi pour un autre lieu et une autre période. La deuxième étude de cas, présentée ci-dessous, utilise une autre variante courante de la méthode multiplicateur/référence, fondée sur les statistiques officielles des décès parmi les héroïnomanes. Là encore, le chiffre de référence provient d'une source existante et les auteurs font des concessions analogues pour établir la valeur du multiplicateur.

Étude de cas n° 2. Décès par surdose en Nouvelle-Galles du Sud (Estimation multiplicateur/référence au moyen d'un indicateur d'usage de drogues)

Voici une deuxième illustration simple de la méthode multiplicateur/référence, tirée d'une étude utilisant un "multiplicateur de décès". Cette étude visait à confirmer les estimations du nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud (Australie) résultant d'autres méthodes et d'autres études, en appliquant un coefficient multiplicateur aux données nationales relatives aux surdoses d'héroïne et plus précisément aux décès liés à ces surdoses. Il n'est rendu compte ici que de la méthode du multiplicateur de décès, mais celle-ci fait partie d'une étude bien plus vaste (voir les autres études de cas ci-après).

Obtention d'un multiplicateur

Le plus difficile dans une estimation par la méthode du multiplicateur consiste généralement à déterminer le multiplicateur. Dans le cas présent, aucune étude directe ou distincte n'a été menée à cet effet, mais un multiplicateur de décès par surdose d'héroïne a été établi à partir des publications existant sur le sujet. En l'occurrence, le regroupement des résultats d'un certain nombre d'études de cohortes d'usagers réguliers d'héroïne [par exemple, Frischer (1998) et Reuter (1993)] fait apparaître un taux de décès compris entre 0,8 % et 1 % par an parmi ces usagers, ce qui donne un "multiplicateur de décès" compris entre 125 (0,8 %) et 100 (1 %).

Utilisation du chiffre de référence

Le nombre moyen de décès par surdose d'héroïne enregistré en Nouvelle-Galles du Sud a été d'environ 360 par an pendant la période considérée. En appliquant les multiplicateurs probables, il faudrait donc qu'il y ait entre 36 000 (avec un multiplicateur de 100) et 45 000 (avec un multiplicateur de 125) usagers réguliers d'héroïne en Nouvelle-Galles du Sud pour expliquer les taux de mortalité observés (voir le tableau 3). Cette estimation approximative est semblable aux estimations précédentes réalisées par d'autres moyens.

Tableau 3. Utilisation du nombre de décès par surdose pour estimer le nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud sur la base d'un taux de mortalité annuel de 1 %
(Moyennes calculées sur une période de cinq ans)

Élément	Valeurs utilisées	Estimations
Référence (R)	Nombre de décès par surdose par an ^a	360
	Pourcentage d'usagers réguliers d'héroïne qui meurent chaque année d'une surdose ^b	1 %
Multiplicateur (M)	Multiplicateur calculé: 1,0 / 0,01 (soit un sur cent)	100
Estimation de la population	Référence par multiplicateur (R * M)	36 000

^aDéterminé à partir des registres de décès disponibles.

^bDéterminé à partir des taux de mortalité cités dans des travaux de recherche.

Mise en garde

Bien sûr, la méthode ci-dessus comporte aussi des limites, notamment parce que l'essentiel des études de cohortes, qui ont servi à estimer les taux annuels de mortalité par surdose, n'ont pas été réalisées en Australie et portent sur une période de temps qui ne correspond que grossièrement à celle couverte par les données de référence. Elle ne peut donc pas tenir compte des variations temporelles et géographiques éventuelles des taux de décès par surdose. Quoique

utile pour fournir une première approximation dans les régions où l'on peut facilement obtenir des données exactes sur la mortalité, compte tenu de la nature assez rudimentaire du multiplicateur retenu, cette méthode ne donne, au mieux, qu'une estimation imprécise. On pourrait peut-être en améliorer la précision en réalisant des études distinctes par région et par période de temps des taux de mortalité par surdose parmi les usagers réguliers d'héroïne, même si, faute d'informations plus détaillées, le multiplicateur employé devrait être le même dans tous les cas.

Cette simple description de la méthode ne rend pas compte des diverses concessions qui doivent être faites dans la pratique; il faut bien sûr définir précisément ce que l'on entend par "usager de drogues", remplacer le "nombre d'usagers de drogues qui meurent dans l'année" par le "nombre de décès enregistrés comme étant liés à la drogue" et supposer que le "taux de mortalité" des études de cohortes publiées sur la mortalité constitue une estimation raisonnable du "rapport entre le nombre de décès liés à la drogue et le nombre d'usagers de drogues pour l'année en question", même s'il a été calculé pour une autre période de temps et dans un pays différent.

Les difficultés propres à l'emploi d'un multiplicateur de décès résident dans le fait que, pour estimer le taux de mortalité parmi les usagers de drogues, l'idéal serait d'effectuer une étude locale longitudinale spécialisée et que, à moins qu'il ne s'agisse d'une étude de très grande ampleur, il faudrait beaucoup de temps pour parvenir à un résultat. Un autre problème tient au fait que le multiplicateur lui-même est particulièrement élevé: si 1 % seulement de la population est visible, le manque de fiabilité de l'estimation est d'autant plus grand.

Toutefois, le fait que cette estimation concorde avec d'autres effectuées antérieurement permet de s'y fier davantage et montre l'intérêt de combiner des estimations obtenues au moyen de diverses méthodes, comportant chacune de nombreuses limites pour accroître la confiance dans le résultat final.

Source: Frischer, "Estimating the prevalence of drug abuse..." (1997); Hall *et al.* (2000); et Reuter (1993).

La précédente étude de cas montre que la population totale d'usagers réguliers d'héroïne est estimée en multipliant le chiffre connu ("officiellement visible") d'usagers décédés dans l'année par le multiplicateur de décès (la réciproque du taux de mortalité annuel parmi les usagers). Dans les pays où il n'est pas facile d'obtenir des statistiques sur les décès liés à la drogue, on utilise aussi souvent comme référence le nombre d'usagers de drogues en traitement. Pour utiliser le sous-groupe des usagers "en traitement" comme référence, il faut:

- a) Connaître le nombre total d'usagers qui suivaient un traitement à un moment donné au cours de l'année considérée;
- b) Disposer d'une estimation tirée d'une enquête par sondage de la proportion d'usagers de drogues qui ont suivi un traitement au cours de cette année-là (par exemple, un sur cinq est un multiplicateur souvent mentionné dans les travaux de recherche publiés au Royaume-Uni).

Étude de cas n° 3. Calcul du multiplicateur à partir de données relatives aux traitements et aux arrestations en Nouvelle-Galles du Sud (*Estimations multiples utilisant différentes sources de données*)

Contexte

Au cours des années 90, il semble que l'offre et l'usage d'héroïne aient augmenté en Australie (et dans de nombreux autres pays). L'usage d'héroïne et les questions qu'il soulève se sont retrouvés au centre des préoccupations du public et des responsables politiques, et un débat intense s'est engagé dans les médias sur l'ampleur du problème et les stratégies envisageables pour y remédier. Cependant, cela s'est produit sans que l'on dispose d'aucune donnée sur le nombre de personnes faisant usage d'héroïne ou dépendantes de l'héroïne. En effet, l'opinion largement répandue selon laquelle le nombre d'usagers d'héroïne s'était considérablement accru en Australie ne s'appuyait sur aucune donnée concrète, mais reposait en réalité essentiellement sur les quatre types d'information ci-après :

- a) Des données venant de l'étranger qui montraient que la production mondiale d'opium avait augmenté;
- b) Des informations locales provenant de la police et d'entretiens avec des usagers d'héroïne et d'autres informateurs privilégiés, qui montraient que le prix de détail de l'héroïne avait baissé alors que le degré de pureté de la drogue et les quantités disponibles sur le marché avaient augmenté;
- c) Une augmentation progressive mais constante du nombre d'héroïnomanes demandant à bénéficier d'un traitement;
- d) Enfin, la montée en flèche largement attestée du nombre de décès par surdose d'héroïne dans les années 90 constituait peut-être l'indication la plus convaincante d'une augmentation de l'usage d'héroïne.

C'est dans ce contexte que les autorités ont demandé à un groupe de chercheurs de tenter d'estimer le nombre de personnes faisant régulièrement usage d'héroïne ou dépendantes de cette substance en Australie.

Sources de données

Pour un certain nombre de raisons, cette tâche est peut-être relativement plus aisée en Australie qu'elle ne le serait dans de nombreux pays. En particulier, le Bureau australien de statistique recueille des données complètes sur la mortalité, y compris sur les décès liés à la drogue, codées selon la dixième révision de la Classification statistique internationale des maladies et des problèmes de santé connexes (CIM-10), et la police comme les centres de traitement tiennent des archives complètes. Cependant, malgré ces atouts, comme indiqué plus haut, il reste difficile d'estimer la taille de la "population cachée".

L'objectif était d'établir une estimation du nombre d'usagers réguliers d'héroïne dans l'ensemble de l'Australie, mais il était beaucoup plus facile d'estimer d'abord le nombre de ces personnes en Nouvelle-Galles du Sud, le plus peuplé des huit États et territoires australiens. Les données disponibles concernant cet État provenaient principalement de deux sources :

- a) *Données relatives aux arrestations.* Les services de police de Nouvelle-Galles du Sud ont fourni des données concernant les arrestations pour infractions liées à l'héroïne (possession et fourniture d'héroïne) pour la période allant de 1997 à 1999;
- b) *Données relatives aux programmes d'entretien à la méthadone.* Le principal mode de traitement de la dépendance aux opioïdes en Australie est

le traitement d'entretien à la méthadone. La méthadone est prescrite par des médecins qui travaillent dans des centres de traitement spécialisés ou qui dispensent des soins de santé primaires. Les personnes auxquelles est prescrite de la méthadone sont d'abord immatriculées auprès des services pharmaceutiques du Ministère de la santé de la Nouvelle-Galles du Sud, de façon à éviter que des personnes ne reçoivent de la méthadone de plusieurs sources.

Calcul du multiplicateur

Il a été possible d'utiliser les données relatives au traitement d'entretien à la méthadone comme référence pour estimer le nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud. D'après ces données, le nombre total d'usagers réguliers d'héroïne ayant commencé un traitement d'entretien à la méthadone pendant la période concernée s'élevait à 13 000.

En utilisant ce chiffre comme référence, il ne restait plus qu'à trouver le multiplicateur approprié pour estimer le nombre total d'usagers. D'après des études antérieures, environ un tiers des usagers d'héroïne interrogés avaient commencé un traitement d'entretien à la méthadone au cours de l'année précédente. On pouvait donc estimer à 39 000 (13 000 x 3) le nombre total d'usagers réguliers (voir le tableau 4).

Tableau 4. Estimation du nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud à partir du nombre de patients suivant un traitement d'entretien à la méthadone

Élément	Valeurs utilisées	Estimations
Référence (R)	Nombre de patients suivant un traitement d'entretien à la méthadone en dispensaire au cours de l'année ^a	13 000
	Pourcentage d'usagers réguliers d'héroïne ayant suivi un traitement d'entretien à la méthadone au cours de l'année écoulée ^b	33,33 %
Multiplicateur (M)	Multiplicateur calculé: 1,0 / 0,3333 (soit un sur trois)	3
Estimation de la population	Référence par multiplicateur (R * M)	39 000

^aDéterminé à partir des registres médicaux disponibles.

^bDéterminé à partir d'études publiées sur des échantillons d'usagers d'héroïne.

Robustesse du résultat

Ces données ont été analysées au moyen de diverses techniques, notamment par saisie avec recoupement (capture-recapture) ainsi que par des méthodes multiplicateur/référence. Des données portant sur de nombreuses années et des méthodes de rétroprojection initialement mises au point pour évaluer le VIH/sida ont été utilisées pour déterminer l'évolution dans le temps du nombre de personnes faisant régulièrement usage d'héroïne ou dépendantes de cette substance.

Il est surprenant de constater que les estimations obtenues au moyen de ces différentes méthodes se situent dans une fourchette étroite: les six estimations du nombre d'usagers réguliers d'héroïne en Nouvelle-Galles du Sud ont été comprises entre 32 000 et 45 000, 37 000 étant la valeur médiane. Étant donné l'étroitesse de cette fourchette, il a été simplement décidé de considérer que cette valeur médiane constituait la "meilleure estimation", mais il serait aussi possible d'employer une méthode plus élaborée qui permettrait de combiner les différentes estimations en les pondérant en fonction de la confiance relative qu'on leur accorde.

Conséquences de l'imperfection des procédures

Cette étude a mis en évidence l'une des difficultés que pose la réalisation de calculs analogues à partir des arrestations enregistrées par la police. En s'appuyant sur des études antérieures, on avait estimé qu'environ 20 % des usagers réguliers d'héroïne interrogés avaient été arrêtés au cours de l'année écoulée et qu'il fallait donc employer un multiplicateur de 5,0 pour obtenir le nombre total d'usagers réguliers. D'après les registres de la police, quelque 2 400 personnes avaient été arrêtées au cours de la période considérée pour des infractions liées à l'héroïne. L'application de ce multiplicateur au nombre total d'arrestations donne 12 000 usagers réguliers d'héroïne en Nouvelle-Galles du Sud, chiffre qui est bien inférieur à d'autres estimations, y compris à celle établie à partir du nombre de personnes suivant un traitement d'entretien à la méthadone.

Plusieurs raisons peuvent expliquer pourquoi les données de la police relatives aux arrestations ne conviennent peut-être pas pour la méthode multiplicateur/référence. Il est probable que certaines des arrestations déclarées par les usagers réguliers qui avaient été interrogés aient été comptabilisées comme des arrestations pour des infractions non liées à l'héroïne. Il se peut aussi que les personnes interrogées aient constitué un groupe d'usagers davantage susceptibles de se faire arrêter que l'ensemble des usagers réguliers, même si aucune raison particulière ne permettait de le penser. Cela signifierait, dans le premier cas, que le multiplicateur établi d'après les entretiens ne correspondait pas exactement à la référence et, dans le second cas, que le multiplicateur ne représentait pas réellement la situation générale. Dans les deux cas, la proportion réelle d'usagers réguliers, dont les arrestations avaient été enregistrées comme étant liées à l'héroïne, était probablement inférieure à celle établie d'après les études. En d'autres termes, le multiplicateur retenu aurait été trop petit et, en l'appliquant au chiffre de référence, on aurait donc sous-estimé le nombre total d'usagers réguliers.

Source: Hall *et al.* (2000); et McKetin *et al.* (1999).

Collecte primaire de nouvelles données

Il est recommandé, dans la mesure du possible, que les chercheurs réalisent une enquête par sondage sur la population cible – usagers injecteurs ou à problème –, dans le cadre de l'étude d'estimation de la prévalence. Cela présente plusieurs avantages:

- a) Ces enquêtes sont susceptibles de fournir plusieurs multiplicateurs et, si des données permettant d'identifier les personnes interrogées sont recueillies, elles peuvent constituer une source de données complémentaires pour les méthodes d'estimation de la prévalence par capture-recapture (voir ci-après la section intitulée "Méthodes de saisie avec recoupement (capture-recapture)");
- b) Elles permettent de recueillir des données sur les principaux risques et comportements de protection, ainsi que d'estimer la prévalence des virus transmis par le sang et de mesurer la portée des mesures de réduction des risques;
- c) Enfin, la conduite d'entretiens avec la population cible permet de poser des questions et de recueillir des informations susceptibles d'être utilisées dans une étude multiplicateur/référence (voir ci-dessus la section intitulée "Les méthodes multiplicateur/référence").

La méthode employée dans l'exemple précédent permet, en particulier, de réaliser une étude spéciale, afin de déterminer la valeur du multiplicateur. Dans l'exemple suivant (étude de cas n° 4), de nouvelles données sont recueillies, afin d'estimer le multiplicateur selon des principes analogues et de déterminer le chiffre de référence.

Étude de cas n° 4. Évaluation nationale au Pakistan en 2000
(Emploi de la méthode multiplicateur-référence avec un multiplicateur de traitement établi d'après des entretiens avec des informateurs privilégiés)

Situation de départ et évaluation

Cette étude visait à déterminer la prévalence de la dépendance aux drogues dures au Pakistan. On disposait de bonnes données concernant la taille et la structure de la population de toutes les provinces pakistanaïses, mais de peu d'informations fiables sur l'usage de drogues lui-même. Une estimation en avait déjà été faite selon des méthodes d'enquête nationales, mais elle était beaucoup trop ancienne.

On disposait pourtant de quelques informations sur les centres de désintoxication, en l'occurrence des centres spécialisés destinés aux usagers de drogues (essentiellement aux personnes faisant abus d'alcool ou d'héroïne). Ces personnes étaient également traitées dans des hôpitaux publics. Il a donc été décidé d'essayer d'utiliser les données du système de soins sur le nombre de personnes en traitement pour estimer, avec un multiplicateur, le nombre total de personnes concernées dans plusieurs agglomérations pakistanaïses.

L'établissement par extrapolation, à partir des chiffres de ces études locales, d'estimations pour chacune des grandes provinces du Pakistan est traité ci-après (voir la section intitulée "Estimations nationales de la prévalence établies par extrapolation à partir d'estimations locales"). La méthode du coefficient multiplicateur appliqué aux traitements décrite ici constituait le premier volet de l'estimation de la prévalence, le second consistant à utiliser le nombre de toxicomanes incarcérés et un multiplicateur correspondant. On espérait que ces estimations fondées sur des données différentes présenteraient une certaine concordance qui en renforcerait la validité.

Définition de la population cible

Selon la définition adoptée, la population cible était celle des usagers d'héroïne ou des injecteurs (de toute substance); à défaut de meilleure appellation, ces personnes ont été qualifiées de "toxicomanes". La consommation de feuilles ou de résine de cannabis (haschisch ou charros) est la forme la plus courante d'abus de drogues au Pakistan – probablement plus encore que celle d'alcool, dont l'usage est illégal –, mais les chercheurs ont estimé qu'ils ne pourraient pas, dans la pratique, conduire des entretiens ou des enquêtes ou réaliser des sondages efficaces pour estimer simultanément la prévalence du cannabis et de l'héroïne ou de l'injection.

Pour les raisons pratiques indiquées ci-dessous, il a été décidé de réduire, aux fins de l'estimation de la prévalence, la population cible aux hommes âgés de 15 à 45 ans:

- a) D'une façon générale, les mœurs sociales excluaient que des femmes faisant usage de drogues puissent être vues dans les centres de traitement,

bien que l'on considère que la consommation d'opiacés en vente libre constituait un problème grave parmi les femmes;

- b) Parce que des travaux antérieurs indiquaient que cette tranche d'âge comprenait près de 90 % des personnes ayant une dépendance aux drogues dures, et qu'on ne voyait pratiquement jamais aucun homme plus jeune ou plus âgé en traitement.

Le chiffre de référence et le multiplicateur ont été déterminés en ne prenant en compte que les centres de traitement spécialisés pour toxicomanes. Les personnes suivant un traitement dans les hôpitaux publics ont été exclues pour les raisons suivantes:

- a) Elles étaient peu nombreuses;
- b) Il n'était pas certain qu'une distinction puisse être établie entre les toxicomanes et les autres patients traités;
- c) Les données consignées concernant les antécédents et l'abus de drogues étaient inadéquates.

Détermination du chiffre de référence

Il a fallu recueillir toutes les informations de référence sur le nombre de personnes en traitement – à tous les niveaux – en recensant à part (avec une couverture de 100 %) tous les centres de traitement spécialisés pour toxicomanes. Pour ce faire, on a mis à jour d'anciennes listes de centres spécialisés afin de constituer un registre national des centres de traitement, en espérant que cela favoriserait la réalisation de nouveaux travaux sur la prévalence de l'usage de drogues. Au total, 73 centres ont été recensés au niveau national.

La collecte des données dans les centres a été effectuée selon des procédures permettant d'envisager différentes définitions de la référence, mais la définition retenue pour l'analyse a été le "nombre de toxicomanes traités en tant que patients stationnaires au cours de l'année écoulée" (pour les autres définitions envisagées, voir l'appendice 1 du rapport principal du Pakistan). Les directeurs de chacun des 73 centres spécialisés ont été interrogés par téléphone ou en personne, pour déterminer le nombre de toxicomanes traités au cours de l'année écoulée. Ces entretiens ont été conduits par du personnel médical.

Détermination du multiplicateur

Là encore, la tâche la plus difficile a été de déterminer le multiplicateur. Celui-ci a été estimé au moyen d'une enquête spécialement conçue pour recueillir les informations voulues. Les calculs ont été fondés sur des informations fournies par un échantillon d'informateurs privilégiés lors d'entretiens individuels. Pour les critères de sélection des informateurs privilégiés, voir ci-après la section intitulée "Travail sur le terrain". Il a été demandé à chaque informateur:

- a) Combien de toxicomanes il avait personnellement rencontrés au cours des 12 derniers mois;
- b) Combien d'entre eux avaient été, à sa connaissance, traités en tant que patients stationnaires dans un centre spécialisé au cours des 12 mois précédents.

Il leur a également été demandé combien de toxicomanes avaient séjourné en prison au cours des 12 mois précédents, mais cette information a uniquement été utilisée pour déterminer le multiplicateur secondaire permettant d'établir

une estimation à partir du nombre de personnes emprisonnées. Comme l'entretien avec chaque informateur privilégié était long, ces questions essentielles ont été posées en premier.

Pour le calcul du multiplicateur à appliquer au nombre de personnes traitées, on a exclu les informateurs privilégiés ayant participé d'une façon ou d'une autre aux services de traitement et, pour le calcul du multiplicateur à appliquer au nombre de personnes incarcérées, on a exclu ceux qui avaient un rapport avec les services de police ou les établissements pénitentiaires. Les informateurs privilégiés qui n'avaient pas rencontré un nombre minimal de toxicomanes ont également été exclus des analyses.

Échantillonnage

Les informateurs privilégiés ont été sélectionnés sur une base nationale. Trente-six localités géographiques ont été expressément choisies – et non déterminées aléatoirement – sur l'ensemble du territoire pakistanais, afin de représenter la structure sociale générale du pays. Elles ont servi de premières unités d'échantillonnage (grappes) et ont été stratifiées pour constituer dix-huit paires, comprenant chacune une localité globalement urbaine et une localité globalement rurale géographiquement proches. Les paires ainsi formées, plus nombreuses dans les régions les plus peuplées, couvraient les quatre provinces du Pakistan.

Les enquêteurs (40 spécialistes) ont reçu pour instruction d'interroger, dans chaque localité, au moins cinq informateurs privilégiés qu'ils devaient sélectionner eux-mêmes à partir d'une liste de catégories socioprofessionnelles, en choisissant si possible des informateurs dans au moins cinq catégories différentes dans chaque localité. Les catégories répertoriées étaient les suivantes: policiers, juges, médecins, agents de santé, autres agents publics, maires, conseillers, mollahs, prêtres, travailleurs sociaux, enseignants, chefs de tribu et anciens toxicomanes. Les refus de collaborer n'ont pas été consignés, puisque aucune approche formelle ni aucune structure d'échantillonnage n'avait pu être définie pour une même localité.

Travail sur le terrain

L'organisation des travaux sur le terrain étant complexe, la responsabilité d'en diriger l'exécution courante a été confiée à un chercheur local spécialisé afin de limiter les retards et les difficultés. Les entretiens ont été conduits par des personnes formées spécialement pour l'enquête: personnel médical, médecins, travailleurs sociaux, stagiaires dans ces professions et ex-toxicomanes associés à l'opération, plus un ou deux volontaires remplissant les conditions voulues. Pour assurer la comparabilité des données recueillies par un si grand nombre d'enquêteurs et faire en sorte que ces données correspondent exactement aux besoins, les entretiens ont tous été conduits sur la base d'un questionnaire.

Pour contrôler la qualité des informations provenant d'une étude aussi large, quatre responsables régionaux désignés ont contrôlé de près la circulation des questionnaires et veillé à ce qu'il soit rendu compte de l'emploi de chacun d'entre eux. En outre, chaque questionnaire a été signé par l'enquêteur et contre-signé après vérification par un responsable régional.

Même en limitant la collecte des données à 18 localités du pays, les déplacements ont posé un problème (l'un des enquêteurs a ainsi dû voyager deux jours à dos de chameau).

Analyse et résultats

Les calculs prenant comme référence le nombre de patients stationnaires de sexe masculin âgés de 15 à 45 ans ont été effectués au niveau géographique le plus bas, c'est-à-dire pour chaque localité distincte où il y avait un ou plusieurs centres. En ce qui concerne l'extrapolation et l'agrégation des données aux fins de la détermination de la prévalence nationale au Pakistan, le lecteur est invité à se reporter à la section intitulée "Estimations nationales de la prévalence établies par extrapolation à partir d'estimations locales".

Le calcul de la référence primaire a été effectué de la façon suivante.

Comme il y avait des centres de traitement spécialisés dans environ 30 % des localités étudiées et aucun ailleurs, des références ne pouvaient être déterminées que pour ces "localités de traitement". Des totaux de référence ont été calculés pour chacune de ces localités en dénombrant les héroïnomanes et les usagers de drogues par injection de sexe masculin, âgés de 15 à 45 ans, traités en tant que patients stationnaires. Ces chiffres ont été ajustés pour tenir compte du fait qu'environ 10 % de ces patients venaient de régions éloignées. Les chiffres de référence ajustés ont été regroupés pour obtenir un chiffre global pour la principale ville et toutes les autres localités de chaque province.

Des multiplicateurs ont été calculés pour chaque localité en divisant le nombre de toxicomanes rencontrés par un informateur privilégié au cours de l'année écoulée par le nombre de ces toxicomanes ayant reçu un traitement. Parmi les différentes façons possibles de déterminer un multiplicateur unique, on a opté pour la médiane, parce que ce chiffre n'était pas trop influencé par les réponses extrêmes et ne variait pas selon qu'il était calculé à partir du pourcentage de personnes traitées ou des multiplicateurs eux-mêmes. Les multiplicateurs ont été regroupés de la même façon que les chiffres de référence, c'est-à-dire pour la principale ville de la province et pour toutes les autres localités de la province.

Le chiffre de référence global a été multiplié par le multiplicateur médian pour obtenir le nombre total de toxicomanes dans la principale ville et à l'extérieur de celle-ci dans chacune des quatre provinces. Deux autres estimations destinées à donner une idée de la variabilité de cette procédure ont consisté à prendre les quartiles inférieur et supérieur du multiplicateur au lieu de la valeur médiane. En répétant l'ensemble du calcul avec ces multiplicateurs et les mêmes chiffres de référence, on a défini la fourchette dans laquelle se situe le nombre de toxicomanes dans chaque province.

Les résultats de ces estimations ne sont pas présentés ici. Pour plus d'informations, on se reportera au rapport principal du Pakistan.

Mise en garde

Les estimations de la prévalence ont été obtenues en divisant le nombre estimatif total de toxicomanes de chaque localité par le nombre d'hommes âgés de 15 à 45 ans dans cette localité. Ces estimations ont été jugées plausibles, mais elles sont très basses par rapport à l'enquête précédente. Cependant, celle-ci, comme les enquêtes ultérieures destinées à l'actualiser, reposait sur des postulats beaucoup plus vastes que la présente étude. En conduisant ensuite des enquêtes de moindre ampleur à différents intervalles pour déterminer les facteurs d'actualisation de l'enquête initiale, on ne peut que gonfler l'erreur d'estimation, quelle que soit la fiabilité de l'enquête initiale.

Comme pour les autres études, le plus difficile a été d'estimer le multiplicateur. Cela pouvait être fait de plusieurs façons, et la procédure retenue était susceptible d'exercer une influence (à la hausse ou à la baisse) sur le résultat. Par exemple, compte tenu des différentes réponses des informateurs privilégiés dans chaque localité, il aurait pu être déterminé en calculant la moyenne des différents multiplicateurs, en calculant la moyenne des différentes estimations du pourcentage de personnes en traitement, en calculant ces moyennes au niveau des localités ou au niveau de la province et en additionnant les nombres de toxicomanes rencontrés et traités avant d'établir la moyenne. En fait, en calculant la prévalence selon cette dernière procédure, on a obtenu des résultats très semblables à ceux obtenus avec la méthode retenue.

En outre, il est certain que le fait d'exclure les informateurs privilégiés qui n'avaient pas rencontré un nombre minimal de toxicomanes, ainsi que ceux qui avaient participé aux traitements ou qui étaient en rapport avec les établissements pénitentiaires et les services de police, a eu une incidence sur les résultats. Bien sûr, les informations réunies auprès des informateurs privilégiés risquent de comporter des erreurs. La distinction entre hôpital public et centre de traitement spécialisé peut ne pas avoir été claire dans l'esprit des informateurs privilégiés. En outre, ils n'ont pas nécessairement exclu les patients traités en dehors de l'année considérée, et il a été impossible de demander ou de vérifier à quel endroit les toxicomanes connus avaient été traités.

Enfin, même si les données des recensements de population peuvent être exactes, la définition des zones desservies par les centres de traitement peut compliquer leur utilisation pour le calcul des taux de prévalence.

Dans la mesure où il est extrêmement difficile de réaliser des estimations de la prévalence en l'absence de données, il est absolument indispensable d'employer différentes méthodes et procédures.

Source: Rehmann, Griffiths et Taylor (2002).

Autres méthodes d'estimation de la valeur du multiplicateur

Une fois que le groupe de référence a été défini, il est généralement assez facile de déterminer la taille de ce groupe à partir de la source pertinente. Ce qui est toujours le plus difficile, au moins dans le domaine de l'épidémiologie de l'usage des drogues, c'est l'estimation du multiplicateur. Outre les études de cohortes permettant d'estimer les taux de mortalité, et les études spécialisées servant à déterminer les taux de traitement, d'immatriculation ou d'arrestation, il existe des techniques d'échantillonnage en chaîne parfois appelées "sondages en boule de neige", qui consistent à demander, lors d'entretiens individuels, à chaque personne d'un échantillon principal, de désigner un certain nombre de personnes qu'elle connaît et qui sont ensuite intégrées à l'échantillon dans un deuxième temps.

Il est ainsi possible de réunir des informations complémentaires concernant l'usage de drogues par les personnes désignées, soit auprès de membres de l'échantillon

principal, soit auprès des personnes désignées elles-mêmes. Pour estimer la proportion d'usagers de drogues en traitement, on peut, par exemple, poser la question suivante à chaque membre de l'échantillon principal: "Parmi vos 10 amis les plus proches qui font usage de drogues, combien ont suivi un traitement dans la région au cours des 12 derniers mois?", puis regrouper les informations ainsi recueillies.

Hétérogénéité et stratification de la population

L'étude de cas ci-dessus n'a porté que sur les hommes âgés de 15 à 45 ans, car on a estimé que peu de données seraient disponibles en dehors de cette tranche d'âge, en partie parce que l'on pensait y trouver très peu de "toxicomanes". Le fait de se limiter à cette tranche d'âge a permis de mieux se concentrer sur l'essentiel.

Bien que la toxicomanie chez les femmes soit une question dont on se préoccupe beaucoup, on a pensé que l'estimation aurait été faussée si celles-ci avaient été incluses avec les hommes dans l'étude. Si elles avaient été incluses, il aurait certainement fallu appliquer un autre multiplicateur au nombre de personnes traitées, car les femmes en traitement étaient très peu nombreuses. Cette subdivision du problème – d'abord les hommes, puis les femmes – est un exemple de stratification de la population. Il s'agit de décomposer une population hétérogène en sous-groupes plus homogènes, de façon à améliorer la précision des procédures d'estimation de la prévalence. Selon la population cible, d'autres subdivisions sont envisageables: injecteurs/non-injecteurs, personnes ayant un emploi/personnes sans emploi, etc. Toute caractéristique utile pour déterminer des multiplicateurs et des chiffres de référence distincts devrait être prise en considération. On ne perd pas grand-chose si la stratification se révèle inutile – seul le recours à des échantillons extrêmement petits conduit à une réduction sensible de la précision –, mais on a beaucoup à gagner si elle améliore l'homogénéité des sous-groupes cibles.

Cette étude de cas illustre l'emploi d'un autre grand type de stratification: la stratification géographique. Au lieu d'établir pour l'ensemble de l'étude une seule estimation du multiplicateur et un seul chiffre de référence, des chiffres distincts ont été calculés pour chacune des quatre grandes régions géographiques et, à l'intérieur de celles-ci, la principale ville a été traitée à part. On a procédé de la sorte, car on pensait que la valeur du multiplicateur applicable à chacune de ces huit divisions pourrait être différente et qu'un groupage risquait d'introduire des inexactitudes. La stratification géographique a en outre pour avantage d'avoir permis la réalisation d'estimations distinctes de la prévalence pour chacune des quatre régions.

Conditions requises et risques d'erreurs

Cette procédure a pour avantage de pouvoir s'appliquer à bon nombre de situations, à condition de disposer de deux éléments:

- a) Le chiffre de référence requis – par exemple, le nombre de décès parmi les usagers de drogues;

- b) Le multiplicateur requis – par exemple, le taux de mortalité dans l'ensemble de la population d'utilisateurs de drogues.

Le multiplicateur peut être déterminé au moyen de n'importe quelle technique d'échantillonnage applicable, notamment par échantillonnage aléatoire, comme dans l'étude de cohorte sur la mortalité citée plus haut, ou selon diverses techniques d'échantillonnage par désignation (sondages en boule de neige). Cette souplesse fait tout l'intérêt de cette méthode.

La définition du sous-groupe de référence, qui vise avant tout à faciliter la tâche des chercheurs, n'est qu'une première étape du travail d'estimation: les exemples cités plus haut ont montré que l'on pouvait utiliser aussi bien le nombre de décès parmi les usagers de drogues que le nombre de personnes en traitement, le nombre des arrestations enregistrées par la police ou le nombre de patients des centres de traitement de l'infection à VIH. Toute définition claire et précise convient. Mais, la même définition doit être employée pour la détermination du chiffre de référence et la collecte des données de construction de l'échantillon, qui servira à déterminer le multiplicateur. Du point de vue de la robustesse de l'estimation de la prévalence, il s'agit là d'un atout majeur.

Dans la pratique, il est nécessaire de définir exactement la référence et le multiplicateur correspondant. Par exemple, dans une étude fondée sur le nombre de personnes traitées, il faut savoir quelle sorte de traitement est appliqué; il se peut que "traitement à la méthadone" constitue une définition suffisante, mais il peut être préférable de disposer d'une liste des centres de traitement. Il convient de souligner qu'il faut définir également l'étendue géographique de la localité ou de la région considérée, ainsi que la période précise à laquelle les données s'appliquent.

Il faut aussi tenir compte des suppositions qui sont faites. Tout d'abord, on suppose que les données "de référence" sont correctes. Or on sait que, malheureusement, les sources de données courantes peuvent être inexactes en raison des sous-déclarations ou du fait que les données recueillies sont incomplètes. Ainsi, dans l'étude de cas n° 1, les auteurs ont évoqué la possibilité que les laboratoires sous-estiment le nombre de tests de dépistage du VIH effectués et que les médecins qui prescrivent les tests n'indiquent pas toujours que les personnes concernées font usage de drogues par injection. En conséquence, il peut être nécessaire d'ajuster le total de référence pour tenir compte de ce type de sous-déclaration.

La méthode suppose que l'on ait correctement défini le multiplicateur. Lorsqu'on calcule le multiplicateur, la question essentielle est de savoir si les personnes visées figurent dans la référence. Si tel est le cas, le multiplicateur correspond effectivement à la référence (même si celle-ci ne rend pas entièrement compte de la population cible). Quand le nombre de référence est celui des personnes traitées, par exemple, il peut être nécessaire d'établir la liste des centres de traitement pris en considération, de façon à assurer une stricte équivalence des définitions de la référence,

et du multiplicateur. Ceci vaut tout particulièrement pour les études fondées sur une stratification géographique, où les usagers de drogues d'une région peuvent être traités dans une autre.

La méthode suppose également que l'on dispose d'une estimation non biaisée du multiplicateur. Dans l'idéal, cette estimation est déterminée à partir d'un échantillon représentatif d'usagers de drogues à problème, sur une période et dans un lieu précis correspondant aux données de référence que l'on compte utiliser. C'est rarement le cas. Dans l'étude de cas n° 1, les auteurs ont utilisé un multiplicateur tiré d'une enquête sur des usagers de drogues par injection menée dans une autre ville et ont supposé que le multiplicateur serait le même pour Toronto en 1996. Comme aucun échantillon aléatoire d'usagers de drogues à problème n'est véritablement représentatif, le mieux est de recruter des sujets d'une façon qui limite tout risque de biais. Ainsi, si l'on voulait effectuer une estimation non biaisée de la proportion d'usagers de drogues par injection inscrits à un programme d'échange de seringues, il serait malavisé de recruter directement ces usagers en dehors d'un tel programme d'échange et de demander combien sont inscrits.

Si les informations servant au calcul du multiplicateur sont recueillies au cours d'entretiens, on doit supposer que l'événement de référence est suffisamment fréquent et marquant pour être détecté dans un échantillon d'usagers de drogues à problème ou pour que ceux-ci s'en souviennent. Par exemple, il est presque certain qu'ils rendront compte fidèlement du fait qu'ils suivent un traitement ou qu'ils ont été arrêtés pour possession de drogues. En outre, la question qui leur est posée doit être claire, de sorte que, si, par exemple, on prend comme référence les usagers de drogues par injection inscrits dans un programme spécial d'échange de seringues, les personnes interrogées répondent "non" quand elles ne sont allées que dans des pharmacies pour se procurer des seringues propres et "oui" quand elles se sont rendues dans un centre d'échange de seringues ou dans des pharmacies et un tel centre.

L'une des principales conditions requises est, bien sûr, que le multiplicateur représente réellement le rapport entre le chiffre de référence et l'ensemble de la population cible. Par exemple, si les taux de traitement sont très différents dans les zones urbaines et les zones rurales, il pourrait être erroné de regrouper ces zones pour obtenir un multiplicateur unique. Pour éviter ce problème, on peut recourir à une stratification de la population, comme indiqué ci-dessus, si des données sont disponibles pour chaque strate. Pour savoir si c'est nécessaire, il faut d'abord examiner les données et se prononcer sur la base des certitudes déjà acquises.

Il est bien sûr possible qu'une ou plusieurs, voire la totalité, de ces conditions ne soient pas réunies; dans ce cas, il est tout à fait possible que l'étude aboutisse à une estimation incorrecte de la prévalence. Il serait malavisé de ne se fonder que sur une seule estimation par multiplication (voir ci-après la section intitulée "Exactitude des estimations de la prévalence").

Méthodes de saisie avec recouplement (capture-recapture)

L'estimation par saisie avec recouplement (capture-recapture) consiste à exploiter des sources de données courantes qui identifient les usagers de drogues à problème, afin d'estimer la prévalence de l'usage de drogues. En épidémiologie ou dans les recherches effectuées sur l'homme, les sources de données sont souvent des listes de personnes; dans le cas présent, ce sont des listes répertoriant des usagers de drogues à problème. La méthode de capture-recapture est une méthode d'estimation indirecte de la prévalence qui utilise les recouplements entre des listes ou des sources de données incomplètes. Ces listes sont incomplètes, car aucune source ne répertorie tous les usagers de drogues à problème (si c'était possible, il n'y aurait aucune difficulté à estimer la prévalence) et parce qu'il est rare d'obtenir un taux de déclaration de 100 %. Le recouplement met en évidence les personnes qui figurent sur plusieurs listes à la fois.

L'objectif est de déterminer combien d'usagers de drogues à problème devraient figurer sur la liste si celle-ci répertoriait l'ensemble de la population cible, par exemple tous les usagers de drogues à problème dans une ville et pendant une année donnée. Ainsi, pour Hook et Regal (1995), l'estimation par capture-recapture est une technique permettant d'ajuster des données incomplètes. Elle présente clairement un intérêt dans les études épidémiologiques, qui visent à estimer la prévalence d'une maladie à partir de sources de données courantes.

Principes de base

Cette méthode consiste à classer les données dans une table de contingence multiple incomplète, et à procéder à une analyse en employant une méthode statistique standard de modélisation log-linéaire (ou régression de Poisson). On estime, ensuite, le nombre manquant pour compléter la liste, d'après le modèle. On verra plus loin ce qu'est une table de contingence et l'intérêt de cette analyse.

La méthode de capture-recapture a été mise au point par des spécialistes de l'écologie animale pour estimer le nombre de cerfs, de poissons ou d'autres animaux dans une région. Elle consiste à capturer, marquer puis relâcher un échantillon d'animaux, puis à capturer un second échantillon et à déterminer le nombre d'animaux marqués issus du premier échantillon qu'il contient, d'où le nom de "capture-recapture".

En épidémiologie humaine, les séries d'animaux capturés et marqués sont remplacées par des listes de personnes "capturées" par certaines sources de données ordinaires. La capture-recapture est devenue une méthode importante en épidémiologie de l'usage des drogues, qui peut être adaptée et appliquée à la plupart des situations locales. Des exemples d'études par capture-recapture et des directives connexes qui ont été établis à la demande de l'Observatoire européen des drogues et des toxicomanies sont cités dans le présent document et sont également disponibles sous forme électronique (voir le chapitre I et l'annexe pour plus de détails).

Une introduction générale à l'estimation de la prévalence, et notamment à la méthode par capture-recapture, figure dans le *Bulletin des stupéfiants* des Nations Unies

[Hickman *et al.*, (2002)]. Pour de plus amples informations, trois excellentes études de l'historique et des emplois épidémiologiques de la méthode par capture-recapture sont aussi indiquées dans les références et pourraient être mises à disposition des utilisateurs, sous forme électronique, sur le site Web de l'Organisation des Nations Unies (voir le chapitre I et l'annexe).

Capture-recapture avec deux sources de données

Dans cette section, on présente d'abord l'exemple le plus simple, une étude menée avec deux sources de données, puis une étude plus complexe avec des sources de données multiples et, enfin, des études effectuées en l'absence de sources de données ordinaires.

On donnera d'abord un exemple, et les hypothèses théoriques nécessaires pour justifier les calculs viendront après une illustration de la méthode.

L'exemple concret ci-dessous est tiré d'une étude menée à Bangkok en 1991 par Mastro *et al.* Les deux sources de données ci-après ont été utilisées:

Étude de cas n° 5. Exemple concret tiré d'une étude menée à Bangkok (Étude simple par capture-recapture avec deux sources)

- a) Listes d'usagers d'opiacés inscrits dans des programmes de traitement à la méthadone, obtenues auprès de centres de traitement spécialisés de Bangkok. Il s'agissait de registres ordinaires de consultation pour avril et mai 1991;
- b) Personnes arrêtées par la police de Bangkok testées positives aux opiacés entre juin et septembre 1991. Les équipes d'enquêteurs se sont rendues 891 fois dans 72 postes de police pour y recueillir des échantillons d'urine et rassembler des informations sur les personnes arrêtées, afin d'identifier les usagers d'opiacés. Les personnes testées positives aux opiacés sont devenues la seconde source de données.

Les enquêteurs ont identifié 4 064 usagers d'opiacés dans les registres des centres spécialisés dispensant des traitements d'entretien à la méthadone, et 1 540 usagers d'opiacés parmi les 8 212 personnes couvertes par l'enquête qu'ils avaient effectuée dans les postes de police. Ces deux sources constituent des listes incomplètes des usagers d'opiacés à Bangkok à ce moment-là.

À partir du nom complet, du sexe et des dates de naissances, il a été établi que 171 personnes figuraient sur les deux listes, celle des usagers d'opiacés déclarés comme suivant un traitement à la méthadone et celle des personnes arrêtées dont les tests d'urine étaient positifs aux opiacés.

Le tableau 5 présente, sous la forme d'une table de contingence, le nombre de personnes apparaissant dans l'une ou l'autre des sources de données ou dans les deux, puis indique les calculs requis pour estimer la prévalence. Pour les deux sources de données (personnes se rendant dans un centre de traitement à la méthadone et personnes arrêtées dont les tests d'urine étaient positifs aux opiacés), on suppose que la proportion de la population totale d'usagers d'opiacés que l'on peut trouver dans les centres est la même que parmi les personnes arrêtées et celles qui ne l'ont pas été. Si l'on remplit la table sur la base de cette hypothèse, on estime à 36 600 (0,5 % de la population) le nombre d'usagers de drogues par injection à Bangkok en 1991.

Tableau 5. Estimation du nombre d'usagers d'opiacés à Bangkok, en 1991

Présentes dans S1 ^a	Présentes dans S2 ^b	Nombre de personnes recensées	Légende
Oui	Oui	171	Présentes à la fois dans S1, S2 (m)
Oui	Non	3 893	Uniquement présentes dans S1 (c)
Non	Oui	1 369	Uniquement présentes dans S2 (b)
Non	Non	??	"Invisibles" (x)
Total S1	Total S2	Total	Nombre d'usagers d'opiacés dans la population (N)
4 064	1 540		
(m + c)	(m + b)	-	

^aS1: Personnes suivant un traitement à la méthadone dans un centre.

^bS2: Personnes arrêtées dont le test d'urine est positif.

Le nombre d'usagers d'opiacés à Bangkok en 1991 est donc estimé de la façon suivante:

$$\text{Nombre de personnes observées} = m + b + c = 171 + 3\,893 + 1\,369 = 5\,433$$

$$\text{Nombre inconnu (x)} = 1\,369 * 3\,893 / 171 = 31\,166$$

$$\text{Estimation de la population} = 5\,433 + 31\,166 = 36\,599$$

Estimation (chiffre arrondi) du nombre d'usagers de drogues par injection à Bangkok en 1991 = 36 600 (entre 32 000 et 40 800)
(Intervalle de confiance de 95 %)

Dans cet exemple, les chercheurs ont retenu deux intervalles de temps distincts, avril-mai et juin-septembre, comme dans une étude écologique où la seconde capture a lieu après la première. Mais cela n'est pas indispensable en épidémiologie. Ils auraient pu utiliser deux sources de données portant sur la même période (voir ci-après la section intitulée "Postulats"). En outre, ils ont estimé le nombre total d'usagers pour l'ensemble de l'année 1991. En réalité, ils ont estimé le nombre total d'usagers d'opiacés par injection à Bangkok pendant la période étudiée (avril à septembre), mais il est raisonnable de supposer que cette estimation correspond au nombre annuel. En d'autres termes, on peut supposer que le nombre potentiel d'usagers d'opiacés non compris dans cette estimation – c'est-à-dire le nombre de nouveaux usagers et d'usagers qui ont cessé de l'être, qui sont décédés ou qui ont quitté la région – est négligeable (voir ci-après la section intitulée "Postulats").

Pour revenir aux définitions de l'introduction, le tableau ci-dessus est un exemple de table de contingence. Il s'agit, en l'occurrence, d'une table de contingence 2 x 2 à quatre cellules. Elle est incomplète, car nous ne connaissons pas le nombre de personnes qui n'apparaissent dans aucune des deux sources de données: les usagers de drogues par injection non observés (cachés). En d'autres termes, les deux sources de données ne déterminent pas "entièrement" la population cible, c'est-à-dire le nombre total d'usagers de drogues par injection.

Les méthodes de capture-recapture utilisent cette table pour estimer le nombre de personnes nécessaires pour compléter les deux listes. Les postulats nécessaires sont décrits plus loin, mais il y en a un qu'il convient d'évoquer tout de

suite. Dans une capture-recapture à deux échantillons, on suppose que le fait de figurer dans une source de données est indépendant du fait de figurer dans l'autre. Ainsi, dans l'étude ci-dessus, on a supposé que la probabilité qu'un usager ait suivi un traitement dans un centre était la même, que celui-ci ait été arrêté ou non; inversement, on a supposé que le fait qu'un usager avait suivi un traitement à la méthadone n'avait aucune influence sur la probabilité qu'il soit arrêté.

On a observé 5 433 usagers dans les deux sources de données combinées et estimé à 36 600 le nombre total d'usagers, c'est-à-dire qu'environ 1 usager d'opiacés sur 7 ($5\,433/36\,600$) a été observé dans le cadre de cette étude. Ce ratio est relativement élevé et sa validité repose sur le postulat que les deux sources de données sont indépendantes. Si elles ne l'étaient pas – ce qu'une étude sur deux échantillons ne permet pas de déterminer –, la marge d'erreur serait d'autant plus grande que le ratio entre la population observée et la population totale serait plus élevé.

Il est important de garder cela à l'esprit lorsqu'on planifie une étude. Le mieux serait de recueillir des données sur un grand nombre de personnes afin de réduire le ratio personnes observées/personnes non observées, mais cela n'est pas toujours possible dans les études sur l'usage problématique de drogues. On peut aussi opter pour une étude par capture-recapture avec sources de données multiples, comme celle présentée à la section suivante. Il est toujours possible de procéder à une étude sur deux échantillons, mais, dans ce cas, il est conseillé de chercher à savoir si les deux sources de données peuvent être considérées comme indépendantes, ou d'évaluer le degré de dépendance de façon à pouvoir traiter les résultats comme des estimations minimales s'il existe une dépendance positive entre les sources de données, et comme des estimations maximales s'il existe une dépendance négative entre elles (voir, plus loin, la section concernant l'indépendance).

Source: Mastro *et al.* (1994).

Estimation par capture-recapture avec sources de données multiples

Par sources de données multiples, on entend au moins trois listes: par exemple, usagers de drogues à problème en traitement, arrêtés, dans des foyers pour sans-abri, ou s'étant présentés à des services d'urgence. On effectue ensuite des recoupements entre ces sources de données pour identifier les personnes qui figurent au moins sur deux des listes.

Les données obtenues sont ensuite exploitées au moyen d'une régression de Poisson ou d'un programme de modélisation log-linéaire. L'action de remplir la table de contingence de l'exemple précédent est un exemple de modélisation log-linéaire extrêmement simple, qui ne requiert pas de programme, ni même d'ordinateur. On s'abstiendra de décrire les calculs pour lesquels sont utilisées deux sources de données ou plus pour estimer le nombre d'usagers de drogues à problème. Des logiciels statistiques peuvent effectuer ces calculs qui, présentés par écrit, semblent extrêmement complexes. Le postulat d'indépendance, dont le rôle a été essentiel dans les calculs de l'exemple précédent, peut maintenant prendre plusieurs formes; celles-ci sont examinées dans la prochaine section.

Les lecteurs intéressés par les équations d'estimation sont invités à consulter Bishop *et al.* (1975) et Hook et Regal (1995). En général, pour les calculs de capture-recapture avec sources de données multiples, il est recommandé de prévoir une assistance statistique sur place pour expliquer ou exécuter le programme de régression de Poisson (ou de modélisation log-linéaire) nécessaire à l'analyse des données. Les directives de l'Observatoire européen des drogues et des toxicomanies donnent des exemples détaillés de l'utilisation du logiciel de modélisation interactive linéaire généralisée (GLIM) (programme d'analyse statistique interactif) et du logiciel statistique pour sciences sociales (SPSS) pour la capture-recapture.

Dépendances entre les sources de données

On dit qu'une dépendance entre deux sources (S1 et S2) est positive si la probabilité de figurer dans S2 est plus grande pour une personne figurant dans S1 que pour une personne qui n'y figure pas, et qu'elle est négative si la probabilité de figurer dans S2 est moins grande pour une personne figurant dans S1 que pour une personne qui n'y figure pas. Ces dépendances peuvent être analysées en tant que termes d'interaction dans une analyse de régression de Poisson; il faut en tenir compte et les vérifier pour ajuster le modèle aux données. Il est possible d'analyser des interactions ou des dépendances plus complexes faisant intervenir, par exemple, une combinaison de trois sources de données. Un certain nombre de modèles de régression de Poisson différents peuvent être ajustés aux données, selon les dépendances prises en compte et celles qui sont omises. Le meilleur est généralement celui qui donne les meilleures prévisions, d'après des mesures types du manque d'ajustement.

Ajustement d'un modèle

Pour ajuster un modèle à des données de sources multiples, il faut procéder comme suit:

- a) Obtenir trois sources de données ou plus sur les usagers de drogues à problème;
- b) Déterminer les personnes qui apparaissent dans plusieurs sources de données, ainsi que les sources où elles apparaissent;
- c) Établir un tableau à entrées multiples pour analyse dans un fichier de données (voir l'exemple ci-après);
- d) Analyser le tableau au moyen d'une régression de Poisson;
- e) Ajuster le modèle en intégrant les interactions entre les sources de données correspondant aux dépendances potentielles;
- f) Sélectionner le meilleur modèle d'après des mesures types du manque d'ajustement (un statisticien prêtera conseil à ce sujet – voir les directives de l'Observatoire européen des drogues et des toxicomanies);
- g) Utiliser le modèle pour estimer la population non observée ainsi que pour calculer les intervalles de confiance;
- h) Répéter, si possible, l'ensemble de l'analyse pour différents sous-groupes (hommes, femmes, tranches d'âge différentes, etc.).

On trouvera ci-dessous un exemple d'analyse par capture-recapture avec sources multiples tiré d'une étude effectuée par Frischer *et al.* à Glasgow en 1993.

Étude de cas n° 6. Étude sur les usagers de drogues par injection à Glasgow (Écosse)
(Étude par capture-recapture avec sources multiples et stratification)

Bien qu'ayant en fait recueilli des données pour différents types de drogues (héroïne, autres opiacés, amphétamine, cocaïne et benzodiazépines), Frischer et ses collaborateurs ont retenu comme population cible les usagers de drogues par injection plutôt que les injecteurs d'héroïne, car, à Glasgow (Écosse), un grand nombre de drogues différentes sont prises par injection.

Les chercheurs ont recueilli des données à partir des quatre sources suivantes sur une période d'un an:

- a) Les tests positifs et négatifs de dépistage du VIH chez les personnes exposées, du fait de l'usage de drogues par injection;
- b) Les personnes qui se sont présentées dans des centres spécialisés afin de suivre un traitement (notamment pour un problème lié à l'héroïne, à d'autres opiacés, à la cocaïne, aux amphétamines et aux benzodiazépines);
- c) Les personnes arrêtées pour une infraction liée à une drogue (autre que le cannabis);
- d) Les personnes inscrites auprès de programmes locaux d'échange de seringues.

Dans ce type d'étude avec sources de données de recapture multiples, il faut faire des recoupements entre les quatre sources pour pouvoir déterminer le nombre d'usagers de drogues par injection à chaque combinaison de points d'échantillonnage. Au total, 3 444 entrées ont été recueillies: 508 auprès de la police (S1), 1 179 auprès de programmes d'échange de seringues (S2), 507 auprès de laboratoires effectuant des tests de dépistage du VIH (S3) et 1 250 auprès de centres de traitement spécialisés pour toxicomanes (S4). On a établi que 578 personnes figuraient dans plus d'une source de données (4 figuraient dans les 4 sources, 41 dans S2, S3 et S4, et 147 dans S2 et S4). Au total, 2 866 usagers injecteurs ont donc été observés.

Dans l'étude de Glasgow, on a utilisé un certain nombre de modèles allant du plus simple (aucun ajustement, toutes les sources étant supposées indépendantes) au plus complexe (ajusté pour tenir compte de toutes les interactions entre trois sources). Le modèle d'indépendance totale ne présentait pas une bonne concordance.

Le modèle le mieux ajusté aux données du tableau 5 tenait compte des interactions entre trois sources de données: tests de dépistage du VIH (S2), programmes d'échange de seringues (S3) et centres de traitement (S4). En revanche, la source de données S1 (personnes arrêtées par la police pour possession de drogues) n'apparaissait dans aucun des termes d'interaction de ce modèle, ce qui donne à penser qu'elle était indépendante de toutes les autres sources de données. L'interaction entre les sources était généralement positive, ce qui signifie que les personnes figurant sur une des listes avaient une probabilité plus grande de figurer sur une autre liste. Lorsque les sources de données sont dépendantes, en utilisant un modèle qui suppose que toutes les sources sont complètement indépendantes, on risque de sous-estimer la fraction non observée de la population. Dans cette étude, l'estimation fondée sur une indépendance complète était effectivement inférieure à celle obtenue avec le modèle final.

Les chercheurs ont estimé qu'il y avait, au total, 8 500 injecteurs à Glasgow, ce qui représente environ 1,35 % de la population adulte âgée de 15 à 54 ans. Suffisamment de données ont été recueillies au cours de cette étude pour permettre la réalisation d'autres analyses (c'est-à-dire d'autres travaux de modélisation) par sexe et par tranche d'âge, comme le montre le tableau 6. Cette méthode, que l'on a appelée stratification, est l'un des moyens de faire face au problème de l'hétérogénéité (voir ci-après la section intitulée "Postulats").

Tableau 6. Données de capture-recapture multisources tirées d'une étude sur les usagers de drogues par injection à Glasgow (Écosse), en 1990

Dans S1 ^a	Dans S2 ^b	Dans S3 ^c	Dans S4 ^d	Nombre d'usagers identifiés
Oui	Oui	Oui	Oui	4
Oui	Oui	Oui	Non	2
Oui	Oui	Non	Oui	13
Oui	Oui	Non	Non	56
Oui	Non	Oui	Oui	8
Oui	Non	Oui	Non	17
Oui	Non	Non	Oui	50
Oui	Non	Non	Non	358
Non	Oui	Oui	Oui	41
Non	Oui	Oui	Non	52
Non	Oui	Non	Oui	147
Non	Oui	Non	Non	864
Non	Non	Oui	Oui	116
Non	Non	Oui	Non	267
Non	Non	Non	Oui	871
Non	Non	Non	Non	-
Total S1	Total S2	Total S3	Total S4	Total population
508	1 791	507	1 250	-

^aS1 = usagers de drogues par injection arrêtés pour possession de drogues (sauf cannabis)

^bS2 = usagers de drogues par injection inscrits à un programme d'échange de seringues

^cS3 = usagers de drogues par injection ayant subi un dépistage du VIH

^dS4 = usagers de drogues par injection suivant un traitement spécialisé

Le modèle le mieux ajusté incluait des interactions entre les sources S2 et S3, S2 et S4, S3 et S4, et une triple interaction entre S2, S3 et S4; la source S1 était indépendante de toutes les autres. Ce modèle a donné les estimations suivantes, comme l'indique le tableau 7 ci-dessous.

Tableau 7. Estimation de l'usage de drogues par injection à Glasgow – Nombre d'usagers et prévalence

	Nombre d'usagers observés/connus	Nombre d'usagers non observés (estimation)	Nombre total d'usagers (estimation)	Prévalence (estimation, en pourcentage)
Total	2 866	5 628	8 494	1,4
Hommes	1 977	3 567	5 544	1,8
Femmes	889	2 349	3 238	1,0
15-19 ans	264	640	904	1,0
20-24 ans	1 137	1 613	2 750	2,6
25-29 ans	878	1 724	2 602	2,7
30-34 ans	342	796	1 138	1,4
35 ans et plus	245	1 273	1 518	0,6

Ce qui est assez exceptionnel dans l'étude de Glasgow, c'est que les données de la police étaient complètement indépendantes des autres sources. En outre, les trois autres sources de données, S2, S3 et S4, étaient toutes dépendantes. On aurait pu les regrouper en une seule source et effectuer une étude simple

à deux sources avec les données de la police. Il est intéressant, pour faire un exercice simple, de revenir sur l'étude de cas n° 1 plus haut et de calculer la prévalence en utilisant les calculs sur deux sources qui y sont décrits. Les données seraient les suivantes:

- a) Total des personnes traitées, des personnes ayant subi un test de dépistage du VIH et des personnes inscrites à un programme d'échange de seringues: 2 508;
- b) Nombre de personnes arrêtées par la police: 508;
- c) Nombre de personnes apparaissant dans les deux sources: 150.

Le résultat est le même que celui obtenu plus haut au moyen de la régression de Poisson avec des sources multiples. Pourquoi? Parce que les caractéristiques très particulières des résultats de l'analyse multisources montrent que l'on aurait pu aussi bien utiliser deux sources de données indépendantes. C'est en fait assez normal: les personnes répertoriées dans la liste des personnes traitées ont probablement plus de chances d'être inscrites à un programme d'échange de seringues, et sans doute plus encore si elles ont aussi subi des tests de dépistage du VIH; les trois sources sont interdépendantes. En revanche, il est plausible que les personnes arrêtées par la police représentent un échantillon aléatoire de tous les usagers, indépendamment du fait qu'ils aient suivi un traitement, participé à un programme d'échange de seringues ou subi un test de dépistage du VIH.

Ce qui est démontré ici, c'est qu'un chercheur n'aurait pas pu savoir que les données de la police étaient indépendantes et que toutes les autres sources de données étaient dépendantes, sans effectuer une étude multisources et analyser les données au moyen de modèles de régression de Poisson. Il se peut qu'une étude à deux sources donne la bonne réponse, mais aucun des postulats sur lesquels elle se fonde, même les plus évidents, ne peut être vérifiable.

Cet exemple montre encore autre chose: ce n'est que dans des circonstances assez particulières – telles que celles révélées dans cette étude – qu'il est judicieux de regrouper différentes sources de données comme si elles n'en constituaient qu'une seule. En général, même s'il est normal que différentes sources n'aient pas les mêmes chances de "capturer" les usagers de drogues à problème, il est important que tous aient les mêmes chances de figurer sur une liste donnée. Ce ne sera probablement pas le cas si l'on regroupe des listes sources très diverses comme celles des personnes ayant subi des tests de dépistage du VIH ou des personnes arrêtées par la police, et cela serait contraire au postulat d'homogénéité (voir ci-après la section intitulée "Postulats").

Source: Frischer et al., "Estimating the population prevalence of injection drug use..." (1993).

Quels logiciels utiliser

On peut conduire des études à deux échantillons avec un crayon et un papier, une calculatrice ou un tableur. En revanche, la plupart des méthodes par capture-recapture avec sources multiples nécessitent des logiciels statistiques. Bishop *et al.* (1975) donnent les formules permettant d'effectuer les calculs à la main. On trouve, dans les directives de l'Observatoire européen des drogues et des toxicomanies, des exemples d'utilisation des logiciels GLIM et SPSS.

Les sources de données et la recherche des éléments communs

Critères de sélection des sources de données à utiliser

Dans les études par capture-recapture, les meilleures sources de données sont celles qui:

- a) Définissent clairement la population à estimer (par exemple, les usagers d'héroïne par injection ou les usagers de drogues à problème);
- b) Recueillent beaucoup de données permettant d'effectuer des recoupements;
- c) Recueillent des variables de stratification possibles;
- d) Fournissent les données sous forme électronique pour éviter d'avoir à les compiler à la main.

Malheureusement, cela est rarement le cas. Il faudra donc veiller à ce que la qualité des sources de données soit suffisante et, en particulier, vérifier les points suivants:

- a) Permettent-elles d'identifier la population cible à estimer?
- b) Quels sont les identificateurs recueillis et quelles données seront communiquées par ceux qui les détiennent?
- c) Comment les données seront-elles recueillies ou fournies?
- d) Combien de cas seront fournis?
- e) Quelle sera la correspondance entre ces sources de données?

Le dernier point se réfère à l'un des postulats de la capture-recapture (voir ci-après la section intitulée "Postulats"), à savoir que les sources de données sont représentatives de la population cible. Si possible, il faudrait donc sélectionner les sources de données dans le domaine de la justice pénale et des soins.

Les sources de données qui peuvent être disponibles sont indiquées plus haut (voir "Technique de multiplication simple"). La meilleure solution est de faire un inventaire, dans la région couverte par l'étude, pour déterminer celles qui sont disponibles sur place et s'il existe d'autres sources potentielles. Ces sources peuvent servir de référence pour les méthodes multiplicateur/référence lorsqu'elles sont agrégées (groupées) et être utilisées dans des études de capture-recapture lorsqu'elles sont désagrégées (dissociées). On se souviendra que, pour effectuer des estimations par capture-recapture, il n'est pas nécessaire que les sources de données soient complètes (qu'elles couvrent tous les cas possibles), mais il faut qu'elles soient exactes et fiables (autrement dit, les identificateurs doivent être corrects et les informations relatives aux drogues doivent être exactes).

On peut, dans une certaine mesure, regrouper plusieurs sources de données portant sur une population réduite, même si cela comporte aussi des risques. Les études par capture-recapture portant sur plus de cinq sources de données peuvent être complexes à analyser, en raison du grand nombre de modèles qu'il est possible d'ajuster.

Recherche des cas communs (doublons)

À moins de devoir traiter des milliers d'enregistrements, le mieux est d'effectuer cette recherche manuellement. Mais, il peut être utile de se servir de tableurs ou de bases de données pour trier les sources de données de différentes façons, afin d'identifier les personnes qui figurent dans deux listes à la fois. Par exemple, on peut prendre deux listes triées par sexe et par date de naissance pour voir si quelqu'un a le même nom ou les mêmes initiales, ou examiner des listes triées par sexe et par nom pour voir si la date de naissance est la même.

Il est bon de définir ce qui constitue un doublon, car les différences entre les sources peuvent être dues à des erreurs de saisie ou au fait que les personnes ont indiqué des noms légèrement différents. Par exemple, quelqu'un peut apparaître sous le nom de William Shakespeare dans une source et de Bill Shakespear dans une autre, la date de naissance indiquée étant le 26 avril 1563 dans l'une et le 20 avril 1563 dans l'autre. Plus on dispose de données sur chaque personne, plus il est facile d'apparier les données, mais on doit toujours tenir compte des erreurs éventuelles entre les sources de données. En général, si la date de naissance et le sexe sont identiques et que l'un des noms est le même, on considère, dans la plupart des études, qu'il s'agit de la même personne.

Quels que soient les critères retenus, il est important qu'ils soient clairs et précis. On peut, par exemple, adopter une définition stricte, mener l'analyse, puis adopter une définition plus souple et comparer les estimations de prévalence. Ainsi, dans l'étude de Bangkok, Mastro *et al.* ont retenu six critères fondés sur six informations différentes pour chaque personne, comme l'indique le tableau 8.

Tableau 8. Critères d'appariement utilisés pour les estimations de la prévalence dans l'étude de Bangkok

Critères d'appariement	Sexe	Prénom	Nom de famille	Âge	Date de naissance	Numéro d'identité en Thaïlande
1	Identique	Identique	Identique	Identique	Identique	Identique
2	Identique	Identique	Identique	Identique	Identique	Manquant
3	Identique	Identique	Identique	Identique	Similaire	-
4	Identique	Similaire/identique	Similaire/identique	Identique	Similaire/identique	-
5	Identique	Différent	Similaire/identique	Identique	Similaire	-
6	Identique	Similaire/identique	Similaire/identique	Différent	Différent	-

Le terme "similaire", dans le tableau 8 ci-dessus, est synonyme de "proche" et signifie, par exemple dans le cas des dates, que le mois et l'année sont les mêmes. Les critères 1 à 4 ont été utilisés pour estimer la prévalence dans le cadre de l'étude de Bangkok.

Dans l'étude de Glasgow, Frischer *et al.* ont utilisé les critères d'appariement suivants, en se fondant sur cinq informations, comme indiqué dans le tableau 9 ci-dessous.

Tableau 9. Critères d'appariement utilisés pour les estimations de la prévalence dans l'étude de Glasgow

Critères d'appariement	Sexe	Initiale du nom de famille	Initiale du prénom	Date de naissance	Code postal
1	Identique	Identique	Identique	Identique	Identique
2	Identique	Identique	Identique	Identique	Similaire/différent
3	Identique	Identique	Identique	Similaire	Similaire/différent
4	Identique	Identique	Similaire/différent	Identique	Similaire/différent

Études en l'absence de sources de données ordinaires

Que faire s'il n'existe pas de sources de données locales, si aucune source ne recueille d'informations identifiant les usagers, ou si les règles de confidentialité empêchent d'utiliser ces informations? Renoncer, opter pour une méthode différente ou bien envisager, éventuellement, de faire des calculs de capture-recapture par le biais d'enquêtes sur les usagers de drogues à problème? Une telle manière de procéder est tout à fait indiquée lorsque la population est géographiquement discrète et présente une caractéristique commune qui peut être utilisée pour l'identifier: par exemple, personnes qui s'injectent la drogue à la sauvette ou prostitué(e)s qui s'injectent la drogue. Deux exemples de cette méthode sont présentés dans les études de cas n° 7 et n° 8.

La méthode employée au Bangladesh est une étude de capture-recapture simple avec deux sources de données (voir *supra*). Bien qu'aucune information permettant d'identifier les individus n'ait été recueillie, elle a permis d'identifier des doublons, de sorte qu'il a été possible d'effectuer les calculs exposés ci-dessus pour estimer la population totale. Il s'agit bien d'une étude de capture-recapture, car les cas communs aux sources de données 2 et 1 ont pu véritablement être vérifiés.

Dans le cadre de l'étude de Glasgow, on a recueilli certains facteurs permettant l'identification des cas communs observés au cours de différentes nuits d'enquête sur le terrain, mais la méthode utilisée reposait sur un ensemble différent de méthodes statistiques pour estimer la taille de la population. Il est préférable de ne recourir à cette méthode qu'en consultation avec un statisticien ayant l'expérience des modèles ouverts de capture-recapture et des travaux d'estimation.

Étude de cas n° 7. Estimation de la prévalence des prostitué(e)s à Dhaka (Bangladesh) (Capture-recapture en l'absence de sources de données ordinaires)

Cette méthode a été utilisée, en l'occurrence, pour estimer le nombre de prostitué(e)s dans les rues de Dhaka (Bangladesh). En l'absence de données secondaires sur la population de prostitué(e)s dans la ville et de sources de données ordinaires identifiant cette population, il a été mené une enquête ethnographique approfondie sur le terrain. Plusieurs groupes d'informateurs clefs ont été interrogés, notamment les prostitué(e)s, les souteneurs, les conducteurs de rickshaw, les chauffeurs de taxi, les policiers et les *mastans* (voyous) locaux. Il est apparu que pratiquement toutes les personnes qui se prostituaient dans cette ville racolaient leurs clients dans des rues, lieux et parcs déterminés, surtout dans la soirée. Il a été minutieusement dressé une carte de la ville indiquant les secteurs où les prostitué(e)s racolaient les clients.

Avec l'aide des personnes qui se prostituent, des cartes de couleur (rouge la première fois) ont été distribuées dans tous les lieux de prostitution notoires de la ville, dans la soirée et jusqu'au milieu de la nuit. Cette carte de couleur pouvait être utilisée pour passer un contrôle médical gratuit dans l'un des centres de soins de santé primaires et de santé génésique de la ville. Comme les cartes étaient numérotées, un échantillon du nombre de cartes distribuées, représentant l'échantillon de capture, a été aisément obtenu. La même procédure a été répétée sept jours plus tard. Cette fois-là, il a été distribué des cartes de couleur verte, qui pouvaient elles aussi être utilisées pour passer une consultation médicale gratuite. La seconde fois, il n'a été posé aux personnes qui se prostituaient qu'une seule question, pour savoir si elles avaient déjà reçu une carte similaire de couleur rouge. Le nombre estimatif de prostitué(e)s racolant dans la rue – 5 000 environ – a été tiré des données en effectuant le calcul simple expliqué dans l'étude de cas n° 5.

Source: Projet SHAKTI, CARE Bangladesh, Dhaka.

Étude de cas n° 8. Étude de Glasgow sur les prostituées (Capture-recapture en l'absence de sources de données ordinaires)

À Glasgow, les enquêteurs se sont intéressés au principal quartier chaud de la ville. Ils ont déterminé combien de prostituées racolaient dans la rue pendant une certaine période de temps, en dénombrant celles qui travaillaient chaque nuit et en vérifiant si elles avaient déjà été observées les nuits précédentes. Sur une période totale de sept mois, 1 145 contacts ont été établis avec 206 femmes (dont 147 qui s'injectaient des drogues, soit 71 %). Des facteurs d'identification comprenant les initiales du nom et l'année de naissance ont été recueillis pour chaque femme contactée, afin d'obtenir son "historique de capture" pour la période considérée. L'analyse de ces historiques par des "méthodes ouvertes de capture-recapture" a montré que la population de prostituées était constante (environ 200 par nuit), mais que son taux de renouvellement était d'environ 8 % par semaine, ce qui donnait un total annuel d'environ 1 150 prostituées.

Source: Frischer *et al.*, "Estimating the population prevalence of injection drug use..." (1993).

Crédibilité de l'estimation

Comment savoir si les estimations sont crédibles? La réponse pure et simple est que c'est impossible, car des difficultés peuvent toujours surgir (voir ci-après la section intitulée "Postulats"). Mais on peut s'en remettre à l'appréciation de la situation locale que font les enquêteurs eux-mêmes pour juger si les estimations sont crédibles. C'est ainsi que procèdent les spécialistes de l'écologie animale pour pallier le caractère incertain de la méthode. Il convient donc de se poser les questions suivantes:

- a) Les estimations sont-elles conformes aux résultats attendus?
- b) Concordent-elles avec d'autres estimations reposant sur des données probantes qui ont été établies à l'aide d'autres méthodes ou pour des années différentes?

- c) Sont-elles exagérément basses, ou en ressort-il au contraire qu'une personne sur cinq s'injecte des drogues? Dans ce cas, elles sont probablement fausses.
- d) Un intervalle de confiance de moins de 95 % est-il gênant? Dans l'affirmative, il peut être recouru à d'autres méthodes, mais à condition de vérifier l'exactitude des sources de données et les recoupements.
- e) Dans les études de capture-recapture multisources, les dépendances entre les sources de données sont-elles crédibles?

Étude de cas n° 9. L'étude de Jersey (Usagers d'héroïne donnant de faux noms)

L'avantage de l'étude par capture-recapture multisources est que, parfois, les problèmes éventuels ne tiennent qu'à une seule source de données – auquel cas rien n'empêche d'abandonner cette source et de recalculer une estimation. Par exemple, un exercice d'estimation de la prévalence a été effectué dans le cadre d'une évaluation rapide à Jersey. Jersey est une île relativement riche, située dans la Manche entre l'Angleterre et la France et qui compte moins de 100 000 résidents. Cinq sources de données ont été utilisées: enquête sur les usagers de drogues par injection, héroïnomanes en traitement, registre des toxicomanes par lequel tous les médecins déclaraient aux autorités les opiomanes et les cocaïnomanes, personnes admises pour surdose d'héroïne dans les services d'urgence et arrestations effectuées par la police pour des infractions liées à l'héroïne. Le chiffre estimé à l'aide du premier modèle reposant sur les cinq sources était de plus de 2 000 usagers de drogues par injection, ce qui donnerait à Jersey une prévalence très élevée (plus de 2 % de la population totale et plus de 4 % des personnes âgées de 15 à 54 ans).

Les auteurs de l'étude ont soupçonné que les personnes interrogées dans le cadre de l'enquête donnaient de faux noms. Cela a été corroboré en comparant le pourcentage de personnes figurant à la fois dans les données de l'enquête et dans les registres des services de traitement des toxicomanes (10 %) avec le pourcentage de répondants à l'enquête ayant indiqué avoir été en traitement au cours de l'année précédente (plus de 30 %). Il y avait clairement un problème, dans la mesure où l'on pouvait s'attendre à un niveau de concordance d'environ 30 % entre les deux sources de données (services de traitement et enquête). Une fois les données de l'enquête abandonnées, il a été obtenu une estimation beaucoup plus raisonnable et crédible d'environ 750 usagers (0,8 % de la population totale).

Voilà un exemple d'"erreur de classification" (voir ci-après la section intitulée "Postulats"). On voit aussi tout l'intérêt que présentent l'expérience et l'appréciation de la situation locale si l'on veut pouvoir identifier les estimations improbables et les biais potentiels.

Source: Fitch *et al.* (2001). Responding to drug and alcohol use in Jersey: key findings report. Drug Strategy Unit, States of Jersey.

Postulats

Enfin, les postulats sur lesquels reposent les études de capture-recapture sont présentés ci-dessous, mais de façon succincte, car les présentes directives se veulent plus pratiques que théoriques. Il est, cependant, important de comprendre ces postulats pour pouvoir interpréter les conclusions de ces études. Les incidences de chaque postulat sur la méthode de capture-recapture seront également examinées.

Population figée – Ni décès, ni nouveaux cas, ni immigration/émigration dans la région considérée durant l'étude

De telles conditions étant évidemment impossibles à garantir, l'étude doit porter sur une période brève par rapport à la durée pendant laquelle les toxicomanes font un usage problématique de drogues. Par exemple, une durée d'un an est normale et ne s'écarte pas trop de ce postulat. Par contre, une étude sur la base de données, portant sur une période de cinq années ou davantage, risque d'inclure de nombreux usagers de drogues à problème qui n'ont pas fait usage de drogues pendant l'intégralité de la période couverte par l'étude, car: ils ont cessé de se droguer; ils sont décédés ou ont quitté la région; ils se droguent depuis peu de temps; ils se sont récemment établis dans la région. En outre, le postulat implique que, quand il est recouru aux techniques de capture-recapture à un moment où l'incidence est élevée – c'est-à-dire à un moment où l'usage de drogues par injection ou de drogues posant problème est en augmentation rapide –, l'étude devrait couvrir une période aussi brève que possible, ou être retardée en attendant que la progression se soit stabilisée.

Il est possible, grâce à des modèles ouverts de capture-recapture, d'estimer la dynamique de la population, notamment les migrations, les décès et les naissances. Utilisés couramment en écologie animale, ces modèles supposent toutefois un ensemble d'équations et un savoir-faire statistique différents. De tels modèles n'auraient été employés que dans une seule étude sur les drogues (étude de cas n° 8 précitée).

Pas d'erreur de classification – Les doublons sont correctement identifiés

Dans les études humaines, cela suppose que l'on identifie correctement quelqu'un qui figure sur plusieurs listes. Il faut veiller à ce que les données recueillies soient exactes et fiables, en espérant que les faux noms, éventuellement fournis par les usagers de drogues à problème, seront soit éliminés, soit concordants dans les sources de données. Mais, comme pour des raisons de confidentialité il peut se révéler impossible de recueillir et d'utiliser le nom complet des personnes, il faut recueillir, à la place, des informations permettant d'effectuer des recoupements entre les sources de données, sans révéler l'identité de la personne (voir au chapitre II la section intitulée "Enquêtes par capture-recapture" sur les problèmes d'éthique dans la collecte de données).

Dans les exemples cités, Mastro *et al.* ont utilisé le nom complet pour identifier les doublons; Frischer *et al.* ont par contre utilisé, comme dans beaucoup d'études de capture-recapture sur l'usage problématique de drogues, la date de naissance, le sexe et les initiales (première lettre du prénom et première lettre du patronyme). Dans l'exemple du Bangladesh, en l'absence de sources de données nominatives, l'étude a été fondée sur la "recapture" des prostitué(e)s racolant dans la rue.

Dans l'étude sur Jersey, le biais dû à l'erreur de classification conduisait à surestimer la prévalence de l'usage d'héroïne, ce qui est le résultat auquel on peut normalement s'attendre. Comme certaines des personnes interrogées dans le cadre de l'enquête ont donné de faux noms, qui ne pouvaient pas être mis en correspondance avec les noms répertoriés par les autres sources de données, un certain nombre de doublons vrais sont passés inaperçus.

Pas d'hétérogénéité – Tous les usagers de drogues à problème ont la même chance d'être répertoriés par une source de données

Les études humaines sont inévitablement entachées d'hétérogénéité. Certains usagers de drogues à problème, du fait de leur sexe, de leur âge, de leur appartenance ethnique ou d'une autre caractéristique, ont tout simplement plus de probabilité que d'autres d'être répertoriés par une source de données. Par exemple, dans certains pays, les hommes jeunes de race noire qui font un usage problématique de drogues ont plus de probabilité d'être arrêtés que d'autres usagers de drogues à problème. Il faut donc être conscient de cette hétérogénéité. Si l'on dispose de suffisamment de données sur un nombre suffisant de personnes, il est possible de stratifier la population pour faire l'estimation. Des estimations séparées peuvent être faites pour les hommes et les femmes, par tranche d'âge ou par groupe ethnique. À condition de disposer de soutien statistique, il est possible de procéder à l'ajustement de modèles plus complexes (voir ci-après la section intitulée "Utilisation de techniques de modélisation évoluées").

Représentativité – Les sources de données sont représentatives de la population d'usagers de drogues à problème qu'il s'agit d'estimer

Il y a des différences entre les études de capture-recapture sur l'usage problématique de drogues et d'autres études épidémiologiques, comme celles sur le diabète, qui utilisent comme sources de données les listes de cas diagnostiqués. À supposer qu'il n'y ait pas d'erreur de classification, l'étude a donc pour objet d'estimer le nombre de diagnostics pertinents, par rapport à chacune des sources de données spécifiques utilisées dans l'étude, qui n'ont pas été recensés. Il s'agirait donc d'une étude du "sous-recensement".

Les sources de données devant être utilisées définissent-elles de manière adéquate la population cible d'usagers de drogues à problème? L'usage problématique de drogues ne se définit pas, par exemple, comme le diabète. Premièrement, comme cet usage est défini par un certain nombre de problèmes d'ordre clinique, social et criminel, ces sources doivent être aussi inclusives que possible, l'idéal étant de recueillir des données provenant des centres de traitement et de la justice pénale. En second lieu, aucune combinaison de listes, même si elle est complète par rapport à son objectif spécifique, ne pourra jamais recenser le nombre total d'usagers de drogues à problème dans la population. C'est la raison pour laquelle les estimations de la prévalence de l'usage de drogues sont utiles, pleines d'intérêt et plus difficiles à réaliser que pour beaucoup d'autres problèmes de santé publique.

Pas de dépendance dans les études avec deux sources de données – Les deux sources de données sont indépendantes l'une de l'autre

Les études avec deux échantillons sont plus faciles à réaliser –, mais ce postulat est leur principale limite. On ne sait pas, en général, si deux sources de données sont indépendantes l'une de l'autre, et il n'existe pas de moyen de le vérifier. Dès lors que la direction et l'importance approximative du biais sont connues, les études avec deux échantillons peuvent donner des indications très utiles sur la taille minimale ou maximale de la population. En effet, la "dépendance négative" – le fait qu'une personne répertoriée par la source de données A a moins de probabilité d'être répertoriée aussi par la source de données B – signifie que la taille véritable de la population est surestimée. La dépendance positive – le fait qu'une personne répertoriée par la source de données A a plus de probabilité d'être répertoriée aussi par la source de données B – conduit à une sous-estimation de la prévalence vraie.

Les études avec trois échantillons permettent de faire l'économie de ce postulat, mais cela accroît la complexité des calculs. Des modèles log-linéaires sont utilisés pour analyser les données, ce qui nécessite, dans la plupart des cas, des logiciels statistiques et, éventuellement, un soutien statistique. Les études de capture-recapture multisources postulent qu'il n'y a pas d'interaction entre toutes les sources – c'est-à-dire que si trois sources de données sont utilisées, il n'existe pas de dépendance à triple sens entre l'ensemble des sources. Toutefois, s'il apparaît qu'il existe des relations entre chaque paire de sources, ce postulat d'une absence de relations à triple sens perd de sa crédibilité.

Au fur et à mesure que le nombre de sources de données utilisées augmente, le nombre de modèles possibles qui doivent être vérifiés augmente encore plus rapidement. Ainsi, avec trois sources de données sept modèles peuvent être ajustés, mais avec cinq sources de données le nombre de modèles potentiels passe à 31. Le seul modèle qui ne peut pas être ajusté est le modèle dit "entièrement saturé", où toutes les sources de données sont en interaction les unes avec les autres. Dans ce cas de figure, avec trois sources de données une interaction triple ne peut pas être ajustée, et avec quatre sources de données une interaction quadruple ne peut pas être ajustée, etc.

Il est conseillé, dans le cadre de toute étude d'estimation par capture-recapture avec sources multiples, d'obtenir, si possible, au moins une source de données qui ait de bonnes chances d'être indépendante de toutes les autres.

Résumé

Pour récapituler: après avoir obtenu des sources de données, il faut rechercher les éléments communs qu'elles contiennent, placer ensuite les données dans un tableau à entrées multiples pour arriver à une estimation du nombre d'utilisateurs de drogues à problème, puis transformer celle-ci en une estimation de la prévalence, c'est-à-dire un pourcentage de la population totale, de la population adulte (par exemple de 15 à 54 ans), ou encore de la population masculine et de la population féminine et par tranches d'âge, s'il a été fait des estimations pour ces sous-groupes.

Les points essentiels peuvent être résumés sous la forme des conseils pratiques suivants:

- a) Soyez conscients des problèmes potentiels;
- b) Essayez, si possible, de trouver des sources de données multiples;
- c) Sélectionnez les sources de données soigneusement (essayez d'obtenir différents types de sources de données);
- d) Vérifiez si les noms ont été fournis et s'ils sont fiables;
- e) Recherchez des éléments qui corroborent l'information – l'estimation est-elle crédible et concorde-t-elle avec d'autres estimations effectuées par d'autres méthodes?
- f) Améliorez les sources de données qui seront utilisées la fois suivante.

Utilisation de techniques de modélisation évoluées

Il existe d'autres techniques de modélisation, plus complexes, qui nécessitent une analyse statistique plus poussée que les méthodes multiplicateur/référence et capture-recapture exposées ci-dessus. Si l'on souhaite y recourir, il est vivement conseillé de se faire aider par un statisticien expérimenté. Par ailleurs, elles requièrent, d'une manière générale, davantage de données, que celles-ci proviennent de sources ordinaires ou qu'elles soient recueillies spécialement pour l'estimation. On trouvera, dans la présente section, des renseignements succincts concernant certaines de ces techniques, ainsi que des références à l'intention de ceux qui souhaiteraient les examiner plus en détail. Ces techniques sont les suivantes: modélisation à covariables par capture-recapture; études événementielles perfectionnées, fondées sur la technique du multiplicateur; distribution de Poisson tronquée; modélisation par rétrocalcul; modélisation dynamique.

Méthodes par capture-recapture utilisant des modèles à covariables

Tilling et Sterne (1999) ont mis au point une variante de la méthode par capture-recapture, qui permet d'intégrer des covariables (âge, sexe, lieu de résidence, appartenance ethnique, etc.) à la procédure de modélisation pour ajuster l'estimation de la prévalence totale et obtenir des estimations distinctes pour les covariables. Ainsi, cette procédure permet de vérifier et de corriger l'hétérogénéité. Normalement, pour tenir compte de l'hétérogénéité, on stratifie l'ensemble de données en sous-groupes et on applique des modèles différents. Toutefois, avec chaque strate, on utilise davantage d'informations, ce qui n'est viable que si l'on dispose d'un nombre suffisant d'enregistrements pour que des modèles par sous-groupes puissent être appliqués. Le modèle à covariables est beaucoup plus efficace, mais il est aussi plus complexe et exige un appui statistique. De plus, aucune étude épidémiologique sur l'abus de drogues faisant appel à cette méthode n'a été publiée à ce jour. C'est donc une méthode qu'il vaut mieux réserver pour l'avenir.

Modèles événementiels et modèles connexes

Simeone *et al.* (1997) ont présenté une méthode du multiplicateur modifiée et améliorée qu'ils ont expérimentée à Chicago (États-Unis) et qui est actuellement appliquée dans une ville mexicaine frontalière (information fournie par Elena Medina Mora). Cette méthode utilise des "événements", comme multiplicateur et référence. Avant d'en exposer les grandes lignes, il convient d'attirer l'attention sur la distinction entre les données d'observation qui recensent les usagers de drogues (c'est-à-dire des individus) et celles qui recensent des événements répétés (c'est-à-dire le nombre de fois que ces individus ont été observés). Les événements pris en considération sont en général les prises de contact avec des institutions (par exemple, nombre de séjours dans des centres de traitement avec indication non seulement du nombre de personnes reçues, mais aussi du nombre de fois que chaque personne a été reçue dans l'année) ou les décès, bien que ceux-ci fassent évidemment partie des "événements" qui, par nature, ne se répètent pas. Un dénombrement de personnes peut généralement être considéré comme une donnée événementielle simple lorsqu'on se fonde sur le critère déjà/jamais (au cours de la période considérée) et non sur un chiffre. Il est parfois plus facile d'échantillonner des événements que des individus quand on veut recueillir auprès des toxicomanes des données sur leur histoire personnelle ou sur leurs antécédents en matière de consommation de drogues. Ainsi, si l'on prend un échantillon de personnes séjournant dans un endroit fixe tel qu'un centre de traitement, il est ensuite relativement aisé d'organiser ces séjours (événements) en un échantillon aléatoire de tous les séjours dans des centres de traitement, ou tout au moins d'obtenir une bonne approximation. Par contre, dans le cas d'un échantillon de personnes, même si l'on ne compte chaque individu qu'une seule fois, on a beaucoup plus de chances d'interroger les personnes qui se rendent fréquemment dans ces centres que celles qui n'y vont qu'occasionnellement. De plus, le fait que certaines sources de données systématiques ne recensent que des événements et ne peuvent pas dénombrer des individus constitue un argument important en faveur de cette méthode.

Dans l'étude de Chicago, on a retenu, comme références, le nombre de fois qu'un héroïnomanes ou un consommateur de crack a fréquenté des foyers sociaux, a été incarcéré et a reçu un traitement, et, comme multiplicateurs, les taux annuels correspondants. Pour recueillir les données nécessaires pour définir les multiplicateurs et ajuster les références, on a interrogé des échantillons de personnes fréquentant un foyer, de personnes incarcérées et de toxicomanes suivant un traitement spécialisé. Leurs déclarations, associées à une analyse capillaire destinée à valider et à ajuster ces déclarations, ont servi à estimer le nombre d'incarcérations, de séjours en foyer et de traitements concernant des héroïnomanes ou des consommateurs de crack.

Des techniques avancées d'analyse statistique et de modélisation ont été appliquées pour établir des estimations "non biaisées" des taux d'événements. Étant donné qu'une enquête est toujours biaisée, les chercheurs ont utilisé des informations provenant des trois enquêtes susmentionnées, afin d'établir une estimation globale du taux d'événements concernant les héroïnomanes et les consommateurs de crack.

Cette méthode est intéressante en principe, mais elle doit être expérimentée dans d'autres endroits. Les résultats de l'étude en cours au Mexique devraient être disponibles sous peu.

Modèles stochastiques de comptage d'événements (modèles de Poisson tronqués)

Plusieurs modèles de Poisson tronqués ont été proposés. Ces modèles utilisent des informations sur des événements répétés pour estimer la population qui n'a été concernée par aucun événement.

Il est à noter que ces modèles permettent de mieux tirer parti des données, à condition, bien sûr, que le comptage soit correct (ce qui suppose en général que l'enquêté puisse se souvenir de points précis). Les modèles de Poisson tronqués sont en fait les plus simples des modèles stochastiques appliqués dans ce domaine.

Postulats restrictifs

Pour que la distribution de Poisson puisse être adoptée comme modèle stochastique (Hay et Smit, à paraître), il faut poser les postulats suivants:

- a) Les événements répétés sont indépendants des antécédents de l'individu, c'est-à-dire que la probabilité qu'il soit arrêté est la même pendant toute la durée de l'étude, indépendamment des arrestations dont il a déjà fait l'objet;
- b) Tous les individus sont également susceptibles de provoquer un événement, bien que l'on puisse assouplir ce postulat, non sans difficultés, s'il y a un nombre suffisant d'individus présentant des antécédents communs (mais, même dans ce cas, il faut définir certains postulats concernant les antécédents des individus n'ayant provoqué aucun événement, car, si la probabilité que certains individus soient arrêtés est quasiment nulle, la procédure sera nécessairement entachée d'une erreur de comptage ou d'estimation).

Ces postulats étant irréalistes et essentiellement invérifiables, il est dangereux d'utiliser cette méthode pour estimer le nombre de toxicomanes "occultes", et elle est sans doute moins fiable même qu'une enquête par capture-recapture fondée sur deux sources.

Modélisation par rétrocalcul

Law *et al.* (2001) ont adapté des méthodes reposant sur le rétrocalcul pour estimer l'incidence et la prévalence de la consommation d'héroïne. La technique de rétrocalcul, élaborée aux fins des enquêtes épidémiologiques sur le sida, part du principe que l'incidence du résultat pathologique final considéré (le sida) et le processus aboutissant à ce résultat final (l'infection par le VIH) sont liés par le temps d'incubation. Si l'on connaît deux de ces éléments, on peut estimer le troisième. Normalement, on suppose que la distribution du temps d'incubation et l'incidence du résultat final sont connues, et le processus d'infection correspondant à l'incidence observée est estimé. Cette estimation est alors utilisée avec les données correspondantes sur le temps d'incubation pour prévoir l'incidence et la prévalence du résultat final. Dans le cas des études épidémiologiques sur l'abus de drogues prises comme exemples, le résultat final dont on a observé l'incidence était la mort par surdose d'opiacés, les tendances diachroniques étant déduites des statistiques systématiques de la mortalité. La distribution de l'incubation correspond à la distribution du temps écoulé entre le début et la cessation des injections, que cette cessation résulte d'une surdose mortelle ou d'un arrêt volontaire. Cela requiert une somme considérable de données, notamment des statistiques fiables sur la mortalité, afin de recenser le nombre de décès par surdose d'opiacés, et des données sur la mortalité imputable à des surdoses d'opiacés et à d'autres drogues consommées par injection et les taux de cessation des injections.

Modèles dynamiques

Enfin, les modèles dynamiques constituent un volet en expansion rapide de la recherche sur les méthodes d'estimation de la prévalence. D'une manière générale, ces modèles exigent une somme considérable de données que l'on peut qualifier d'indicateurs en matière de drogues provenant de sources aussi variées que possible. Ces données sont analysées ensemble au moyen d'un modèle dynamique unique, dont la structure principale détermine comment des données temporelles ponctuelles peuvent conduire à des observations ultérieures. Il s'agit là de modèles spécialisés qui doivent être adaptés à chaque situation par des experts.

Estimations nationales de la prévalence établies par extrapolation à partir d'estimations locales

Dans les sections précédentes, des méthodes spécifiques d'estimation de la prévalence ont été décrites. Les méthodes d'extrapolation ne sont pas, quant à elles, limitées

à l'estimation de la prévalence, mais elles permettent, à partir d'informations connues sur la prévalence locale, de déduire des informations concernant des régions pour lesquelles on n'en dispose pas.

Dans la plupart des cas, les méthodes spécifiques, dont il est question ci-dessus, sont plus faciles à appliquer au niveau local (c'est-à-dire dans une région de superficie restreinte, voire dans une seule ville). Cela s'explique, en partie, par le fait que les sources de données requises pour estimer la prévalence de l'usage de drogues sont souvent plus faciles à obtenir et à manipuler sur le plan local. Ainsi, il faut, lorsqu'on procède par capture-recapture, confronter une liste de noms d'usagers – ou d'autres facteurs d'identification –, tirée d'une source de données, avec une liste de noms provenant d'une autre source, ce qui peut être une tâche ardue, même au niveau local, si l'on ne dispose pas de moyens informatisés, et quasiment impossible au niveau national. Par ailleurs, comme il est probable que les usagers locaux ne sont recensés que par des sources locales, les rechercher sur des listes nationales serait une perte de temps.

De plus, les données disponibles et les méthodes utilisées pour les recueillir ne sont parfois cohérentes qu'au niveau local ou régional. Ainsi, si l'on retient comme source la fréquentation de centres de traitement, les données disponibles pour les zones fortement urbanisées seront plus nombreuses que celles concernant les zones rurales. Il se peut aussi que, dans certaines régions, il n'existe pas de centre de traitement spécialisé, de sorte que l'utilisateur se rend plutôt dans un centre hospitalier public ou dans un établissement analogue, dont les dossiers sont moins précis. En particulier, lorsqu'on a affaire à des dispensaires privés, la qualité des dossiers peut varier considérablement d'une région et d'un établissement à l'autre, de sorte que les informations ne sont ni homogènes, ni concordantes.

Une autre question, enfin, risque de se poser, celle de la propriété des données, difficile à régler au niveau national. En particulier, il arrive souvent, dans les zones rurales, que les dossiers de police ne soient pas tenus de la même manière; il se peut que les données soient la propriété des autorités locales et ne puissent être combinées ni comparées avec celles d'autres localités. Cela vaut également, dans certains pays, pour les données concernant les décès, ou tout au moins les décès de toxicomanes.

Pour toutes ces raisons et bien d'autres encore, il semble beaucoup plus facile d'organiser une estimation de la prévalence au niveau local, alors que les possibilités de le faire au niveau national demeurent très restreintes. Certes, sur le plan pratique, c'est au niveau local qu'il importe de disposer de chiffres concernant la consommation de drogue et les problèmes qu'elle pose. Il s'agit là d'éléments d'information d'une utilité cruciale. Il n'en reste pas moins qu'il est de toute évidence intéressant d'obtenir des informations chiffrées au niveau national, ne serait-ce que pour se faire une idée de la situation d'ensemble, indispensable pour déterminer une politique nationale. La question est alors de savoir si l'on peut se servir des estimations de la prévalence locale pour en extrapoler une

estimation au niveau national. Il est à noter que le même problème se pose lorsque, par exemple, on souhaite généraliser les résultats d'études menées dans une ville donnée pour en dériver des estimations concernant la région où cette ville est située.

Ce qu'il importe de garder à l'esprit, s'agissant d'estimations synthétiques et de toute autre méthode d'extrapolation, c'est que l'on se sert des chiffres de prévalence connus concernant certaines régions pour estimer la prévalence dans d'autres régions. Il faut, pour cela, que les régions cibles disposent non pas d'une estimation de la prévalence régionale, puisque c'est ce que l'on recherche, mais de certaines sources de données qui soient identiques (ou très proches) de celles concernant la région, pour lesquelles les chiffres de prévalence sont connus. Il s'agit donc d'utiliser des données comparables concernant les diverses localités pour calculer par extrapolation les chiffres de prévalence non connus.

Extrapolation à partir d'une seule estimation locale

Le premier exemple présenté est très simple, puisqu'il s'agit d'extrapoler les résultats d'une estimation d'une région à une autre, même si cette dernière est en fait le reste du pays.

Étude de cas n° 10. Extrapolation à partir de l'étude sur la Nouvelle-Galles du Sud *(Extrapolation simple à partir d'un seul point avec un indicateur d'usage de drogues)*

Une étude locale ayant produit une estimation acceptable de la prévalence de l'usage de drogues (37 000 consommateurs réguliers d'opiacés) pour la Nouvelle-Galles du Sud (voir les études de cas n^{os} 2 et 3), il a été jugé extrêmement intéressant d'extrapoler ce résultat à l'ensemble de l'Australie. On aurait pu tout simplement appliquer le taux de prévalence obtenu pour la Nouvelle-Galles du Sud à l'ensemble de la population australienne, c'est-à-dire partir du principe que ce taux était le même en Nouvelle-Galles du Sud et dans le reste du pays. Ainsi, étant donné qu'environ un tiers de la population australienne réside en Nouvelle-Galles du Sud, on en aurait conclu qu'il y avait, en Australie, 111 000 consommateurs réguliers d'opiacés (3 x 37 000).

Cette démarche pose cependant un problème d'importance: en partant du principe que la proportion d'héroïnomanes sur l'ensemble du territoire australien est identique à celle de la Nouvelle-Galles du Sud, elle ne tient pas compte des éventuelles différences régionales. Or on sait que le taux de consommation d'héroïne a toujours été bien plus élevé en Nouvelle-Galles du Sud (et particulièrement à Sydney) que dans d'autres régions. Les données nationales relatives aux traitements à la méthadone et aux décès par surdose d'héroïne indiquent en effet constamment qu'environ la moitié des personnes admises pour un traitement de ce type, ainsi que la moitié des personnes décédées à la suite d'une surdose d'héroïne sont originaires de Nouvelle-Galles du Sud.

Si l'on doit postuler que le nombre d'héroïnomanes est géographiquement proportionnel soit à l'ensemble de la population, soit au nombre de décès par surdose enregistrés, il est préférable de retenir la seconde solution. La première part du principe que la prévalence de la consommation est géographiquement constante, alors que la seconde pose que c'est le taux de décès par surdose qui l'est.

En conséquence, on a retenu un coefficient multiplicateur de 2,0, qui a permis d'estimer le nombre d'héroïnomanes dans l'ensemble de l'Australie à 74 000 (37 000 x 2). Ce chiffre a ensuite été utilisé pour calculer le taux de prévalence pour la tranche d'âge considérée comme étant pratiquement la seule concernée par le risque de décès par surdose, c'est-à-dire les 15-54 ans. Le taux ainsi obtenu est de 6,9 pour 1 000.

Source: Hall et al. (2000); et Lynskey et al. (1998).

Cet exemple, aussi simple qu'il soit, fait ressortir plusieurs points fondamentaux.

Premièrement, on remarquera que les chercheurs ont décidé de ne pas se contenter d'appliquer le taux de prévalence de l'abus de drogues en Nouvelle-Galles du Sud (appelé "point d'ancrage") au reste du pays (parfois appelé région "cible"). Quand peut-on considérer qu'un taux de prévalence régional vaut également pour l'ensemble du pays ou une autre région? De nombreux facteurs, tels que des structures sociales et une répartition démographique différentes, ainsi que le degré d'urbanisation et la proximité d'itinéraires de trafic de drogues, font que l'hypothèse selon laquelle une région donnée est représentative d'une autre région est trop simpliste pour être utile. Dans l'exemple ci-dessus, on a préféré utiliser, aux fins de l'extrapolation, un indicateur d'abus de drogues simple.

Les données requises peuvent être structurées, comme le montre le tableau 10. La population, le nombre de décès par surdose et le nombre de toxicomanes dans la région qui constitue le point d'ancrage sont connus, ce qui permet de calculer le taux de prévalence des décès par surdose et de l'abus de drogues. Pour la région cible, on dispose d'informations qui permettent de calculer l'indicateur, en l'occurrence le taux de décès par surdose, mais il reste à estimer le taux de prévalence de l'abus de drogues.

Sur le plan technique, on procède comme suit:

- a) On calcule l'indicateur pour le point d'ancrage et pour la cible;
- b) On calcule le taux de prévalence de l'abus de drogues pour le point d'ancrage;
- c) On extrapole à la cible la valeur obtenue pour le point d'ancrage, en partant du principe que la corrélation entre l'abus de drogues et l'indicateur est la même pour le point d'ancrage et la cible;

- d) On calcule, le cas échéant, le nombre effectif de toxicomanes plutôt que le taux de prévalence de l'abus de drogues pour la cible;
- e) On calcule, le cas échéant, le taux global de prévalence en additionnant le nombre estimé de toxicomanes dans chacune des régions, puis les chiffres de population, et en déterminant le ratio correspondant.

Les calculs ainsi effectués sont semblables à ceux auxquels on a procédé pour l'étude de cas n° 10; ils illustrent aussi les postulats retenus et permettent de généraliser la procédure de diverses manières (comme il est indiqué ci-après sous deuxièmement et troisièmement).

Tableau 10. Structure des données aux fins de l'estimation de la prévalence par extrapolation

Élément	Population (Chiffres)	Indicateur de l'abus de drogues (décès par surdose)		Prévalence de l'abus de drogues	
		(Chiffres)	(Taux)	(Taux)	(Chiffres)
Point d'ancrage					
Nouvelle-Galles du Sud	Connus	Connus	Calculé (a)	Calculé (c)	Connus
Cible					
Reste du pays	Connus	Connus	Calculé (b)	Non connus	Non connus
Ensemble	-	-	-	-	-

Deuxièmement, bien qu'une seule région soit retenue comme cible, il serait simple d'en ajouter d'autres en bas du tableau 10, à condition que des informations concernant l'indicateur soient disponibles pour chacune d'elles et d'établir des estimations distinctes pour chaque territoire de l'Australie.

Il est aussi possible d'ajouter plusieurs autres points d'ancrage dans la partie supérieure du tableau, à condition que des informations concernant l'indicateur et l'abus de drogues soient disponibles, ce qui permet de déterminer, sur une base plus large, la corrélation entre la prévalence de l'abus de drogues et la prévalence des décès par surdose. Cela soulève une nouvelle question: comment combiner les divers éléments d'information donnés par les points d'ancrage supplémentaires? En fait, il s'agit ici d'un problème classique d'analyse statistique par régression que n'importe quel progiciel permet de résoudre. En substance, on établit une corrélation moyenne entre les deux taux pour l'ensemble des points d'ancrage, en partant du principe que la corrélation est la même pour tous. S'agissant de l'étude de cas n° 10, les calculs sont considérablement simplifiés du fait qu'il n'y a qu'un point d'ancrage et qu'une cible.

Troisièmement, il est évident que l'indicateur n'a pas un rôle spécifique dans ces calculs: tout autre indicateur, quelle que soit la façon dont il est calculé, peut convenir, à condition, toutefois, qu'il concerne l'abus de drogues et que sa corrélation avec la consommation de drogues soit, selon toute vraisemblance, stable et uniforme d'un point d'ancrage et d'une cible à l'autre, et que l'on dispose d'éléments d'information qui permettent de le calculer uniformément et en parallèle pour les différents points d'ancrage et les différentes cibles.

Enfin, s'il existe, pour une région donnée, plusieurs estimations différentes du taux de prévalence résultant de plusieurs études locales, chacune de ces estimations peut simplement être ajoutée en tant que point d'ancrage distinct dans le tableau 10.

Avant de passer à d'autres aspects, il convient de noter que certaines questions d'ordre technique doivent être prises en considération, si l'on retient plusieurs points d'ancrage dans une analyse de régression. Il faut déterminer:

- a) Si l'on doit appliquer la régression de Poisson – qui est la méthode normale lorsqu'on analyse des taux de fréquence ou de prévalence – ou bien une simple analyse de régression linéaire; si l'on se prononce en faveur de cette dernière, il faudra décider soit d'utiliser directement les taux comme données, soit d'utiliser des données résultant de transformations logarithmiques;
- b) S'il faut établir des coefficients de pondération applicables aux différents points d'ancrage pour tenir compte du degré de fiabilité des estimations de la prévalence fournies pour l'extrapolation;
- c) Si les corrélations entre l'indicateur et les taux de prévalence de l'abus de drogues sont similaires pour tous les points d'ancrage, si, dans certains cas, elles sont trop différentes et devraient être exclues, et comment cela peut être vérifié. Cela vaut également pour la similitude entre les points d'ancrage et les cibles, car si pour certaines cibles l'écart au niveau de l'indicateur est trop grand, l'extrapolation risque de ne pas être fiable;
- d) Si le postulat selon lequel la corrélation entre les deux taux de prévalence est valable, si elle est suffisamment solide pour être utilisée aux fins d'une analyse de régression, et comment cela peut être vérifié.

Il est, en général, possible de trouver une réponse à ces questions, mais le mieux serait d'en discuter avec un statisticien confirmé.

Extrapolation avec plusieurs indicateurs

Une dimension supplémentaire peut être ajoutée à la méthode d'estimation par extrapolation exposée ci-dessus dans la section intitulée "Extrapolation à partir d'une seule estimation locale", en ayant recours à plusieurs indicateurs. On parlera alors d'estimation synthétique, d'estimation par indicateurs multiples ou, de façon plus technique, d'analyse par composantes principales.

Il s'agit d'utiliser une série d'indicateurs connus pour le point d'ancrage et la cible, afin d'affiner l'estimation de la prévalence de l'abus de drogues dans les régions cibles. L'étude de cas n° 11 ci-après se fonde sur les taux de saisie de drogues, de condamnation pour faits de drogues, de fréquentation de centres de traitement spécialisé, de tests de dépistage du sida effectués sur des usagers de drogues par injection et de décès liés à la drogue. Ces taux, qui concernent l'ensemble de la

population adulte, sont connus pour tous les points d'ancrage et toutes les cibles. Les spécialistes de l'analyse de régression penseront, tout d'abord, qu'il s'agit simplement d'estimer la prévalence au moyen d'une équation de régression avec cinq variables prédictives distinctes, une méthode qui est efficace lorsqu'on dispose d'un grand nombre de points d'ancrage. En général, toutefois, dans les cas où l'on a recours à ce type d'estimation indirecte, on ne dispose que de très peu de points d'ancrage, de sorte que les prédictions par régression multiple sont très instables et qu'il faut appliquer une procédure par régression modifiée.

Étant donné qu'en principe tous les indicateurs sont corrélés à un indice central d'abus de drogues, on peut les combiner en un indice unique applicable à tous les points de données et obtenir ainsi, toujours en principe, une valeur représentative de l'ensemble des indicateurs. Pour ce faire, on a recours à l'analyse par composantes principales (ACP), dont la raison d'être est justement de fournir un indice unique et fiable résumant l'ensemble des indicateurs. L'analyse de régression, dont le rôle est d'établir la corrélation avec les taux de prévalence, seulement à partir des données des quelques points d'ancrage, utilise cet indice unique comme prédicteur et est en mesure de fournir des estimations plus fiables pour les régions cibles.

Étude de cas n° 11. Estimation de la prévalence par extrapolation à l'aide d'indicateurs multiples (au Royaume-Uni) (Extrapolation à partir de points d'ancrage multiples)

Les données, dont on disposait pour cette étude, sont indiquées au tableau 11 ci-après. Une extrapolation des données sur l'abus de drogues connues, dans une certaine mesure, pour quatre points d'ancrage répartis dans l'ensemble du Royaume-Uni, permet d'estimer le nombre de toxicomanes pour sept cibles également situées en divers endroits du pays.

Objectif

L'Observatoire européen des drogues et des toxicomanies a établi des principes méthodologiques d'estimation de la prévalence au niveau national. On y analyse et commente les méthodes retenues pour estimer la prévalence au Royaume-Uni, ainsi que les résultats obtenus. Trois méthodes différentes ont été retenues, qui font respectivement appel : à des indicateurs multiples; à des multiplicateurs appliqués aux données relatives aux traitements, aux estimations relatives à l'infection par le VIH et aux statistiques de la mortalité; et aux études de la criminalité en Angleterre et au Pays de Galles, et en Écosse. La présente étude de cas ne rend compte que de l'application de la méthode des indicateurs multiples.

Définitions

La définition de l'expression "usage de drogues problématique", retenue aux fins de cette estimation, est celle donnée par le groupe de travail de l'Observatoire européen des drogues et des toxicomanies, à savoir l'usage de

drogues par injection intraveineuse ou l'usage durable ou régulier d'opiacés, de cocaïne et/ou d'amphétamines, la consommation d'ecstasy et de cannabis n'entrant pas en ligne de compte. Cette définition convient pour les études fondées sur des sources qui recueillent des données systématiques, mais qui ne précisent généralement pas le type d'usage. Il serait vain de se référer à la définition du syndrome de dépendance que donne la dixième révision de la Classification statistique internationale des maladies et des problèmes de santé connexes (CIM-10), à savoir un "ensemble de phénomènes comportementaux, cognitifs et physiologiques survenant à la suite d'une consommation répétée d'une substance".

Méthode: indicateur de l'usage de drogues problématique à plusieurs variables

Cette méthode permet d'estimer le nombre d'usagers de drogues à problème dans l'ensemble de la population, en combinant des informations sur la prévalence, qui ne sont connues que pour quelques régions (appelées échantillons de référence, ou encore "points d'ancrage"), et des "indicateurs" ou des "prédicteurs" de l'usage de drogues connus pour toutes les régions (Mariani *et al.*, 1994). Elle a été utilisée, pour la première fois, aux États-Unis (Woodward *et al.*, 1984) et a été exposée plus en détail dans d'autres ouvrages (Wickens, 1993).

Cette méthode repose essentiellement sur l'hypothèse que la corrélation entre la prévalence (variable dépendante) et les prédicteurs (variables indépendantes) dans l'échantillon de référence vaut également pour toutes les autres régions. On part également du principe qu'un facteur unique sous-tend les indicateurs relatifs à la drogue, et que le principal facteur, qui explique la plus grande partie des écarts entre les indicateurs, peut être extrait au moyen de l'analyse par composantes principales.

On trouvera ci-après, récapitulées, les opérations auxquelles on a procédé pour analyser les données relatives au Royaume-Uni.

Opération 1. On a obtenu une série d'indicateurs ou de prédicteurs de la prévalence de l'usage de drogues problématique disponibles pour toutes les régions. À cet effet, la Grande-Bretagne a été divisée en 11 régions: autorités sanitaires régionales pour l'Angleterre (1 à 8), pour le Pays de Galles (9) et pour l'Écosse, celle-ci comportant deux régions, Strathclyde (10) et le reste de l'Écosse (11). Ce découpage répond à des considérations pragmatiques, qui tiennent à la disponibilité des données.

Opération 2. À partir des données des recensements, on a déterminé le chiffre effectif de la population de chacune des 11 régions.

Opération 3. Les données suivantes, concernant l'abus de drogues au cours de l'année précédente, ont été recueillies pour toutes les régions: a) condamnations pour faits de drogue; b) saisies de drogues placées sous contrôle; c) nombre de toxicomanes qui, d'après les bases de données régionales sur l'abus de drogues, ont suivi un traitement; d) nombre de cas d'infection par le VIH en rapport avec l'usage de drogues par injection; et e) nombre de décès liés à la drogue. Par ailleurs, des estimations de la prévalence de l'abus de drogues étaient disponibles pour quatre régions: North Thames, West Midlands, Pays de Galles et Strathclyde, et ont été utilisées comme points d'ancrage pour l'extrapolation. Les données et les sources sont indiquées dans le tableau 11.

Tableau 11. Données relatives à l'abus de drogues concernant 11 régions du Royaume-Uni, y compris quatre points d'ancrage

Régions	Population	A	Indicateurs (voir notes)			E	Usagers de drogues recensés
			B	C	D		
Angleterre	47 055 204	83 533	92 095	51 850	788	2 371	
North Thames ^a	7 190 479	17 696	21 168	7 842	334	352	44 410
West Midlands ^b	5 150 246	7 125	5 398	4 322	26	193	13 130
Northern et Yorkshire	6 600 626	11 356	13 285	9 722	37	344	
Trent	4 606 495	6 451	7 010	3 580	67	207	
Anglia et Oxford	4 521 912	3 761	4 183	3 762	79	216	
South Thames	6 579 403	13 987	16 530	7 774	122	346	
South-west	6 131 705	10 600	12 717	5 890	60	311	
North-west	6 274 338	12 557	11 804	8 958	63	402	
Pays de Galles	2 835 073	6 110	5 870	2 282	14	139	
Pays de Galles ^c	2 835 073	6 110	5 870	2 282	14	139	8 357
Écosse	5 134 105	3 008	13 452	8 614	687	267	
Strathclyde ^d	2 283 671	943	7 989	4 331	97	127	18 000
Reste de l'Écosse	2 850 434	2 065	5 463	4 283	590	140	
Royaume-Uni	55 024 382	92 651	111 417	62 746	1 489	2 777	

Notes:

A Condamnations pour faits de drogue, 1996

B Saisies de drogues placées sous contrôle, 1996

C Nombre de toxicomanes qui, d'après les bases de données régionales sur l'abus de drogues, ont suivi un traitement en Angleterre et au Pays de Galles (octobre 1996-mars 1997), et en Écosse (avril 1995-mars 1996)

D Nombre de cas d'infection par le VIH en rapport avec l'usage de drogues par injection en Angleterre et au Pays de Galles (1996), et en Écosse (1995)

E Nombre de décès liés à la drogue au Royaume-Uni (1995)

^aEnquête sur la population: personnes ayant fait usage de drogues par injection au moins une fois au cours de leur vie (1991).

^bDossiers informatisés des médecins généralistes: extrapolation à partir des diagnostics d'abus de drogues ou de toxicomanie (1996).

^cÉtude par capture-recapture de toxicomanes lourds (1994).

^dProjection établie à l'aide d'un logiciel d'estimation de la prévalence de l'abus de drogues.

Opération 4. Chaque variable (A à E) et chaque chiffre relatif à l'abus de drogues concernant les points d'ancrage ont été convertis en taux de prévalence pour 100 000 habitants.

Opération 5. Pour chaque région, un indicateur d'abus de drogues a été calculé au moyen d'une analyse par composantes principales, à partir des cinq variables retenues. Un contrôle a été effectué afin de s'assurer que l'indice calculé représentait de manière satisfaisante la variation entre les régions telles qu'indiquées par les cinq variables (au-delà de 45 %, en l'occurrence) et pour déterminer s'il fallait recourir à un second indice. La première composante a fait apparaître une forte corrélation entre les indicateurs cruciaux: condamnations (0,049), saisies (0,842), traitements (0,789), VIH (0,147) et décès (0,864). Par contre, la composante 2 (non retenue aux fins de l'analyse) était plus fortement corrélée aux condamnations et au VIH.

Opération 6. Enfin, on a eu recours à une analyse de régression pour estimer, à partir des taux de prévalence connus des points d'ancrage et de l'indice calculé, le taux de prévalence de l'abus des drogues pour les sept cibles. Ce taux de prévalence a ensuite été converti en une estimation du nombre de personnes faisant abus de drogues dans chaque région.

Les données retenues pour la méthode à indicateurs multiples sont indiquées au tableau 12.

Tableau 12. Données et estimations de la prévalence au moyen d'indicateurs multiples concernant le Royaume Uni (1996)

Région	Population	Indice	Taux de prévalence calculé observé	Taux de prévalence calculé par extrapolation	Nombre d'usagers calculé par extrapolation
Angleterre					(225 358)
North Thames ^a	7 190 479	57	0,618 %	0,573 %	41 213
West Midlands ^b	5 150 246	19	0,255 %	0,187 %	9 643
Northern et Yorkshire	6 600 626	56		0,557 %	36 786
Trent	4 606 495	31		0,313 %	14 410
Anglia et Oxford	4 521 912	28		0,279 %	12 600
South Thames	6 579 403	58		0,581 %	38 234
South-west	6 131 705	47		0,473 %	28 997
North-west	6 274 338	69		0,693 %	43 475
Pays de Galles					(12 629)
Pays de Galles ^c	2 835 073	45	0,295 %	0,445 %	12 629
Écosse					(30 221)
Strathclyde ^d	2 283 671	75	0,788 %	0,749 %	17 110
Reste de l'Écosse	2 850 434	46		0,460 %	13 111
Royaume-Uni					(268 208)

Note: Les chiffres entre parenthèses correspondent à des totaux.

^aEnquête sur la population: personnes ayant fait usage de drogues par injection au moins une fois au cours de leur vie (1991).

^bDossiers informatisés des médecins généralistes: extrapolation à partir des diagnostics d'abus de drogues ou de toxicomanie (1996).

^cÉtude par capture-recapture de toxicomanes lourds (1994).

^dProjection établie à l'aide d'un logiciel d'estimation de la prévalence de l'abus de drogues.

À propos de l'exemple ci-dessus, il est deux points sur lesquels il convient d'appeler l'attention. Premièrement, un examen rapide des données relatives aux 11 régions montre que le nombre des condamnations présente une corrélation inverse avec tous les autres indicateurs, ce qui s'explique essentiellement par les taux de condamnation particulièrement bas pour les deux régions d'Écosse. Il s'agit là d'un scénario plausible, dans la mesure où, là où le taux de condamnation est élevé, on peut s'attendre à une baisse de l'abus observable, mais on peut aussi envisager d'autres types d'analyse. Une option consisterait simplement à écarter l'indicateur relatif aux condamnations et à appliquer une analyse par composantes principales, pour établir un indice unique à partir des quatre indicateurs restants qui présentent une corrélation positive, puis à procéder comme ci-dessus. Le tableau 12 serait alors fondé sur cet indice, établi à partir de quatre indicateurs seulement.

Les auteurs de l'étude ont toutefois décidé de retenir une méthode plus avancée et ont établi deux indices à partir de l'ensemble des cinq variables, au moyen d'une analyse par composantes principales qui suppose l'extraction de deux composantes obliques, l'une portant essentiellement sur les condamnations et l'autre sur les traitements. On peut, par la suite, retenir, selon que l'on juge bon, l'un ou l'autre des indices dérivés, voire les deux, aux fins de l'analyse de régression, mais, avec quatre points d'ancrage seulement, il est sans doute préférable de n'en utiliser qu'un seul, comme indiqué dans le tableau 12. Une fois de plus, il se trouve que, dans cet ensemble de données, les résultats globaux de l'extrapolation ne sont pas très différents, quelle que soit l'option choisie. On ne saurait toutefois trop insister sur le fait que cette complexification de la procédure ne devrait pas être entreprise sans consulter un statisticien expérimenté.

En second lieu, il est possible de comparer directement avec les données connues pour les points d'ancrage les estimations de prévalence obtenues par extrapolation,

au moyen du modèle de régression, pour ces mêmes points d'ancrage. Cela est utile pour contrôler si le modèle de régression permet de déterminer correctement la corrélation entre l'indice et l'abus; c'est ce que l'on appelle, en termes généraux, une analyse des résidus. Les auteurs soulignent que, dans l'étude ci-dessus, l'accord entre le modèle et le taux de prévalence observé pour le Pays de Galles n'est pas très bon et que, dans la mesure où ce taux représente un quart des données des points d'ancrage, il convient de faire preuve de prudence dans l'interprétation des résultats. L'analyse des résidus doit être systématique lorsqu'on procède par extrapolation.

Limitations

Cette méthode suppose que la prévalence non observée est en corrélation avec les indicateurs observés et que la corrélation entre ces indicateurs et les points d'ancrage est la même pour toutes les régions. Toutefois, d'autres facteurs influent sur les indicateurs et peuvent invalider ce postulat et les résultats qui en découlent. Ainsi:

- a) Le nombre d'usagers de drogues en traitement peut être fonction de la capacité des centres spécialisés; il risque aussi d'être affecté par la sous-déclaration, laquelle varie d'une région à une autre;
- b) La répression des faits de drogue et l'attention qui leur est accordée peuvent varier d'une région à une autre;
- c) La prévalence de l'infection par le VIH parmi les usagers de drogues par injection varie selon les régions;
- d) Les variables devraient toutes correspondre aux lieux de résidence, mais certaines d'entre elles (par exemple les statistiques de la police) correspondent à l'endroit de la déclaration.

La fiabilité et la validité des estimations établies pour les points d'ancrage sont d'une importance capitale. Les données relatives aux quatre points d'ancrage retenus pour l'étude ci-dessus ont été obtenues au moyen de diverses techniques, qui reposent sur divers postulats, qui n'ont pas pu être évalués. De plus, seules deux des quatre estimations avaient été établies pour la même année que les indicateurs (1996).

La variation régionale des estimations de l'usage problématique de drogues, résultant de l'application de la méthode à indicateurs multiples, constitue l'une des observations les plus importantes de l'étude ci-dessus. Ces estimations vont de 19 pour 1 000 habitants (West Midlands) à 75 pour 1 000 (Strathclyde) (voir le tableau 12). Pour interpréter ces écarts, il convient de garder à l'esprit le fait que ces estimations dépendent des points d'ancrage disponibles, qui peuvent représenter diverses formes d'usage problématique de drogues. Dans deux régions (West Midlands et Pays de Galles), l'écart entre la prévalence effectivement observée et les estimations obtenues à l'aide du modèle à indicateurs multiples est considérable.

Conclusions

L'une des méthodes retenues pour l'étude ci-dessus, la méthode des indicateurs multiples, a permis d'obtenir les estimations les plus élevées, et sans doute aussi les plus fiables, de l'usage problématique de drogues au Royaume-Uni. Cette méthode est économique dans la mesure où elle n'exige pas que l'on recueille

de nouvelles données, sauf si l'on a besoin de procéder à des études distinctes pour déterminer des points d'ancrage supplémentaires. L'une des difficultés a été d'obtenir des données auprès de différents organismes, publics et autres. Il a fallu procéder à des estimations au niveau sous-national, car il n'a pas été possible (compte tenu des délais fixés pour l'étude) d'obtenir des données pour des effectifs moins nombreux et plus représentatifs. Les autorités sanitaires régionales, en effet, s'occupent de populations hétérogènes et numériquement importantes, aussi bien en zone rurale qu'en zone urbaine, de sorte que les différences géographiques sont gommées. Il conviendrait, à l'avenir, d'envisager d'autres moyens permettant de faire apparaître les différences d'ordre géographique (par exemple, distinguer Londres et sa banlieue de la région de la Tamise, et obtenir des données distinctes pour d'autres grandes villes telles que Liverpool, Manchester et Birmingham).

Pour conclure, l'étude ci-dessus donne à penser qu'il est possible d'envisager une approche plus différenciée pour faire face à l'abus de drogues, mais qu'il faudra procéder à de nouveaux travaux pour mieux décomposer les caractéristiques démographiques.

Source: Projet de l'Observatoire européen des drogues et des toxicomanies intitulé "Study to obtain comparable national estimates of problem drug use prevalence in all European Union member States" (CT.97.EP.04); Frischer, M., *et al.*, "A comparison of different methods for estimating the prevalence of problematic drug abuse in Great Britain..." (2001) et Frank *et al.* (1978).

L'étude ci-dessus porte sur sept cibles et quatre points d'ancrage. Bien que des totaux partiels aient dû être établis pour l'Écosse et pour l'Angleterre, les données retenues pour l'analyse par composantes principales ne devraient concerner que les 11 régions prises séparément, à l'exclusion des totaux partiels.

Dans l'étude, telle qu'elle a été publiée, toutes les données ont été converties en taux pour 100 000 habitants pour chaque région, et ces taux ont été utilisés directement (le choix d'un autre dénominateur, par exemple 1 000 ou 100, n'aurait aucune incidence sur l'analyse, l'important étant d'utiliser un taux et non un chiffre absolu). On aurait pu aussi utiliser des transformations logarithmiques. Bien que, dans cette étude, cela n'affecte guère les résultats obtenus par extrapolation, il est souvent plus sûr, lorsqu'on analyse des taux, d'utiliser des transformations logarithmiques dès le départ, notamment si la fourchette des taux est très large. Cela garantit, à tout le moins, que les taux estimés d'usage problématique de drogues sont supérieurs ou égaux à zéro (c'est-à-dire que les estimations sont techniquement plausibles).

Postulats

Pour conclure, on notera que la méthode d'estimation synthétique – ou à plusieurs indicateurs – peut être appliquée de diverses façons légèrement différentes, selon les données disponibles et l'objet de l'estimation. Par exemple, puisque l'on n'utilise que les taux de prévalence, l'extrapolation peut être faite uniquement sur la base

de ces taux, sans que l'on connaisse les chiffres absolus correspondants. Toutefois, il est clair que l'on part des postulats de base suivants dans tous les cas:

- a) L'analyse doit se fonder sur les taux de prévalence des indicateurs et de l'usage problématique de drogues;
- b) On part du principe qu'il existe une corrélation stable entre la prévalence et l'ensemble des indicateurs, ce qui suppose aussi que les indicateurs soient définis et construits de la même façon pour tous les points de données;
- c) L'extrapolation donne des résultats moins valables et moins fiables si la corrélation présumée ne vaut pas pour tous les points de données;
- d) Les résultats de l'extrapolation sont fonction du modèle de régression ajusté (Poisson, linéaire, transformation logarithmique);
- e) L'extrapolation donne des résultats moins fiables et les estimations sont davantage sujettes à caution si l'on utilise plus d'un indice, à moins que le nombre des points d'ancrage ne soit élevé (on peut poser comme principe qu'il faut disposer d'au moins trois points d'ancrage pour chaque indice utilisé);
- f) Une analyse par composantes principales requiert un nombre de points de données suffisant (on peut poser comme principe qu'il en faut au moins deux à trois fois plus que d'indicateurs).

Exactitude des estimations de la prévalence

Les méthodes d'estimation indirecte peuvent donner des résultats très inexacts. Il importe donc d'en évaluer l'exactitude d'une manière ou d'une autre. Les statisticiens font normalement la distinction entre la fiabilité et le biais. Une méthode est fiable lorsqu'elle donne toujours le même résultat chaque fois qu'elle est appliquée de la même façon à la même population. Les résultats obtenus sont fiables lorsque tout écart mineur qu'ils peuvent présenter s'explique par le hasard (c'est-à-dire par le fait que les différents échantillons, bien qu'étant tous représentatifs de la population, sont composés d'individus différents). On parle de biais lorsqu'on observe systématiquement un écart entre le résultat obtenu et la valeur vraie, quel que soit le degré de fiabilité de la méthode. Le biais résulte du fait que, compte tenu des conditions propres à l'étude en cours, les postulats qui sous-tendent la méthode retenue ne peuvent s'appliquer. La "robustesse" d'une méthode est la mesure de sa capacité à fonctionner dans des conditions pour lesquelles elle n'a pas été explicitement conçue, c'est-à-dire du degré de biaisage qu'elle génère.

La fiabilité est, d'une certaine manière, intrinsèquement fonction de la méthode d'estimation, mais elle dépend aussi de la taille de l'étude; en effet, plus celle-ci est petite, moins elle tend à être fiable. En ce qui concerne l'estimation indirecte de la prévalence de l'abus de drogues, on peut dire, d'une manière générale, que sa fiabilité est fonction du rapport entre le nombre estimé de toxicomanes et le nombre d'enregistrements relatifs à ces derniers que l'on retient pour l'étude. Certes, la fiabilité varie aussi d'une méthode à l'autre, et d'autres facteurs encore entrent en jeu.

Ainsi, les études fondées sur un multiplicateur de mortalité égal ou supérieur à 100 sont en général moins fiables que celles utilisant des références plus communes et des multiplicateurs moins élevés. De même, plus les données sont nombreuses, et plus il y a de recoupements entre les sources de données dans une étude par capture-recapture, plus l'estimation est robuste, à condition, bien sûr, que l'étude ne soit pas "biaisée".

Pour ce qui est du biais, toutefois, les facteurs qui entrent en ligne de compte sont d'un autre ordre. La taille de l'étude ne joue guère lorsqu'il s'agit d'éliminer la distorsion engendrée par la méthode retenue. La plupart des méthodes les plus couramment utilisées donnent des résultats non biaisés lorsqu'elles sont correctement appliquées. Le biais intervient lorsqu'on n'applique pas de bonnes pratiques de recherche et lorsqu'on ne se rend pas compte (par un contrôle ou une observation) que les postulats qui sous-tendent l'analyse statistique n'ont pas été respectés. Les bonnes pratiques de recherche sont exposées dans les deux sections suivantes. La question du non-respect des postulats est traitée ci-après, dans la section intitulée "Robustesse des estimations de la prévalence". Le recours à des procédures statistiques classiques pour déterminer les intervalles de confiance relatifs aux estimations de la prévalence est traité ci-dessous.

Calcul des intervalles de confiance par des méthodes statistiques classiques

L'intervalle de confiance est la fourchette dans laquelle on est en droit de penser que se situe la valeur réelle, en l'occurrence la prévalence. Cette façon fort utile de présenter la fiabilité (plus l'intervalle est grand, moins l'estimation est précise) ne fait que donner une indication de la variabilité d'échantillonnage, c'est-à-dire des écarts aléatoires entre l'estimation et la valeur réelle. Il existe des formules standard qui permettent de déterminer des bandes de confiance autour des estimations ponctuelles de la prévalence, calculées au moyen de la plupart des méthodes statistiques. Avec un logiciel, cette opération est simple. L'intervalle de confiance peut être exprimé par des chiffres absolus (par exemple, entre 12 000 et 27 000 toxicomanes) ou par une marge (par exemple, 18 000 toxicomanes, avec une marge de 1,5). Il est à noter que la validité de l'intervalle de confiance calculé est fonction de la validité des hypothèses retenues aux fins de l'analyse et qu'elle ne rend compte que de la variabilité d'échantillonnage et non du biais. C'est pourquoi, compte tenu du fait que les méthodes d'estimation de la prévalence peuvent être sujettes à de nombreux biais, ces intervalles ne sont pas aussi décisifs que le recours à des indicateurs multiples ou que les comparaisons entre estimations.

Calcul des intervalles de confiance dans le cas des méthodes multiplicateur/référence

Dans le cas d'une estimation de prévalence par la méthode multiplicateur/référence, l'intervalle de confiance est facile à calculer lorsque cette estimation est considérée comme une régression (log-linéaire) de Poisson. Si l'on a recours à un progiciel pour

établir l'estimation, l'intervalle de confiance sera calculé automatiquement. D'une manière générale, lorsqu'on estime le nombre de toxicomanes dans une région ou un pays donné, l'intervalle de confiance (exprimé par une marge) fait apparaître que plus le multiplicateur est élevé, moins l'estimation est fiable et moins la valeur du multiplicateur lui-même est fiable. Pour ce qui est du degré de fiabilité du multiplicateur, on peut dire, en gros, qu'il est fonction de la taille de l'échantillon à partir duquel il est calculé: plus l'échantillon est grand, plus le multiplicateur est fiable.

Il arrive souvent que l'on ne dispose pas des informations requises pour calculer l'intervalle de confiance, et peu d'analyses effectuées par la méthode multiplicateur/référence en comportent. Cela n'est toutefois pas un problème majeur; en effet, il existe des incertitudes plus dommageables que celles qui tiennent à la variabilité d'échantillonnage, comme on le verra ci-après dans la section intitulée "Robustesse des estimations de la prévalence".

Calcul des intervalles de confiance dans le cas des méthodes par capture-recapture

Les intervalles de confiance des estimations par capture-recapture peuvent être calculés très facilement, quelle que soit la complexité du modèle, par le progiciel utilisé pour ajuster ce modèle. Par ailleurs, Bishop *et al.* donnent des équations qui permettent de les calculer manuellement. On peut aussi les calculer en se fondant sur l'intervalle de probabilité ou la qualité de l'ajustement, les limites inférieures et supérieures de l'intervalle de confiance étant déterminées par approximations successives en obtenant pour la population non observée les valeurs qui s'approchent le plus d'un écart de 3,84 (95 %) de G^2 , qui est une mesure de l'ajustement du modèle. Pour plus de précisions sur cette méthode, voir Hook et Regal (1995). L'incertitude d'une estimation par capture-recapture est d'autant plus grande que les échantillons prélevés aux points de capture sont petits, que le pourcentage et le nombre de correspondances entre les sources sont faibles et que le modèle requis pour décrire les données est complexe.

Ici aussi, il importe de noter que l'intervalle de confiance ne rend absolument pas compte de la viabilité du modèle; il n'est donc pas le seul élément à retenir si l'on veut déterminer le degré de certitude d'une estimation.

Les intervalles de confiance peuvent normalement être calculés de façon analytique, comme indiqué dans la présente section, mais on a aussi très souvent recours, pour ce faire, à des techniques de simulation de données, lesquelles requièrent un moins grand nombre de postulats. Bien que l'on puisse, pour déterminer le degré de fiabilité, appliquer des techniques de Monte Carlo et d'autres méthodes axées sur les données, on n'y recourt que rarement dans les faits. Les techniques polyvalentes axées sur les données, telles que les méthodes *bootstrap* ou *jack knife*, donnent des résultats plausibles mais pas toujours satisfaisants. En général, elles exigent tant d'efforts que l'analyse proprement dite devient un aspect mineur du travail. C'est pourquoi on ne les applique que rarement, et uniquement avec l'aide d'un programmeur expérimenté.

Robustesse des estimations de la prévalence

Les méthodes standard de calcul des intervalles de confiance applicables dépendent de la validité des postulats de la méthode d'estimation indirecte utilisée, quelle qu'elle soit. Toutefois, ce n'est pas la variabilité d'échantillonnage mais la robustesse qui est véritablement importante et le non-respect des postulats qui pose le plus gros problème.

Il est bien plus probable que les erreurs dans les estimations obtenues soient principalement dues à des situations non conformes aux postulats requis et non à une simple variabilité d'échantillonnage tel qu'on l'entend en statistique. Cette non-conformité a tendance à générer des biais qui peuvent parfois être importants, et le manque d'exactitude qui en résulte ne peut être corrigé par les méthodes standard de calcul des intervalles de confiance.

En général, dans le cadre des méthodes multiplicateur/référence, on recourt rarement aux méthodes statistiques classiques pour calculer les intervalles de confiance. On établit plutôt une estimation haute et une estimation basse de la prévalence en faisant varier le multiplicateur ou la référence, sur la base d'informations relatives à leur incertitude. Par exemple, Hartnol *et al.* (1985) ont commencé par utiliser un multiplicateur de valeur 100 – correspondant à un taux de mortalité présumé de 1 % –, puis ont refait les calculs en appliquant un multiplicateur de valeur 200 – correspondant à un taux de mortalité présumé de 2 %. La fourchette d'estimations ainsi obtenue n'a rien à voir avec le concept d'intervalle de confiance, mais elle rend compte de façon tout à fait informelle de l'incertitude inhérente à l'information utilisée.

Compte tenu du nombre de postulats que requiert la méthode par capture-recapture et du risque implicite qu'ils ne soient pas respectés, il est toujours possible que le modèle utilisé pour l'analyse soit erroné, quoi que le chercheur fasse, et que l'estimation ne corresponde donc pas à la réalité. Dans les sections pertinentes, on a attiré l'attention sur les postulats qui risquaient de poser problème, par exemple:

- a) Il est dangereux de n'utiliser que deux sources pour une estimation par capture-recapture;
- b) Il faut supposer qu'au moins une interaction du premier ordre est nulle.

On trouve également dans la littérature statistique une méthode standard permettant de vérifier si les modèles log-linéaires de Poisson décrivent de façon adéquate les données (test de Pregibon).

Peu de travaux portent sur le degré exact de sensibilité de ces diverses méthodes d'estimation indirecte au non-respect des postulats, mais l'on s'accorde à penser que les estimations peuvent être complètement fausses dans le pire des cas. C'est pourquoi il est courant – selon une approche qui a l'avantage de la simplicité – que l'on vérifie la concordance et la convergence des estimations faites au moyen de différentes méthodes indirectes. En général, le seul critère utilisé pour juger l'exactitude de

ces méthodes est la mesure dans laquelle les estimations qu'elles donnent dans n'importe quelle situation donnée sont convergentes. Là encore, comme il n'existe pas de critère formel pour juger la convergence et la concordance des diverses procédures d'estimation, on doit se contenter d'émettre un jugement informel.

Étant donné que la concordance est sans doute le meilleur indicateur de la fiabilité d'une estimation, on peut faire les recommandations suivantes:

- a) Ayez recours à des méthodes à plusieurs indicateurs:
 - i) Appliquez, si possible, la méthode de capture-recapture et la méthode multiplicateur-référence;
 - ii) Appliquez des méthodes faisant appel à plusieurs multiplicateurs, le mieux étant que ces multiplicateurs soient dérivés de plus d'une source (au cas où l'échantillon retenu pour établir le multiplicateur serait biaisé);
- b) Utilisez différents modèles pour l'analyse d'un ensemble de données, afin d'obtenir une fourchette d'estimations possibles: dans le cas des méthodes de capture-recapture, comparez différents modèles log-linéaires sous-optimaux;
- c) Vérifiez la plausibilité et la cohérence des estimations concernant différents comportements ou différents sous-groupes de population:
 - i) La prévalence de l'usage de drogues par injection et la prévalence de l'usage d'héroïne devraient présenter une certaine corrélation, sauf si l'on a de bonnes raisons de penser que tel n'est pas le cas (par exemple lorsque l'héroïne est principalement consommée sous la forme de produit à fumer);
 - ii) Le modèle retenu pour une estimation par capture-recapture devrait être crédible pour ce qui est des interactions envisagées entre les sources de données.

Principes directeurs pour l'établissement d'estimations fondées sur les résultats des activités de recherche

Chapitre III

Principes généraux

Les données utilisées dans les méthodes d'estimation indirecte décrites dans le présent manuel sont soit obtenues auprès de sources existantes, soit recueillies spécialement pour les besoins des projets de recherche entrepris. Les tâches ne sont pas les mêmes dans les deux cas mais il existe, à un niveau très général, un recoupement important, et il est souhaitable que toute activité de recherche offre des possibilités de coopération et de renforcement des capacités de recherche, de façon à préparer les activités futures et à créer les conditions de base qui permettront de les mener à bien plus facilement. Les sections ci-après insistent donc sur l'importance de consultations avec d'autres chercheurs et d'autres institutions, qui exercent leurs activités dans des domaines en rapport avec les drogues, en partie pour répondre aux priorités immédiates de ces activités et en partie pour constituer des réseaux en vue de recherches futures. Ainsi, en saisissant les possibilités de coopération, une étude devrait favoriser la constitution d'une capacité durable de surveillance et de recherche dans le domaine de l'abus des drogues, grâce à la constitution d'un corpus de connaissance et d'informations et à la mise en place de mécanismes de coopération entre groupes par ailleurs isolés. C'est dans cette optique que les principes généraux de "bonne pratique" ci-dessous ont été formulés.

Conditions minimales à remplir pour mener à bien un projet d'études de la prévalence

Une recherche ne peut être de qualité si elle est engagée à la légère. Elle suppose une diversité considérable de compétences et un travail en équipe. La liste des compétences, des moyens et des connaissances générales qui figure ci-après n'est bien entendu qu'indicative, mais sa diversité parle d'elle-même.

Nombre des éléments qui figurent dans la liste ci-dessous sont considérés comme allant de soi par n'importe quel chercheur, mais peut-être que l'intérêt de certains n'apparaît qu'a posteriori.

Liste récapitulative des conditions minimales à remplir en matière d'expertise et de ressources

1. Contacts avec l'ensemble de la communauté scientifique (épidémiologie) internationale;
2. Expertise en matière de planification et de gestion de projets;
3. Groupe consultatif composé de chercheurs éminents possédant une expérience de la question;
4. Possibilité de consulter des statisticiens dès les phases de conception;
5. Compétences en matière de rédaction de questionnaires;
6. Personnel de terrain possédant une expérience des études dans le domaine des drogues ou une expérience similaire; formation du personnel de terrain;
7. Possibilité d'utiliser des ordinateurs et compétences en matière de saisie des données et de création de fichiers;
8. Conseils de statisticiens lors des analyses ultérieures;
9. Constitution avec d'autres chercheurs intéressés de groupes de discussion, sur Internet ou par d'autres moyens.

Le premier point, à savoir l'établissement de contacts au niveau international, n'est pas simplement une évidence: les organismes de recherche sur les drogues, partout dans le monde, ont acquis une expertise, une expérience et des connaissances importantes, et il est essentiel – aussi bien pour définir les objectifs et les caractéristiques de base du projet que pour en déterminer la faisabilité – d'établir avec eux des contacts dès le début. Entreprendre un nouveau projet qui ne peut être mené à bien est tout aussi inutile qu'entreprendre un projet similaire, sous une forme ou une autre, à un projet antérieur. De même, concevoir un projet de telle façon que celui-ci soit difficile à réaliser, alors qu'un projet plus simple pourrait permettre d'atteindre pratiquement les mêmes objectifs beaucoup plus efficacement, revient à se créer des problèmes inutiles qui risquent d'en contrarier le bon déroulement. Les facteurs, le cadre général et les spécificités d'un projet donné d'évaluation de la prévalence sont beaucoup trop variés pour que l'on puisse, dans un manuel comme celui-ci, formuler des recommandations de portée générale, et on ne saurait sous-estimer combien il importe de prendre conseil auprès de ceux qui ont déjà acquis une expérience en la matière. Les nombreux exemples qui figurent ci-après dans la section intitulée "Principes techniques" devraient fournir quelques indications quant à la façon de tenir compte de la situation locale, à laquelle le projet devra être adapté.

Une compétence en planification et en gestion est évidemment indispensable, mais elle fait partie de ces compétences qui sont trop souvent considérées comme allant de soi. Faire des recherches n'est pas aisé. Évaluer la prévalence de l'abus des drogues

est difficile. Une bonne planification est un élément nécessaire, mais non suffisant: il faut également être en mesure de parer aux imprévus et de faire preuve de suffisamment de souplesse sans que cela ne porte atteinte à la validité de la méthode employée. C'est un compromis difficile à réaliser et qui implique nécessairement une aide extérieure. Il est, par conséquent, très fortement recommandé de faire appel à un groupe d'experts qui aidera à définir les objectifs et les méthodes du projet, ainsi qu'à en évaluer le déroulement. Ce groupe d'experts devrait pouvoir participer à la prise des décisions nécessaires pour faire face aux difficultés d'exécution et remédier aux erreurs de conception qui ne deviennent apparentes qu'une fois le projet engagé. Il serait inconsidéré de croire que la planification, aussi élaborée soit-elle, aura tout prévu, et tout projet devra donc être conçu de telle façon que l'on puisse y apporter les ajustements jugés nécessaires pour en assurer le succès.

Il arrive trop fréquemment que l'on ne fasse appel à un statisticien qu'une fois qu'un ensemble de données a été recueilli et que l'on a besoin de conseils quant à la méthode d'analyse à appliquer. Or ce type de conseil est encore plus important lors de la conception du projet que lors des phases ultérieures. Si le projet est conçu de telle façon qu'il ne permette pas d'effectuer une analyse statistique satisfaisante, il sera impossible d'y apporter, par la suite, les modifications qui s'imposent, et il existe en général de nombreuses insuffisances – mauvais échantillonnages, mauvaise formulation du questionnaire, absence d'étude de faisabilité, taille inappropriée du projet, etc. – qui peuvent rendre une étude inutilisable. Ce type de problème est souvent trop technique pour être aisément décelé par un non-statisticien, tandis qu'un statisticien expérimenté sera toujours en mesure de fournir des conseils quant à la façon de les résoudre. Il faudrait également se préoccuper, lors de la phase de conception, de la nature des formulaires destinés à enregistrer ou à recueillir les données – feuilles de codage – et, le cas échéant, du questionnaire. Dans le cas d'études qui impliquent des entretiens avec des toxicomanes, l'intervention d'un statisticien pourrait se révéler extrêmement utile pour la rédaction du questionnaire.

Une fois les diverses activités de planification terminées, la responsabilité d'effectuer les recherches incombe au personnel et à l'organisation de terrain. En dernière analyse, on peut dire que le travail sur le terrain constitue l'élément clef de la collecte des données, que celles-ci soient recueillies à partir des dossiers conservés par divers organismes ou qu'elles proviennent d'entretiens personnels. À cet effet, il est vivement souhaitable de faire appel à du personnel de terrain expérimenté, en particulier si le projet prévoit des entretiens avec des toxicomanes. Il est essentiel, pour recueillir des données utiles et fiables, que les toxicomanes interrogés se sentent à l'aise, que la personne qui les interroge ne porte pas sur eux de jugement moral et que la confidentialité des réponses soit respectée. Les personnes dont l'activité est en rapport avec les drogues sont utiles à cet égard, et il faut également envisager la possibilité de faire appel à d'anciens toxicomanes. Par contre, les personnes qui, d'une façon ou d'une autre, sont associées aux services chargés de faire respecter la loi ne feront pas, à l'évidence, de bons interviewers. Que les données soient recueillies auprès de sources officielles, de dossiers semi-officiels ou lors d'entretiens, la coopération doit être le maître mot pour le personnel de terrain.

Que le personnel de terrain soit expérimenté ou non, il devra nécessairement être formé aux méthodes, questions et procédures de collecte qui seront utilisées dans le cadre de l'étude. Il ne faut pas partir du principe que ces procédures sont évidentes ou ne demandent aucune explication. La formation a, entre autres, pour objectif de faire en sorte que les critères, les définitions et les interprétations soient appliqués de la même façon par l'ensemble du personnel de terrain. Une certaine forme de supervision et d'organisation est nécessaire pour coordonner l'action du personnel. Elle permet de s'assurer à la fois de la qualité des données collectées et du respect des procédures, ainsi que de contrôler le flux d'informations – fiches de renseignement ou questionnaires – à destination du siège. Il est surprenant de constater combien de questionnaires ou de fiches sont perdus ou inutilisables. En dernière analyse, la validité des données utilisées pour l'étude dépend de la qualité du personnel et de l'organisation du travail sur le terrain.

Aucune étude moderne de prévalence ne peut être réalisée sans ordinateur et sans savoir comment utiliser les logiciels et construire des fichiers. La saisie des données peut prendre du temps et, si l'étude est de grande ampleur, il faudrait envisager de l'effectuer au fur et à mesure. Si les principaux chercheurs n'ont pas les compétences informatiques nécessaires, il est indispensable de s'attacher la coopération d'une personne qui les possède. Un statisticien peut toujours fournir des conseils à ce sujet également. On espère que le présent manuel constituera une source d'information utile en matière d'analyse statistique pour les chercheurs qui disposent déjà d'une certaine expertise dans ce domaine, et le texte précise les aspects plus complexes de l'analyse qui nécessiteront l'intervention de personnes plus qualifiées. Ceux qui n'ont pas d'expérience à cet égard devront absolument chercher à obtenir les conseils d'un statisticien lors de l'application de ces méthodes indirectes d'estimation.

Enfin, une fois que l'essentiel des travaux a été réalisé et qu'il s'agit principalement d'en rédiger les conclusions, il importe de ne pas le faire de manière isolée. Le travail d'équipe est utile à ce stade, car il permet non seulement d'obtenir un appui statistique, mais aussi de demander conseil à d'autres personnes de l'équipe, que ce soit de manière informelle à l'occasion de contacts avec des confrères ou dans le cadre de groupes de discussion organisés, y compris sur Internet.

Considérations éthiques dans différents contextes sociaux

Il importe de toujours agir avec prudence lorsqu'une enquête, ou tout autre type d'étude, suppose d'obtenir des informations personnelles des répondants. Les entretiens avec ces derniers présentent de nombreux risques. Il faut tenir compte, entre autres, des coutumes sociales de chaque pays pour déterminer les modes de comportement éthiques, par opposition aux modes de comportement non éthiques, les bonnes manières et les procédures socialement acceptables. On trouvera, dans les paragraphes ci-dessous, les principaux éléments dont il faudra tenir compte dans de nombreux pays et dans de nombreux contextes sociaux.

Censure sociale

Ce type de recherche porte sur un comportement réprouvé par la société et doit, par conséquent, toujours être mené avec une grande précaution. Les personnes interrogées doivent toujours être certaines que leurs réponses seront confidentielles et que leur anonymat sera respecté, aussi bien dans le traitement des réponses que dans la façon dont elles sont contactées. Le chercheur doit toujours être en mesure de garantir que les réponses resteront confidentielles et anonymes, non seulement parce que cela contribuera à la validité des données, mais également pour la simple raison qu'il intervient dans la vie de la personne interrogée. Les assurances à donner seront différentes et spécifiques d'un groupe à l'autre. Par exemple, dans certains pays, ce qui touche aux jeunes et aux femmes est extrêmement sensible et nécessite une attention particulière.

Cadre des entretiens

La façon dont les répondants sont contactés est importante: elle devrait être discrète et aussi privée que possible. Il en est de même en ce qui concerne le cadre dans lequel sera mené l'entretien: il est généralement recommandé que celui-ci se déroule dans une pièce ou un endroit privé et isolé. Si cela n'est pas possible, le chercheur doit alors faire particulièrement attention aux questions familiales, non seulement en ce qui concerne le rôle des femmes dans la famille en général – qui, dans de nombreux pays, doit être traité avec une attention particulière –, mais également en ce qui concerne les rapports entre mari, femme et enfants.

Un autre facteur important dont il faut tenir compte tient au statut social éventuel de la personne interrogée dans le pays. Les personnes chargées de mener ces entretiens doivent être conscientes de toute différence éventuelle de statut social entre elles-mêmes et les répondants, notamment lorsque ces derniers occupent une position particulièrement importante dans la société. Il n'existe pas de règles générales applicables aux très diverses situations sociales que l'on peut rencontrer d'un pays à l'autre. Tout ce qu'il est possible de faire est d'appeler l'attention sur cette question de façon que le chercheur adopte des méthodes appropriées.

Considérations de sécurité

De par sa nature même, l'étude d'un comportement illégal s'accompagne de certaines difficultés, au nombre desquelles figure au premier chef le type de relation que le chercheur entretient avec la police et la justice, ainsi qu'avec tout autre organisme public dont il souhaite consulter les dossiers. Il est essentiel que le chercheur soit en mesure d'assurer l'anonymat et la confidentialité de toutes les informations fournies. En outre, cette confidentialité suppose que ses données et ses dossiers soient séparés des documents et écrits administratifs plus accessibles. Si une certaine forme d'identification des personnes interrogées est maintenue, celle-ci devrait être aussi peu détaillée que possible et être conservée en lieu sûr. Il arrive fréquemment que les dossiers officiels contenus dans un système soient "contaminés" par des indications contenues dans un autre système public, qui permettent d'identifier les personnes

concernées. Il importe de limiter, non seulement pendant toute la durée du projet, mais aussi par la suite, l'accès aux données recueillies et aux autres dossiers, en particulier aux réponses fournies par les personnes interrogées. Afin d'assurer l'anonymat et la confidentialité après le projet, il faudra généralement détruire toutes les informations administratives et celles permettant d'identifier des personnes, et ne conserver au maximum que des données de recherche chiffrées et des dossiers entièrement anonymes.

Principes techniques

La présente section expose brièvement un certain nombre de principes généraux concernant les diverses étapes d'une étude, depuis la définition de la cible et la planification jusqu'au choix du mode et du lieu de publication finale des résultats, en passant par la sélection de la méthode appropriée.

Définitions de la cible

Cadre d'exécution

Quelle que soit l'étude réalisée, un certain nombre d'objectifs précis doivent être fixés dès le début si l'on veut éviter toute "dérive". En particulier, il faut définir clairement le cadre à l'intérieur duquel le projet sera exécuté. S'agissant de l'aspect financier, la plupart des organismes de financement voudront avoir la preuve que le projet a bien atteint l'objectif fixé. Du point de vue de la planification, il est indispensable de freiner la tendance naturelle à fixer de plus en plus d'objectifs au fur et à mesure que de nouvelles questions intéressantes se font jour: un projet peut facilement devenir trop ambitieux, aussi bien du point de vue financier que du point de vue pratique. La tendance à renoncer aux aspects les plus difficiles du projet est également naturelle et doit être combattue. Si le projet comporte plusieurs éléments – qu'il s'agisse de plusieurs phases de collecte de données ou de plusieurs méthodes d'estimations –, il faut veiller à les tenir aussi séparés que possible et à ce qu'ils se déroulent dans un cadre unifié.

Population cible et population de référence

Dans les études d'estimation de la prévalence, la population cible, à savoir la population de toxicomanes, définie d'une façon ou d'une autre, correspond au numérateur du taux de prévalence, tandis que la population de référence, au sein de laquelle se trouve le groupe cible, correspond au dénominateur. Les critères appliqués aux aspects démographiques, géographiques et temporels doivent être les mêmes pour la population cible et la population de référence.

Il est clair qu'il faut, avant tout, délimiter géographiquement la zone dans laquelle se trouvent les populations étudiées, par exemple un pays ou une région bien définie

d'un pays s'il existe des incertitudes quant aux frontières ou d'autres raisons pour se limiter à une région donnée. Dans certains cas, la prévalence de l'abus des drogues peut être impossible à estimer, soit parce que la région considérée est difficile d'accès, soit parce que l'on ne dispose que de peu d'informations, soit encore parce que cette prévalence y est tellement faible qu'elle en devient difficile à mesurer. Il se peut également que l'on ne cherche à déterminer la prévalence que pour une région limitée, voire une ville. En tout état de cause, il est indispensable d'adopter une définition précise et claire.

Il faudra sans doute aussi prendre en considération la nationalité et le lieu de résidence. Par exemple, les ressortissants étrangers en visite dans le pays ou qui y résident doivent-ils être inclus dans l'étude? Dans certains cas, la réponse à cette question n'aura guère d'incidences sur le résultat obtenu, mais, dans d'autres cas, elle pourrait en avoir. Si l'étude est limitée à une zone géographique restreinte, par exemple une ville, la question de la résidence peut alors prendre une très grande importance. La décision en la matière doit être fonction de l'objectif principal assigné à l'étude, et les considérations d'ordre méthodologique détermineront les critères à appliquer à la population cible comme à la population de référence.

Il faut également fixer avec précision les caractéristiques démographiques des populations étudiées. Par exemple, l'étude ne doit-elle porter que sur les personnes âgées de 15 à 65 ans, ou sur l'ensemble de la population? Doit-elle porter aussi bien sur les hommes que sur les femmes? Sur les personnes ayant un travail que sur les chômeurs? Dans la plupart des cas, le choix de ces critères démographiques et, le cas échéant, d'autres caractéristiques sociales est déterminé par des raisons pratiques ainsi que par la disponibilité et l'accessibilité des données nécessaires, et pas seulement par les préférences des chercheurs.

La période couverte par l'étude est généralement un peu plus difficile à définir. Elle sera fonction à la fois de la définition des comportements cibles (voir les paragraphes ci-dessous consacrés aux principales mesures en matière de comportement) et de la définition des comportements retenus pour définir un "toxicomane" au sein de la population cible. Par exemple, la période généralement recommandée fait partie intégrante de la définition "a consommé des opiacés au cours des 12 derniers mois".

Principales mesures en matière de comportement

Afin d'identifier la population cible, il est également nécessaire de définir les types de drogues sur lesquels portera l'étude. L'introduction de ce manuel présente les catégories de drogues qui figurent dans la classification de la toxicomanie du GAP. Il est bien entendu recommandé d'estimer, dans le cadre de la même étude, les taux de prévalence pour plusieurs, voire la totalité, de ces catégories. Cela n'est toutefois pas toujours possible, en particulier si certaines sources de données ne concernent, par exemple, que les opiacés et si la probabilité de trouver des utilisateurs d'autres drogues est limitée. Il se peut que le critère essentiel soit la consommation par

injection (de n'importe quelle drogue et pas seulement d'héroïne) ou bien encore qu'il faille définir les concepts de "toxicomanie" ou d'"usage problématique de drogues". À cet égard, il convient de noter qu'il est généralement très difficile d'appliquer sur le terrain, ou lors de l'administration de questionnaires, les définitions de ces concepts utilisées dans la Classification internationale des maladies ou le Manuel diagnostique et statistique des troubles mentaux (DSM). Des définitions proches mais toujours précises sont souvent ce que l'on peut faire de mieux.

L'introduction définit les périodes de prévalence pouvant être utilisées. Afin d'assurer la compatibilité avec le questionnaire destiné aux rapports annuels, le choix d'une période d'un an devait être la règle, mais il est également possible d'obtenir simultanément des informations sur l'utilisation des drogues au cours d'autres périodes. Si l'étude prévoit des entretiens, la période retenue est généralement la période de 12 mois précédant l'entretien. Si les données proviennent de documents déjà existants, le chercheur n'aura peut-être pas d'autre choix que d'utiliser l'année civile écoulée. Si des rapprochements sont effectués entre des données provenant de diverses sources, il est de loin préférable d'appliquer la même période pour chacune des sources. Si cela n'est pas possible, il faudra procéder à des pondérations spécifiques parfois difficiles [Simeone *et al.* (1997); et Fischer *et al.*, "A comparison of different methods for estimating the prevalence of problematic drug abuse in Great Britain..." (1999)].

Si, dans le cadre de l'étude, il est nécessaire de définir ce qu'est "un consommateur de drogue" ou un "toxicomane", ou un autre concept similaire, le critère de base sera toujours la fréquence d'utilisation au cours d'une période donnée. On considère généralement qu'une personne est un consommateur actif si sa consommation est égale ou supérieure à un seuil de fréquence donné – même si ce seuil est exprimé par la formule "au moins une fois". Il est toujours trop difficile, dans la pratique, d'estimer les quantités de drogues consommées. Comme le font les auteurs de nombreuses études [par exemple, Simeone *et al.* (1997)], on considère qu'une personne a un antécédent de consommation de drogues dès l'instant où sa consommation atteint ou dépasse un seuil de fréquence donné.

Bien que la consommation puisse tomber en deçà du seuil, puis le dépasser à nouveau, c'est-à-dire être intermittente, ce type de détail n'est généralement pas pris en considération, ou bien les définitions de la période de référence excluent cette possibilité (par exemple, "tout cas d'utilisation au cours des 12 derniers mois" pour une étude couvrant une période de 12 mois). Par ailleurs, plusieurs événements de la vie des consommateurs de drogues qui ont une incidence sur ces techniques d'estimation – arrestations, admission et traitement et séjours dans des lieux d'accueil pour sans-abri – sont généralement considérés comme ponctuels, bien qu'ils puissent, dans certains cas, déboucher sur des situations durables. Par exemple, il est peu probable qu'une personne soit arrêtée alors qu'elle se trouve à l'hôpital, mais il n'en est généralement pas tenu compte dans les analyses. Si, toutefois, l'étude peut exploiter utilement ce type d'information, il faudra alors demander à un statisticien quelle est la meilleure façon de procéder [voir Simeone *et al.* (1997)].

Définition des sources de données

Les objectifs et les définitions dépendent, bien entendu, des sources de données existantes. On a déjà dit que si plusieurs sources étaient utilisées, il importait de retenir la même période de référence pour chacune d'entre elles. Il est tout aussi important que toutes les autres définitions correspondent d'une source à l'autre. Si c'est impossible, il faudra alors effectuer des contrôles ou des calculs, afin de déterminer dans quelle mesure les disparités entre les différentes définitions perturbent l'analyse globale des données. Cela peut être compliqué ou au contraire relativement facile. La définition du consommateur d'héroïne par injection est un exemple simple: il est possible que la définition retenue par la police ne se fonde pas sur les mêmes critères que celle appliquée par un centre de traitement des héroïnomanes. De même, si des données sont recueillies sur les tests de dépistage du VIH chez les consommateurs de drogues par injection, il se peut que ces données concernent la consommation de n'importe quelle drogue par injection (et pas seulement d'héroïne). "En traitement" constitue un exemple très courant de définition ambiguë, différentes sources faisant référence à différents types de traitement et, s'il s'agit d'une étude locale, à différentes unités de traitement, par exemple à l'intérieur ou à l'extérieur de la zone. Ce type d'ambiguïté est examiné dans plusieurs études de cas.

Sources de données possibles et méthodes à suivre

Évaluation préalable de la situation

Il est capital, au stade de la planification, de s'attacher à évaluer l'état des connaissances en matière d'usage de drogues. Peu importe que ces connaissances soient fiables ou sujettes à caution; ce qui compte, c'est de savoir si elles soulèvent des questions dont on pourra tirer parti pour la conception de l'étude. À cet égard, il convient non seulement de prendre en considération les statistiques et publications officielles, mais aussi d'envisager de recourir à des informateurs privilégiés.

Il faudrait, à tout le moins, évaluer rapidement et simplement toutes les sources de données dont on peut disposer et passer en revue, d'une manière générale, les possibilités de recueillir des données primaires. L'examen d'études du même type, qu'elles concernent le pays cible ou non, ainsi que d'études réalisées dans un contexte social analogue, peut être d'une grande importance.

Audit des sources nationales ou locales de données systématiques

Il est utile, avant d'entreprendre des travaux de recherche, d'avoir une idée très précise des sources de données et d'information qui sont constamment disponibles et de celles dont on peut extraire des informations utiles, d'une façon ou d'une autre, pour l'étude. Ces sources fournissent non seulement des données chiffrées exploitables à des fins épidémiologiques, mais aussi une liste d'établissements, de lieux et d'individus pouvant être utilisée pour constituer des échantillons aux fins de la

recherche. Il est toujours difficile d'obtenir des cadres d'échantillonnage adéquats, mais la difficulté est encore plus grande lorsqu'il s'agit de travaux sur la drogue. Il convient donc de vérifier ces sources avant de finaliser le projet.

Dans la mesure où la disponibilité des sources de données varie d'un pays à l'autre, et parfois d'un milieu social à l'autre dans un même pays, il est difficile de donner des conseils d'ordre général. La liste de pointage ci-dessous peut aider les chercheurs à passer en revue les sources susceptibles de faciliter le contact avec les usagers de drogues.

Liste de pointage Sources de données possibles

1. Registres et dossiers des centres de traitement
2. Médecins et professions médicales en général
3. Centres hospitaliers
4. Hôpitaux psychiatriques
5. Services spécialisés aux toxicomanes
6. Registres relatifs aux infections par VIH et autres registres de santé
7. Registres spécialisés de toxicomanes
8. Registres des décès
9. Registres des décès liés à la drogue
10. Rapports de police et casiers judiciaires

Il convient également de consulter, à ce sujet, les principes méthodologiques d'estimation de la prévalence établis par l'Observatoire européen des drogues et des toxicomanies (voir, dans l'introduction, la section intitulée "Documents complémentaires").

Caractéristiques des sources de données

Quelles que soient les sources d'information recensées que l'on juge utiles, il importe d'en vérifier certains aspects fondamentaux en rapport avec leur contenu, leur structure et leur accessibilité. La liste de pointage ci-dessous indique certains aspects généraux à examiner, mais il va de soi que les données requises seront fonction de l'objet de chaque projet de recherche. Cette liste indique, de façon générale, certaines des questions qu'il convient de se poser en ce qui concerne les sources de données et leur contenu.

Liste de pointage**Questions à se poser en ce qui concerne les bases de données et d'informations**

1. Qui possède et qui tient à jour la base de données?
2. Qui possède les données?
3. Quels types de drogues sont pris en compte?
4. La base comporte-t-elle des informations relatives à l'usage de drogues par injection?
5. Quelles types d'information la base renferme-t-elle concernant l'usage de drogues?
6. Les données concernent-elles des individus ou des événements/contacts?
7. La base fait-elle une distinction entre le premier contact et des contacts répétés?
8. Est-il possible d'identifier les individus à partir des données relatives aux événements/contacts?
9. Est-il possible d'identifier les individus à partir d'autres sources de données?
10. Quelle est la couverture géographique?
11. Sur quelle période les données portent-elles?
12. Sur quel support les données sont-elles conservées?
13. Existe-t-il des obstacles éventuels à l'accès aux données?
14. Peut-on envisager une amélioration des données dans l'avenir?
15. Peut-on envisager une utilisation de ces données en réseau?

Ici aussi, il convient de consulter les principes méthodologiques d'estimation de la prévalence établis par l'Observatoire européen des drogues et des toxicomanies (voir, dans l'introduction, la section intitulée "Documents complémentaires").

Sources de données spécifiques et possibilité de recueillir des données primaires

Si l'on entreprend de recueillir spécialement des données pour les besoins d'une étude menée selon l'une quelconque des procédures exposées dans le présent référentiel, on peut envisager de donner à la source que l'on va ainsi développer un caractère permanent en prévision de futures activités de surveillance. Dans le cadre de l'étude menée au Pakistan (étude de cas n° 4), par exemple, un registre des centres de traitement spécialisés a été créé, afin d'établir une référence pour la procédure d'estimation du multiplicateur de traitement qui a été utilisée. Sans que cela

soit l'objectif principal de l'étude, on a décidé de continuer à tenir ce registre, de façon qu'il puisse être utilisé à l'avenir pour d'autres études de la prévalence ou à toute autre fin. Bien qu'il ne soit pas une source officielle de données, le registre a été confié à des personnes capables de le tenir à jour.

Méthodes de collecte de données primaires

Études pilotes aux fins de l'enregistrement des données et de l'établissement des questionnaires

La collecte des données exige un effort d'organisation, de coopération et de planification considérable. Il faut généralement établir des formulaires papier, que ce soit pour transcrire, sous une forme convenant davantage à l'étude, les informations consignées dans les dossiers officiels ou pour enregistrer les informations recueillies lors des entretiens. Certes, dans certains cas, plutôt rares, les données existent déjà sur support électronique, de sorte qu'il suffit de leur appliquer un simple traitement informatisé pour les adapter à l'étude. Il convient, chaque fois que possible, de tirer pleinement parti de cet avantage et d'éviter la transcription sur papier. Normalement, cependant, lorsque des formulaires papier doivent être utilisés pour la transcription ou la collecte des données, il faut garder à l'esprit un certain nombre de considérations d'ordre général au stade de leur conception. Tout d'abord, il y a de très fortes chances pour que les données qu'ils contiennent doivent ensuite être entrées sur ordinateur; il importe donc de s'assurer les services d'un informaticien qui puisse aider à leur conception ou à tout le moins donner son approbation. Cela concerne aussi bien le mode d'enregistrement et de codage des données que la façon dont les personnes et les formulaires seront numérotés ou désignés par ce système de codage informatisé. Il est, par ailleurs, impératif de se demander qui remplira les formulaires: le chercheur, un enquêteur qualifié, ou l'enquêté?

Le respect d'un certain nombre de principes de base lors de la conception d'un questionnaire contribue à assurer la qualité des données recueillies. Si le contenu de chaque question relève à proprement parler de la compétence du chercheur, les aspects techniques de la conception du questionnaire doivent être envisagés en collaboration avec un spécialiste ou à l'aide d'un bon manuel (par exemple, Oppenheim, "*Questionnaire Design*"). Un statisticien peut aider à déterminer les aspects généraux des questions se rapportant expressément à la prévalence et tous les points de détail dont il faut tenir compte au stade de la conception. Bien que, s'agissant de l'estimation de la prévalence, il soit en général plus facile de choisir les questions à poser puisqu'il s'agit de mesurer des comportements spécifiques (et non des attitudes, qui sont plus difficiles à mesurer), il faut quand même avoir une certaine expérience de la conception de questionnaires. La liste de pointage ci-dessous indique certaines des considérations spécifiques à prendre en compte. Il convient aussi d'évaluer le temps qui sera probablement nécessaire pour poser les questions et consigner les réponses dans le cadre de l'entretien, car la patience de l'enquêté n'est pas illimitée.

Liste de pointage**Considérations fondamentales concernant la conception des questionnaires**

1. Veiller à ce que tout un chacun puisse répondre adéquatement et judicieusement à toutes les questions;
2. S'assurer que l'on puisse toujours indiquer "ne sait pas";
3. Tenir compte du fait que des individus peuvent refuser de répondre à certaines questions, voire à l'ensemble du questionnaire;
4. Expliquer à la personne qui remplit le questionnaire quelles sont les questions qui peuvent être laissées de côté dans des cas particuliers;
5. Faire en sorte qu'il n'y ait pas de réponses "blanches": il doit être répondu à chaque question par une marque ou une indication quelconque;
6. Déterminer quelles sont les catégories de réponses qui sont utiles et veiller à ce que ces catégories soient exhaustives et ne se recoupent pas;
7. Déterminer les cas dans lesquels on peut poser des questions à réponses multiples;
8. Décider des questions qui doivent rester ouvertes;
9. Veiller à ce que les questions impliquant une condition ("dans l'affirmative...") soient correctement structurées et renvoient à la section voulue;
10. Veiller à ce que la période considérée ou l'intervalle de référence soient précisés;
11. S'assurer que, si nécessaire, le lieu est précisé.

Étant donné le nombre de questions que soulève la conception des formulaires et des questionnaires, ainsi que les procédures à suivre pour la conduite des entretiens et l'enregistrement des données, il est indispensable de réaliser des études pilotes appropriées afin de vérifier ce qui se passera dans la pratique. Bien qu'étant généralement de faible envergure, ces études sont précieuses, car elles mettent au jour des erreurs, des ambiguïtés et des difficultés imprévues concernant les questions, la consignation des réponses ou les procédures. Il importe aussi, bien évidemment, qu'elles se déroulent dans des conditions aussi proches que possible de la réalité, afin que l'on puisse en tirer le meilleur parti.

Sélection des enquêtes par des méthodes inhabituelles

Lorsqu'il faut, pour déterminer la prévalence de l'abus de drogues, compter ou interroger des toxicomanes aux fins de la collecte de données, il arrive souvent que la difficulté d'entrer en contact avec ces derniers pousse les chercheurs à établir des

échantillons selon des méthodes inhabituelles. Ces méthodes sont inhabituelles en ce sens qu'il faut pallier, par des solutions inédites, l'impossibilité d'avoir accès à un cadre d'échantillonnage, tel que celui qu'exigent les sondages probabilistes classiques. La plus importante d'entre elles est l'échantillonnage sur place, exposée ci-dessous, mais on a aussi fréquemment recours à l'échantillonnage par désignation ou "échantillonnage en boule de neige", qui consiste à prendre un petit échantillon aléatoire établi par les méthodes classiques, puis à l'agrandir en demandant aux enquêtés de fournir des informations sur leurs connaissances et leurs amis. On se reportera à un ouvrage publié par Intraval [Bielemann et de Bie (1992)] pour de plus amples informations sur la procédure et les moyens de contrôle à appliquer lorsqu'on utilise cette méthode.

Échantillonnage binomial (échantillonnage sur place)

Au stade de l'élaboration de la stratégie de collecte des données, on doit souvent faire face au problème fondamental exposé ci-dessus, à savoir qu'il est difficile de localiser les usagers de drogues. Il est fréquent qu'on interroge les usagers de drogues avérés dans des lieux où l'on a le plus de chances de les trouver, notamment dans des centres de garde à vue ou de détention, des centres de traitement publics ou privés et des foyers pour sans-abri. Ce type d'échantillonnage – qui consiste à établir l'échantillon sur place et non sur la base d'un cadre d'échantillonnage ou d'une liste d'usagers de drogues préétablis – est connu sous le nom d'échantillonnage binomial ou d'échantillonnage sur place [Bielemann et de Bie (1992) et Goodman (1961)]. Lorsqu'on utilise cette méthode, il faut déterminer notamment si la présence fréquente d'un individu sur le site considéré aura pour effet de fausser l'échantillon au point que celui-ci ne sera plus représentatif de la population cible; en effet, si un individu fréquente plus souvent le site qu'un autre, il aura plus de chances d'être inclus dans l'échantillon. Cet aspect doit être pris en considération lorsqu'il s'agit d'établir des échantillons aléatoires d'individus ou bien des échantillons d'événements (en l'occurrence, la fréquentation du site); si l'on souhaite, par exemple, extrapoler les résultats à l'ensemble de la population toxicomane, il faut pondérer les taux d'événements par l'inverse de la probabilité que l'utilisateur ait été inclus dans l'échantillon [voir, par exemple, Simeone *et al.* (1997)], faute de quoi ce dernier sera représentatif non des toxicomanes, mais de la fréquentation du site.

Procédures de sélection des sites

Lorsqu'on applique une procédure qui exige un échantillonnage, un comptage ou encore la compilation d'informations dans différents sites, et lorsqu'il est impossible de retenir tous les sites pertinents aux fins de la collecte de données (ce sont en général les moyens dont on dispose qui font que l'on doit se contenter d'un échantillon de ces sites), il faut, dans la mesure du possible, élaborer une méthode rigoureuse pour établir cet échantillon. S'il est possible d'obtenir un cadre d'échantillonnage des sites qui couvre exhaustivement, ou à tout le moins de manière très satisfaisante, la population cible, des procédures de stratification et de groupement peuvent alors

être appliquées pour établir un échantillon à la fois maniable et représentatif. Les chercheurs qui ne sont pas au fait des procédures d'échantillonnage probabiliste devraient consulter un spécialiste.

Entretiens et contrôle de la véracité des réponses

En ce qui concerne les entretiens, un autre problème se pose: il faut distinguer les usagers de drogues des non-usagers et contrôler la véracité des réponses faites par les enquêtés concernant l'usage de drogues. On est en droit de penser que tout individu qui reconnaît consommer de la drogue répondra honnêtement à toutes les autres questions qu'on lui posera à ce sujet, mais il ne faut pas s'attendre que tous les usagers de drogues qui fréquentent les sites retenus soient disposés à parler ouvertement de leur comportement toxicomane. Si on le juge nécessaire, on pourra alors procéder à des analyses biologiques qui révéleront si l'individu a récemment fait usage de drogue; ces analyses peuvent aussi être effectuées sur un échantillon aléatoire, afin d'estimer la proportion d'usagers qui refusent d'admettre qu'ils consomment de la drogue. Dans la pratique, toutefois, cela n'est en général pas possible, car les individus ne se soumettent pas volontiers à de telles analyses – à moins, peut-être, qu'on leur offre une incitation financière. En général, les chercheurs ne peuvent qu'espérer que les enquêtés disent la vérité.

Études de mortalité par cohortes et multiplicateurs relatifs aux décès

Pour ce qui est en particulier des études faisant appel à des multiplicateurs, on a indiqué qu'il pouvait être nécessaire de recueillir des données primaires, afin de déterminer la valeur du multiplicateur à appliquer à la référence. S'agissant du multiplicateur relatif à la mortalité, il est peu probable que l'on puisse entreprendre, spécialement pour les besoins d'une telle étude, des travaux visant à déterminer le taux de mortalité parmi les toxicomanes étudiés. Les taux de mortalité sont généralement établis par des études de cohorte longues et complexes. On pourra consulter des ouvrages sur la question [par exemple, Frischer *et al.* (1998)] ainsi que la publication de l'Observatoire européen des drogues et des toxicomanes sur l'estimation indirecte (voir, dans l'introduction, la section intitulée "Documents complémentaires").

Publication des résultats et examen par des confrères

Qu'elle se présente sous forme de rapport interne (pour une institution de financement ou un organisme public) ou qu'elle soit destinée à être publiée sous forme d'ouvrage scientifique, l'étude doit impérativement faire l'objet d'un examen par des confrères et par d'autres spécialistes. Aucun chercheur n'est en mesure de prévoir tous les aléas, tous les écueils ni toutes les difficultés qui se poseront, et il a tout à gagner à soumettre les résultats de ses travaux à d'autres chercheurs et spécialistes. Tous les chercheurs, les revues scientifiques et les instituts de financement sérieux et renommés le font systématiquement.

Cet examen est primordial non seulement pour le chercheur, mais aussi pour la préservation de la qualité des travaux de recherche publiés au niveau international. C'est le meilleur moyen d'assurer que la validité d'une étude soit admise par la profession et reconnue au-delà du cercle immédiat des collègues et des bailleurs de fonds de ses auteurs.

Lorsque les travaux sont avant tout destinés à être publiés sous la forme d'un rapport pour un organisme public, il est souvent possible de les publier également dans des revues scientifiques réputées. L'un des problèmes, en effet, est que les résultats des travaux sont enfouis dans de longs rapports auxquels il est difficile d'avoir accès, soit parce qu'il faut obtenir une autorisation, soit parce que le rapport est trop volumineux. Un rapport de recherche publié sous une forme destinée au public est souvent beaucoup plus succinct et met généralement l'accent sur les principaux points d'intérêt, de sorte que les résultats obtenus font l'objet d'une plus grande diffusion. Avant de publier un rapport sous une autre forme, il faudra obtenir l'autorisation de l'institution qui l'a commandé et veiller à ce qu'il n'y ait pas d'autres difficultés d'ordre administratif pour y avoir accès.

Un rapport final devrait exposer, dans le détail, les méthodes retenues tant pour la collecte de données que pour l'analyse. Par ailleurs, il importe de signaler, par souci de crédibilité, les lacunes et les difficultés, prévues ou imprévues, auxquelles il a fallu faire face. D'une manière générale, chacune des quatre phases d'un projet de recherche, à savoir la planification des travaux, la collecte des données pertinentes, l'analyse des données et la rédaction du rapport, revêt la même importance et peut exiger autant de travail que les autres. La description des modalités de la collecte de données et la présentation des résultats de l'analyse sont tout aussi importants. Le meilleur moyen de veiller à l'exactitude est de rédiger le rapport au fur et à mesure des travaux. Si l'on attend le dernier moment, il est possible que l'on oublie des détails et des réserves qui ont leur importance, ainsi que des difficultés auxquelles il a fallu faire face. Décrire les modalités de la recherche au fur et à mesure des travaux non seulement permet d'établir un rapport plus complet, mais aussi allège la charge de travail au cours des étapes finales.

Ressources générales utiles pour les études de prévalence

Annexe

Ressources du Programme mondial d'évaluation de l'abus de drogues

Site Web du GAP sur l'épidémiologie – Objet et utilisation

Le GAP tient à jour un site Web que le lecteur devrait consulter. L'un des objectifs prioritaires du projet est de développer ce site, afin qu'il donne accès à des rapports, des publications et d'autres sites Web utiles. Consulter l'adresse Internet: www.unodc.org, puis rechercher GAP. Voir également le centre de ressources du GAP (<http://gap.gide.net>).

Sites Web utiles

Département santé mentale et toxicomanie, *Guide to Drug Abuse Epidemiology* (WHO/MSD/MSB 00.3) (Genève, Organisation mondiale de la santé, 2000), consulter l'adresse Internet suivante: www.who.int/substance_abuse/PDFfiles/EPI_GUIDE_A.pdf.

Observatoire européen des drogues et des toxicomanies, *Estimation de la prévalence de la consommation problématique des drogues en Europe*, Monographie scientifique n° 1 (Luxembourg, Bureau des publications officielles de l'Union européenne, 1997). Consulter l'adresse Internet suivante: www.emcdda.org/situation/themes/problem_drug_use.shtml.

Documentation générale

Ouvrages cités dans le texte

Archibald, C. P., *et al.*, "Estimating the size of hard-to-reach populations: a novel method using HIV testing data compared to other methods", AIDS, vol. 15 (supplément) (1992), S41-S48.

- Bieleman, B., et de Bie, E., "Between the lines – a study of the nature and extent of cocaine use in Rotterdam" (Rotterdam, IntraVal Foundation, 1992).
- Bishop, Y. M. M., Fienberg, S. E., et Holland, P. W., "Estimating the size of a closed population", dans *Discrete Multivariate Analysis: Theory and Practice* (Cambridge, Massachusetts, MIT Press, 1975), p. 229 à 256.
- European Monitoring Centre for Drugs and Drug Addiction, *Methodological Guidelines to Estimate the Prevalence of Problem Drug Use at Local Level* (Lisbonne, 2000), consulter l'adresse Internet suivante: www.emcdda.org/situation/themes/problem_drug_use.shtml.
- Fitch, C., et al., *Responding to Drug and Alcohol Use in Jersey: Key Findings Report* (Jersey, States of Jersey Drug Strategy Unit, 2001).
- Frank, B., et al., "Seeking truth in heroin indicators: the case of New York City", *Drug and Alcohol Dependence*, vol. 3 (1978), p. 345 à 358.
- Frischer, M., "Estimating the prevalence of drug abuse using the mortality multiplier method: an overview", dans OEDT, *Estimation de la prévalence de la consommation problématique des drogues en Europe...*
- Frischer, M., et al., "A comparison of different methods for estimating the prevalence of problematic drug abuse in Great Britain", *Addiction*, vol. 96, n° 10 (2001), p. 1465 à 1476.
- Frischer, M., et al., "Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug users in Glasgow, Scotland", *American Journal of Epidemiology*, vol. 138, n° 3 (1993), p. 170 à 181.
- Frischer, M., et al., "Mortality among injecting drug users: a critical reappraisal", *Journal of Epidemiology and Community Health*, vol. 47, n° 1 (1993), p. 59 à 63.
- Goodman, L., "Snowball sampling", *The Annals of Mathematical Statistics*, vol. 32 (1961), p. 148 à 170.
- Hall, W., et al., "How many opioid users are there in Australia?", *Medical Journal of Australia*, vol. 173 (2000), p. 528 à 531.
- Hay, G., et Smit, F., "Estimating the number of drug injectors from needle exchange data", dans *Addiction Research and Theory* (à paraître).
- Hickman, M., et al., "Estimating the prevalence of problematic drug use: a review of methods and their application", *Bulletin des stupéfiants* (2002).

- Hook, E. B., et Regal, R. R., "Capture-recapture methods in epidemiology: methods and limitations", *Epidemiologic Reviews*, vol. 17 (1995), p. 243 à 264.
- Hser, Y., *et al.*, *Techniques for the Estimation of Illicit Drug User Prevalence: an Overview of Relevant Issues* (Washington (D.C.), National Institute of Justice, 1992).
- International Working Group for Disease Monitoring and Forecasting, "Capture-recapture and multiple record systems estimation I: history and theoretical development", *American Journal of Epidemiology*, vol. 142 (1995), p. 1047 à 1057.
- Lynskey, M., et Hall, W., "Jurisdictional trends in opioid deaths 1988-1995", *Australian and New Zealand Journal of Public Health*, vol. 23 (1998), p. 519 à 524.
- Mariani, F., Guaiana, R., et Di Fiands, T., "An epidemiological overview of the situation of illicit drug abuse in Italy", *The Journal of Drug Issues*, vol. 24 (1994), p. 579 à 595.
- Mastro, T. D., *et al.*, "Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method", *American Journal of Public Health*, vol. 84 (1994), p. 1094 à 1099.
- McKetin, R., *et al.*, *Drug Trends 1998. A Comparison of Drug Use and Trends in Three Australian States*, National Drug and Alcohol Research Centre Monograph No. 41 (Sydney, University of New South Wales, 1999).
- Observatoire européen des drogues et des toxicomanies, Monographie scientifique n° 1, *Estimation de la prévalence de la consommation problématique des drogues en Europe* (Lisbonne, 1997).
- Rehmann, Griffiths et Taylor, "Rapid assessment report on the Pakistan national assessment exercise" (UNODC, 2002).
- Reuter, P., "Prevalence estimation and policy formulation", *Journal of Drug Issues*, vol. 23 (1993), p. 167 à 184.
- SHAKTI Project, CARE Bangladesh (Dhaka).
- Simeone, R., *et al.*, "A plan for estimating the number of "hardcore" drug users in the United States" (Washington (D.C.), Drug Policy Research Group, Office of National Drug Control Policy, 1997).
- Tilling, K., et Sterne, J. A. C., "Capture-recapture models including covariate effects", *American Journal of Epidemiology*, vol. 149, n° 4 (1999), p. 392 à 400.
- Wickens, T. D., "Quantitative methods for estimating the size of the drug using population", *Journal of Drug Issues*, vol. 23 (1993), p. 185 à 216.

Woodward, J. A., Retka, R., et Nig, L., "Construct validity of heroin abuse estimators", *The International Journal of the Addictions*, vol. 19, n° 1 (1984), p. 93 à 117.

Autres ouvrages utiles

Estimations par capture-recapture

Bonett, D. G., Woodward, J. A., et Bentler, P. M., "A linear model for estimating the size of a closed population", *British Journal of Mathematical and Statistical Psychology*, vol. 39 (1986), p. 28 à 40.

Brownie, C., "Reader reaction – recent models for mark-recapture and mark-resighting data", *Biometrics*, vol. 43 (décembre 1987), p. 1017 à 1022.

Chao, A., "Estimating the population size for capture-recapture data with unequal catchability", *Biometrics*, vol. 43 (décembre 1987), p. 783 à 791.

Cormack, R. M., "Log-linear models for capture-recapture", *Biometrics*, vol. 45 (juin 1989), p. 395 à 413.

Cowan, C. D., et Malec, D., "Capture-recapture models when both sources have clustered observations", *Journal of the American Statistical Association*, vol. 81, n° 394 (juin 1986).

Doscher, M. L., et Woodward, J. L., "Estimating the size of subpopulations of heroin users: applications of log-linear models to capture-recapture sampling", *The International Journal of the Addictions*, vol. 18, n° 2 (1983), p. 167 à 182.

Haber, M., "Testing for pairwise independence", *Biometrics*, vol. 42 (juin 1986), p. 429 à 435.

Hook, E. B., et Regal, R. R., "Effect of variation in probability of ascertainment by sources upon 'capture-recapture' estimates of prevalence", *American Journal of Epidemiology*, vol. 137, n° 10 (1992).

Seber, G.A.F., "A review of estimating animal abundance", *Biometrics*, vol. 42 (juin 1986), p. 267 à 292.

Wittes, J. T., "Capture-recapture methods for assessing the completeness of case ascertainment when using multiple information sources", *Journal of Chronic Diseases*, vol. 27 (1974), p. 25 à 36.

Wolter, K. M., "Capture-recapture estimation in the presence of a known sex ratio", *Biometrics*, vol. 46 (mars 1990), p. 157 à 162.

Échantillonnages communautaires et en "boule de neige"

Bieleman, B., et de Bie, E., *Between the Lines – a Study of the Nature and Extent of Cocaine Use in Rotterdam* (Rotterdam, IntraVal Foundation, 1992).

Biernacki, P., "Snowball sampling. Problems and techniques of chain referral sampling", *Sociological Methods and Research*, vol. 10, n° 2 (novembre 1981), p. 141 à 163.

Goodman, L., "Snowball sampling", *The Annals of Mathematical Statistics*, vol. 32 (1961), p. 148 à 170.

Griffiths, P., *et al.*, "Reaching populations of drug users by the use of privileged access interviewers: methodological and practical issues", *Addiction*, vol. 88 (1993), p. 617 à 1626.

Kaplan, C. D., Korf, D., et Sterk, C., "Temporal and social contexts of heroin-using populations. An illustration of the snowball sampling technique", *The Journal of Nervous and Mental Disease*, vol. 175, n° 9 (1987).

van Meter, K. M., "Methodological and design issues: techniques for assessing the representatives of snowball samples" (source non communiquée).

Multiplicateurs de mortalité et de décès

Darke, S., Ross, J., et Hall, W., "Overdose among heroin users in Sydney, Australia: II. Responses to overdose", *Addiction*, vol. 91, n° 3 (1996), p. 413 à 417.

Davoli, M., *et al.*, "Risk factors for overdose mortality: a case-control study within a cohort of intravenous drug users", *International Journal of Epidemiology*, vol. 22, n° 2 (1993), p. 273 à 277.

Farrell, M., *et al.*, "Suicide and overdose among opiate addicts", *Addiction*, vol. 91, n° 3 (1996), p. 321 à 323.

Frischer, M., *et al.*, "Mortality among injecting drug users: a critical reappraisal", *Journal of Epidemiology and Community Health*, vol. 4 (1993), p. 59 à 63.

Gossop, M., *et al.*, "Frequency of non-fatal overdose", *British Medical Journal*, vol. 313 (1996).

Hammersley, R., Cassidy, M., et Oliver, J., "Drugs associated with drug-related deaths in Edinburgh and Glasgow, novembre 1990 à octobre 1992", *Addiction*, vol. 90, p. 959 à 965.

Home Office, "Statistics of drug addicts notified to the Home Office of the United Kingdom of Great Britain and Northern Ireland, 1994", *Home Office Statistical Bulletin 17/95* (Londres, 1995).

Joe, G. W., Lehman, W., et Simpson, D., "Addict death rates during a four-year post-treatment follow-up", *American Journal of Public Health*, vol. 2, n° 7 (1982), p. 703 à 709.

Ruttenber, A., et Luke, J., "Heroin-related deaths: new epidemiological insights", *Science*, vol. 226, p. 14 à 20.

Walsh, R., "Opioid drug accidental deaths in the Newcastle area of New South Wales, 1970-1987", *Drug and Alcohol Review*, vol. 10, p. 79 à 83.

Zador, D., Sunjic, S., et Darke, S., "Heroin-related deaths in New South Wales, 1992: toxicological findings and circumstances", *Medical Journal of Australia*, vol. 164, p. 204 à 207.

Ouvrages généraux

Bishop, Y. M. M., Fienberg, S. E., et Holland, P. W., "Estimating the size of a closed population", dans *Discrete Multivariate Analysis: Theory and Practice* (Cambridge, Massachusetts, MIT Press, 1975), p. 229 à 256.

Fitch., C., *et al.*, *Responding to Drug and Alcohol Use in Jersey: Key Findings Report* (Jersey, States of Jersey Drug Strategy Unit, 2001).

Frank, O., "Estimation of population totals", dans *Perspectives on Social Network Research*, Holland, P. W., et Leinhardt, S. L., éd. (New York, Academic, 1979).

Ghodse, H., "Casualty departments and the monitoring of drug dependence", *British Medical Journal*, vol. 1 (1977), p. 1381 et 1382.

Hickman, M., *et al.*, "Estimating the prevalence of problematic drug use: a review of methods and their application". *Bulletin des stupéfiants...*

Hook, E. B., et Regal, R. R., "Capture-recapture methods in epidemiology: methods and limitations", *Epidemiologic Reviews*, vol. 17 (1995), p. 243 à 264.

International Working Group for Disease Monitoring and Forecasting, "Capture-recapture and multiple record systems estimation I: history and theoretical development", *American Journal of Epidemiology*, vol. 142 (1995), p. 1047 à 1057.

Mastro, T. D., *et al.*, "Estimating the number of HIV-infected injection drug users in Bangkok: a capture-recapture method", *American Journal of Public Health*, vol. 84 (1994), p. 1094 à 1099.

McKeganey, N., *et al.*, "Female streetworking prostitution and HIV infection in Glasgow", *British Medical Journal*, vol. 305 (octobre 1992), p. 801 à 804.

OEDT, *Methodological Guidelines to Estimate the Prevalence of Problem Drug Use at Local Level* (Lisbonne, 2000), consulter l'adresse Internet suivante: www.emcdda.org/situation/themes/problem_drug_use.shtml.

OEDT, *Recommended Draft Technical Tools and Guidelines – Key Epidemiological Indicator: Prevalence of problem drug use* (Lisbonne, 2000), consulter l'adresse Internet suivante: www.emcdda.org/situation/themes/problem_drug_use.shtml.

Picklands III, J., et Raghavachari, M., "Exact and asymptotic inference for the size of a population", *Biometrika*, vol. 74, n° 2 (1987), p. 355 à 363.

SHAKTI Project, CARE Bangladesh (Dhaka).

Ouvrages divers

Bloor, M., Wood, F., et Palmer, S., *Estimating the Prevalence of Injecting Drug Use and Serious Drug Use in Wales* (Cardiff, Social Research Unit, 1998).

Crabbe, T., Donmall, M. C., et Millar, T., "Validation of the University of Manchester drug abuse database", *Journal of Epidemiology and Community Health*, vol. 53, n° 3 (1999), p. 159 à 164.

Ditton, J., et Frischer, M., "Computerised projection of future heroin epidemics: a necessity for the 21st century?", *Journal of Substance Use and Abuse* (à paraître).

Frischer, M., *et al.*, "A comparison of trends in problematic drug abuse from two reporting systems", *Journal of Public Health Medicine*, vol. 22, n° 3 (2000), p. 362 à 367.

Hickman, M., *et al.*, "Surveillance of problem drug use in the UK: a review of a regional drug abuse database", *Journal of Public Health*, vol. 21 (1999), p. 271 à 277.

Prevalence of HIV in the United Kingdom: Report of the Unlinked Anonymous Prevalence Monitoring Programme in the United Kingdom (Londres, Department of Health, 1999).

