

## Les défis de la mise à jour de l'indice de défavorisation avec les données du recensement de 2011 et de l'Enquête nationale auprès des ménages (ENM)



En 2011, le questionnaire long du recensement du Canada a été remplacé par l'Enquête nationale auprès des ménages (ENM), introduisant un biais potentiel pour l'utilisation des données issues de cette source à petite échelle et ayant une incidence sur la mise à jour de l'indice de défavorisation.

Ce texte donne un aperçu des scénarios méthodologiques qui ont été testés et qui ont permis la mise à jour de l'indice de défavorisation 2011.

### Sommaire

Introduction	2
Méthode	2
Résultats	3
Le choix de la version de l'indice 2011	5
Constats et recommandations	6

### Faits saillants

- Le remplacement du formulaire long du recensement par la participation volontaire à l'Enquête nationale auprès des ménages (ENM) a fait bondir le taux global de non-réponse en 2011.
- La mise à jour de l'indice de défavorisation matérielle et sociale (IDMS) avec les données de 2011 risquait d'introduire un biais de non-réponse et de mener à des résultats imprécis.
- Quatre versions de l'indice de défavorisation de 2011 comprenant des ajustements variés ont été testées pour évaluer sa qualité et pour identifier la version la plus comparable et précise pour des analyses longitudinales.
- Les différences entre les versions se sont révélées minimes. Celle ne procédant à aucune correction présente des résultats similaires aux autres versions. Elle a donc été retenue à des fins de comparabilité avec les indices précédents.
- La mise à jour de l'IDMS 2011 présente des performances très satisfaisantes à l'égard des divers critères. Toutefois, cette mise à jour complexe et longue a nécessité de surcroît la collaboration de Statistique Canada pour l'exploitation de leur fichier maître.
- L'indice de défavorisation demeure une très bonne mesure des inégalités sociales tant à l'échelle canadienne, provinciale ou régionale, surtout en l'absence d'information socioéconomique individuelle dans les bases de données administratives.

## Introduction

De 1991 à 2006, l'indice de défavorisation matérielle et sociale (IDMS) a été construit à partir du Recensement de la population canadienne mené par Statistique Canada aux cinq ans, suivant une méthodologie comparable. Cet indice est composé de six indicateurs dont trois proviennent du questionnaire court et les trois autres du questionnaire long.

En 2011, le questionnaire long du recensement couvrant 20 % des ménages privés, a été remplacé par l'Enquête nationale auprès des ménages (ENM). Le formulaire de l'ENM a été envoyé à 30 % des ménages privés mais son caractère volontaire a fait bondir le taux global de non-réponse. Un taux de non-réponse plus élevé entraîne un risque de biais plus élevé, de sorte que les estimations de l'ENM pourraient être moins fiables que celles obtenues à partir du questionnaire long du recensement. C'est d'autant plus vrai pour des niveaux géographiques fins. Or, l'indice de défavorisation est basé sur l'aire de diffusion (AD), l'unité du recensement la plus fine pour laquelle des données sont diffusées. Dans ce contexte, élaborer un indice de défavorisation selon les mêmes paramètres qu'auparavant demeure-t-il encore valide et fiable? Est-il possible d'apporter des corrections pour tenir compte des biais potentiels?

## Méthode

Pour créer l'indice de défavorisation de 2011, la même méthodologie de base que celle utilisée dans la version originale basée sur les données de 1991 a été retenue (à quelques exceptions près) par souci de comparabilité dans le temps (Pampalon *et al.*, 2009a). Tout d'abord, les mêmes indicateurs qui forment traditionnellement les deux composantes matérielle et sociale de l'indice ont été retenus, soit :

### Composante sociale

- La proportion de personnes vivant seules.
- La proportion de familles monoparentales.
- La proportion de personnes séparées, divorcées ou veuves.

### Composante matérielle

- Le revenu moyen.
- La proportion de personnes sans diplôme d'études secondaires.
- La proportion de personnes occupées.

Ce sont les trois indicateurs de la composante matérielle qui proviennent des données de l'ENM en 2011.

Après avoir ajusté cinq<sup>1</sup> des six indicateurs par la méthode de standardisation directe pour tenir compte de l'âge et du sexe des individus, une analyse en composantes principales a été effectuée afin d'intégrer ces indicateurs calculés pour les aires de diffusion retenues.

Pour tenter de minimiser l'impact des biais potentiels du passage à l'ENM, quatre versions de l'indice de défavorisation de 2011, brièvement décrites ci-après, ont été produites :

1. Version 1 : standardisation (méthode directe) des trois indicateurs provenant de l'ENM, sans aucune correction pour tenir compte de la non-réponse plus élevée. Cette version correspond à l'approche méthodologique traditionnelle pour produire l'IDMS.
2. Version 2 : similaire à la version 1, mais chaque indicateur provenant de l'ENM a subi une correction en fonction de son estimation de 2006 pour les AD présentant un risque de biais élevé.
3. Version 3 : similaire à la version 1, mais la valeur de chaque indicateur provenant de l'ENM est corrigée en fonction des estimations des AD voisines ayant des taux de réponse jugés acceptables et ce, pour les AD présentant un risque de biais élevé.
4. Version 4 : similaire à la version 1, mais la valeur de chaque indicateur provenant de l'ENM est corrigée en fonction de l'estimation obtenue à un niveau géographique supérieur, c'est-à-dire la subdivision de recensement (municipalité) en milieu rural ou le secteur de recensement en milieu urbain. Encore une fois, seules les AD avec un risque de biais élevé subissent une correction. Plus la non-réponse associée à une AD est élevée, plus la correction apportée est importante.

<sup>1</sup> La proportion de familles monoparentales n'a pu être ajustée car l'unité d'analyse est la famille et non l'individu, lequel constitue l'unité d'analyse des cinq autres indicateurs.

## Résultats

Plusieurs critères ont été envisagés afin d'évaluer la qualité de l'indice de défavorisation de 2011 par rapport à celui de 2006, ainsi que pour déterminer si une ou plusieurs des quatre versions étaient supérieures aux autres.

En premier lieu, la structure factorielle de l'indice de 2011 est similaire à celle des recensements antérieurs, et ce pour les quatre versions testées. Les deux mêmes composantes – matérielle et sociale – ressortent clairement et la variabilité totale expliquée par ces composantes est même plus élevée qu'en 2006 (70 % en 2011 contre 65 % en 2006 pour le Canada).

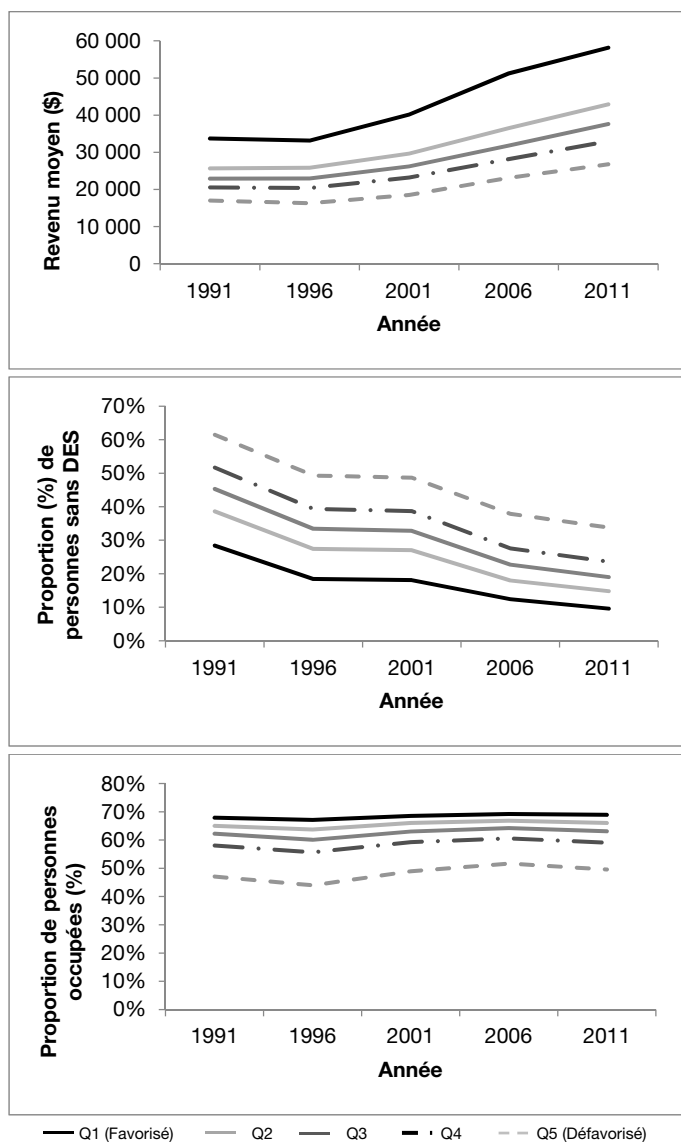
En deuxième lieu, il était important de vérifier si la proportion de la population couverte par l'indice de 2011 représentait toujours une forte majorité de la population. Nous avons constaté que la présence d'aires de diffusion avec un taux global élevé de non-réponse n'a pas d'impact majeur sur la population couverte par l'indice : 85 % de la population est utilisée pour l'analyse en composantes principales et seulement 3,5 % de la population ne reçoit pas d'indice. Cette situation est comparable à celle de l'indice de 2006. La version 3 est la moins adéquate selon ce critère.

Troisièmement, l'évolution des indicateurs par quintile a été étudiée, particulièrement pour ceux de la composante matérielle qui proviennent de l'ENM. Cette dernière n'étant pas obligatoire, nous nous attendions à ce que certains groupes de la population comme les plus riches, les plus pauvres ou encore les jeunes répondent moins. Il y aurait, par conséquent un écart moindre entre les groupes de population formant les cinq quintiles pour les indicateurs de revenu, d'emploi et de scolarité en 2011. La tendance historique observée depuis 1991 aurait donc dû être brisée. Or, la tendance se poursuit. Aux échelles canadienne et québécoise, le revenu, l'emploi et la scolarité montrent toujours des écarts importants entre les quintiles de la composante matérielle (figure 1). Ce constat vaut pour toutes les versions produites.

Quatrièmement, un regard porté aux changements de quintiles des AD dont le lien géographique est le même pour deux recensements successifs indique que la moitié des AD conservent le même quintile de défavorisation. La plupart des AD qui changent de quintiles le font de façon mineure de  $\pm 1$  (par exemple, en passant du quintile 3 au quintile 4). L'analyse de ces changements de quintiles des AD entre 2006 et 2011 a démontré que la distribution des différences est semblable à ce qu'elle était entre 2001 et 2006 (périodes sans changement au recensement). On observe légèrement plus de changements, mais la hausse n'est pas significative. Le tableau 1 présente ces comparaisons pour le Canada avec la version 1. Les résultats sont comparables pour toutes les versions.

Cinquièmement, nous avons testé la capacité de l'indice de défavorisation de 2011 à détecter des inégalités sociales de santé (ISS). Nous avons déjà démontré que l'indice, une mesure écologique, sous-estime les véritables inégalités sociales de santé comparativement à celles calculées à partir des données individuelles (Pampalon *et al.*, 2009b; Pampalon *et al.*, 2009c). Par contre, nous avons observé lors des précédentes mises à jour de l'indice que l'utilisation de la version de l'indice centrée sur la période couverte par les données (par exemple la version 2006 de l'indice utilisée avec la période 2004-2008 de décès, plutôt que la version 2001 de l'indice) mettait en évidence de plus grandes ISS. Par conséquent, la meilleure version de l'indice de 2011 devrait, elle aussi, permettre de détecter les plus grandes ISS, lorsqu'appliquée de façon centrée sur une période de données de santé. Pour ce faire, nous avons comparé, pour la période 2009-2013 et pour les indicateurs suivants : la mortalité prématurée (moins de 75 ans); la mortalité par suicide; et la fécondité des adolescentes (moins de 20 ans), les ratios ajustés des taux de mortalité ou de fécondité des groupes de la population les plus défavorisés (Q5) sur les plus favorisés (Q1), obtenus en utilisant l'indice de 2006 (version non centrée sur la période de données) et les quatre versions de l'indice de 2011 (versions centrées sur la période de données).

**Figure 1** Évolution du revenu personnel moyen, de la proportion de personnes sans diplôme d'études secondaires et de la proportion de personnes occupées selon le quintile matériel de l'indice de défavorisation, de 1991 à 2011, Canada



**Tableau 1** Différences de quintiles des dimensions matérielle et sociale entre les indices de défavorisation de 2001 et 2006 et entre ceux de 2006 et 2011, version 1, Canada

2001 versus 2006		
Différence absolue	Matériel %	Social %
0	<b>45,08 %</b>	<b>51,66 %</b>
1	<b>38,28 %</b>	<b>37,19 %</b>
2	13,12 %	9,16 %
3	3,08 %	1,71 %
4	0,45 %	0,27 %
Total	100,00 %	100,00 %

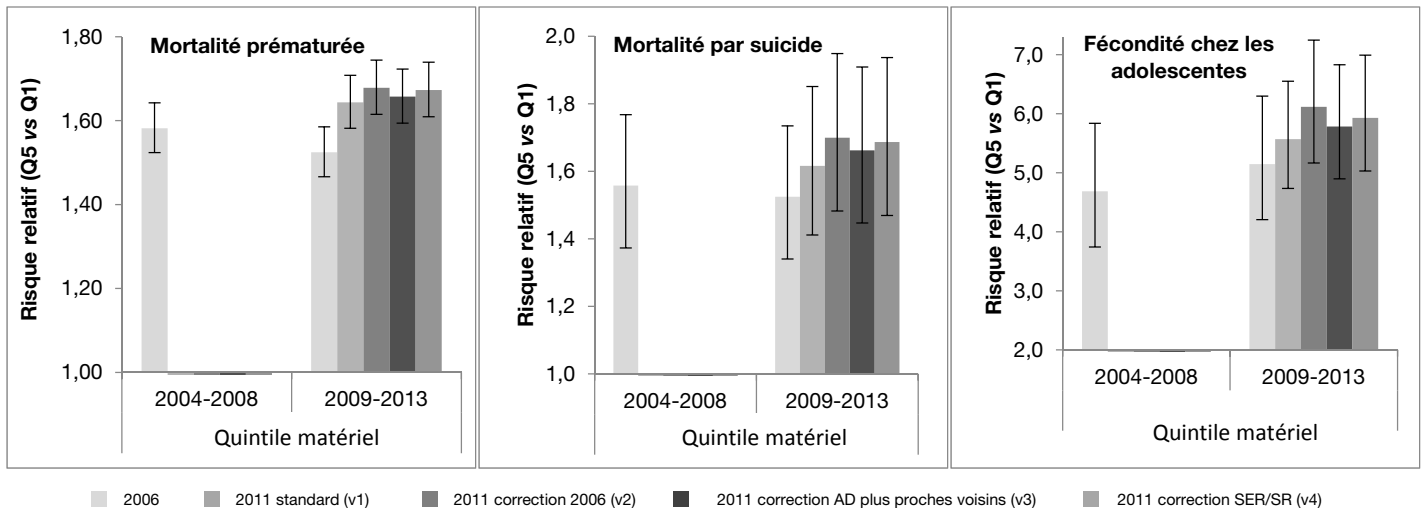
  

2006 versus 2011 (version 1)		
Différence absolue	Matériel %	Social %
0	<b>44,02 %</b>	<b>58,02 %</b>
1	<b>38,37 %</b>	<b>35,35 %</b>
2	13,89 %	5,87 %
3	3,35 %	0,71 %
4	0,38 %	0,06 %
Total	100,00 %	100,00 %

La figure 2 montre que pour la mortalité prématurée (moins de 75 ans), la mortalité par suicide et la fécondité des adolescentes (moins de 20 ans), la composante matérielle de 2011 permet de détecter de plus grandes inégalités que celles mesurées avec l'indice de 2006, et ce, peu importe la version. Dans les trois cas, la version 2 de l'indice de 2011 (corrigé avec les indicateurs de 2006) présente les plus grands écarts, mais les différences avec les trois autres versions ne sont pas substantielles.

En dernier lieu, une validation de l'indice de défavorisation de 2011 à une échelle locale a été réalisée avec l'aide d'intervenants régionaux de deux régions sociosanitaires du Québec : l'Estrie et la Montérégie. Le processus visait à déterminer si les différences de quintiles majeures ( $\pm 3$  ou  $\pm 4$ ) entre 2006 et 2011 de certaines aires de diffusion reflétaient une réalité locale de véritables changements entre les deux recensements. Le processus n'a pas été concluant, en raison d'un biais de mémoire (ex. : les changements ont-ils eu lieu avant ou après 2010-2011? Quelle était la situation sur le terrain en 2005-2006 dans les AD concernées?) et de l'absence d'éléments de comparaison (il s'agissait de la première validation de ce genre). Enfin, selon les intervenants interpellés, seulement un tiers des changements majeurs était pleinement explicable. Ces résultats auraient par ailleurs pu être similaires en 2006 si ce processus de validation avait été utilisé.

**Figure 2** Risque relatif<sup>a</sup> de la mortalité prématurée, de la mortalité par suicide et de la fécondité chez les adolescentes (moins de 20 ans) des personnes les plus défavorisées (Q5) comparativement aux plus favorisées (Q1) d'après la dimension matérielle de l'indice de 2006 et des quatre versions de l'indice 2011, 2004-2006 et 2009-2013, Québec



<sup>a</sup> Risque relatif ajusté pour tenir compte de l'âge, de la zone géographique et de l'autre dimension de défavorisation (composante sociale).

## Le choix de la version de l'indice 2011

L'examen des résultats des quatre versions de la composante matérielle de l'indice indique que la version 2 est légèrement plus performante pour détecter des inégalités. Les différences entre les versions sont toutefois très minimes dans la majorité des cas. À priori donc, l'utilisation des versions 1, 2, 3 ou 4 n'entraîne pas de différence très marquée dans l'estimation des risques relatifs. Sur les plans opérationnel et technique, les versions 1 et 2 présentent le plus d'avantages. La version 2 est légèrement plus performante pour détecter des inégalités et, par définition, pour réduire les différences entre 2006 et 2011. Toutefois, elle présentait un défaut majeur dans l'éventualité où l'ENM allait être reconduite, soit l'impossibilité d'être utilisée dans le futur. En effet, il serait impossible de se rabattre sur les données de 2011 pour avoir des données aussi fiables que celles de 2006.

Les deux autres options testées (versions 3 et 4) présentent plusieurs problèmes. Dans la version 3, on constate notamment qu'il est parfois difficile de trouver des AD voisines avec des taux globaux de non-réponse « acceptables ». En ce qui concerne la version 4, son principal défaut est que l'AD n'est pas indépendante du niveau géographique supérieur. Comme la taille du niveau géographique supérieur peut varier passablement, particulièrement pour les subdivisions de recensement, les comparaisons sont plus problématiques.

La version finale retenue est donc la première, soit celle sans ajustement de la valeur des trois indicateurs formant la défavorisation matérielle. Étant donné que la version 1 présente des résultats similaires aux autres versions mais sans procéder à des corrections, nous la privilégions car c'est la plus simple et la plus comparable aux indices précédents. La version originale (c'est-à-dire la version 1) allait être reproduite en 2016 si l'ENM avait été conservée (ce qui n'aurait pas été le cas, par exemple, avec la version 2 corrigée avec les indicateurs de 2006). Aussi, elle évite d'introduire une nouvelle forme de biais comme ce pourrait être le cas avec les autres versions.

## Constats et recommandations

La plupart des résultats présentés sont encourageants pour la validité de l'indice de défavorisation de 2011 qui, rappelons-le, est construit partiellement avec des données de l'ENM à l'échelle des aires de diffusion. La structure factorielle est non seulement similaire à celles des précédentes versions de l'indice, la variance totale expliquée par les composantes matérielle et sociale est même plus élevée qu'en 2006. La population couverte par l'indice est pratiquement aussi élevée que pour les recensements antérieurs. L'analyse des six indicateurs par quintile de défavorisation continue de faire ressortir d'importantes disparités, particulièrement entre les plus favorisés et les plus défavorisés. Malgré des variations légèrement plus importantes, les changements de quintiles observés au niveau des AD entre 2006 et 2011 ne se distribuent pas différemment des changements relevés entre 2001 et 2006. Par ailleurs, l'indice de 2011 continue non seulement de détecter des inégalités de santé importantes, mais il représente une nette amélioration par rapport à l'indice de 2006 lorsqu'utilisé pour l'analyse de données récentes de 2009 à 2013.

Aucune des quatre versions de l'indice ne se démarque vraiment des autres. La version 1 de l'indice, qui n'applique aucune correction, est par conséquent le choix tout indiqué. C'est la plus simple et la plus comparable aux indices précédents en plus d'être facilement reproductible. Les trois autres versions avec correction risquaient d'introduire une nouvelle forme de biais.

Finalement, bien que la mise à jour de 2011 de l'IDMS présente des performances très satisfaisantes à l'égard des divers critères que nous avons analysés, celle-ci a été complexe, longue et n'aurait pu se faire sans la collaboration de Statistique Canada pour l'exploitation de leur fichier maître.

L'indice n'a jamais été un outil parfait répondant à tous les besoins. Les problèmes liés à l'ENM sont un incitatif à l'utiliser avec davantage de prudence. Nous recommandons donc son usage pour évaluer l'ampleur des inégalités sociales à une échelle canadienne, provinciale ou régionale, ou encore pour introduire une mesure socio-économique comme variable de contrôle et d'ajustement dans des analyses statistiques qui autrement devraient s'en passer. L'indice de

défavorisation représente dans ce cas une très bonne mesure des inégalités sociales, surtout en l'absence d'information socioéconomique individuelle dans les bases de données administratives.

Nous invitons par contre à la prudence si l'objectif est d'utiliser l'indice à une échelle locale, notamment pour comparer des aires de diffusion entre elles. Nous avons en effet vu qu'une proportion non négligeable de celles-ci change de quintile entre les recensements, parfois même de façon très contrastée avec le recensement précédent. Ces changements peuvent refléter de réelles différences sur le terrain mais celles-ci n'ont pu être vérifiées. Dans d'autres cas, ils peuvent être induits par des biais méthodologiques, en particulier le taux global de non-réponse. Pour un tel usage, l'interprétation de l'indice de défavorisation de 2011 devrait toujours être accompagnée du taux global de non-réponse de chaque aire de diffusion et d'un indicateur de discordance entre les effectifs de population estimés par le recensement et ceux fournis par l'Enquête nationale des ménages. Lorsque possible, l'indice de 2011 devrait être comparé à l'indice de défavorisation de 2006 et éventuellement à celui de 2016, ces deux derniers servant d'outils de validation. Malgré tout, à l'échelle locale, l'indice de défavorisation demeure beaucoup plus robuste que les estimations du revenu, de la scolarité et de l'emploi prises une à une.

## Références

Pampalon R, Hamel D, Gamache P, Raymond G. (2009a). A deprivation index for health planning in Canada. *Chronic Diseases in Canada, Maladies Chroniques au Canada*, 29 (4): 178-191.

Pampalon R, Hamel D, Gamache P. (2009b). Health inequalities in urban and rural Canada. Comparing inequalities in survival according to an individual and area-based deprivation index. *Health & Place*, 16: 416-420.

Pampalon R, Hamel D, Gamache P. (2009c). A comparison of individual and area-based socio-economic data for monitoring social inequities in health. *Health Reports*, 20 (4): 85-94.

## Les défis de la mise à jour de l'indice de défavorisation avec les données du recensement de 2011 et de l'Enquête nationale auprès des ménages (ENM)

### AUTEURS

Philippe Gamache, statisticien  
Denis Hamel, statisticien  
Bureau d'information et d'études en santé des populations

### RÉVISION

Christine Blaser, conseillère scientifique  
Bureau d'information et d'études en santé des populations

### MISE EN PAGE

Hélène Fillion, adjointe à la direction  
Vice-présidence à la valorisation scientifique et aux communications  
Lyne Théorêt, agente administrative  
Bureau d'information et d'études en santé des populations

### REMERCIEMENTS

Les auteurs souhaitent remercier l'Agence de la santé publique du Canada pour sa collaboration et sa contribution financière à la mise à jour de l'indice de défavorisation 2011.

Ce document est disponible intégralement en format électronique (PDF) sur le site Web de l'Institut national de santé publique du Québec au : <http://www.inspq.qc.ca>.

Les reproductions à des fins d'étude privée ou de recherche sont autorisées en vertu de l'article 29 de la Loi sur le droit d'auteur. Toute autre utilisation doit faire l'objet d'une autorisation du gouvernement du Québec qui détient les droits exclusifs de propriété intellectuelle sur ce document. Cette autorisation peut être obtenue en formulant une demande au guichet central du Service de la gestion des droits d'auteur des Publications du Québec à l'aide d'un formulaire en ligne accessible à l'adresse suivante : <http://www.droitauteur.gouv.qc.ca/autorisation.php>, ou en écrivant un courriel à : [droit.auteur@cspq.gouv.qc.ca](mailto:droit.auteur@cspq.gouv.qc.ca).

Les données contenues dans le document peuvent être citées, à condition d'en mentionner la source.

Dépôt légal – 1<sup>er</sup> trimestre 2017  
Bibliothèque et Archives nationales du Québec  
ISBN : 978-2-550-77343-6 (PDF)

© Gouvernement du Québec (2017)

N° de publication : 2202