

Un cadre d'évaluation des données administratives

Mylène Lavigne et Christian Nadeau¹

Résumé

La diminution des taux de réponse observée par plusieurs instituts nationaux de statistique, leur volonté de limiter le fardeau de même que l'importance des contraintes budgétaires auxquelles ils font face favorisent une utilisation accrue des données administratives pour la production de renseignements statistiques. Les sources de données administratives qu'ils sont amenés à considérer doivent être évaluées afin d'en déterminer l'adéquation à l'utilisation, et ce en tenant compte de plusieurs aspects. Une démarche d'évaluation des sources de données administratives en vue de leur utilisation comme intrant à des processus de production de renseignements statistiques a récemment été élaborée à Statistique Canada. Celle-ci comprend essentiellement deux phases. La première phase ne nécessite que l'accès aux métadonnées associées aux données administratives considérées alors que la deuxième est effectuée à partir d'une version des données permettant l'évaluation. La démarche et l'outil d'évaluation sont présentés dans cet article.

Mots Clés : Qualité des données administratives; Cadre d'assurance de la qualité.

1. Introduction

La diminution des taux de réponse observée par Statistique Canada au cours des dernières années, la volonté de réduire le fardeau de réponse et les pressions budgétaires croissantes amènent l'organisme à envisager des alternatives efficientes à ses pratiques courantes de collecte et de production de renseignements statistiques. Une approche privilégiée par l'organisme afin de faire face à de tels défis consiste à accroître l'utilisation des données administratives pour remplacer, compléter et étoffer ses programmes statistiques. Un certain nombre de facteurs favorisent une telle approche, notamment la richesse et la diversité des sources de données administratives disponibles, leur accessibilité et les moyens techniques maintenant disponibles pour leur exploitation.

L'utilisation des données administratives par un institut national de statistique est généralement souhaitable (Brackstone, 1987). Il est cependant essentiel que les décisions concernant l'acquisition et l'utilisation d'un ensemble de données administratives précis tiennent compte de leur nature et de leur qualité, du type d'utilisation, de l'incidence de la qualité de ces données sur la qualité des produits statistiques qui découlent de leur utilisation, de la confidentialité et du respect de la vie privée de même que des aspects légaux et financiers. Il apparaît judicieux dans le cadre d'un processus d'acquisition et d'utilisation de fichiers de données administratives à des fins statistiques, de procéder à une évaluation systématique qui tient compte de ces aspects afin que des décisions éclairées soient prises. Daas et al. (2009), Berning et al. (2013) et Australian Bureau of Statistics (2009) ont déjà suggéré de telles approches d'évaluation.

Certaines activités menées par différents programmes statistiques permettent déjà d'évaluer la qualité des données administratives à Statistique Canada. Toutefois, il est devenu opportun de développer un cadre d'évaluation des données administratives visant à soutenir l'ensemble des programmes statistiques dans leurs efforts d'acquisition de nouveaux fichiers. Ce cadre fait l'objet du présent article. Les objectifs liés à la démarche d'évaluation et les principes sous-jacents aux fonctions d'acquisition et d'utilisation des données administratives à des fins statistiques sont d'abord présentés à la section 2. La démarche et l'outil d'évaluation font ensuite l'objet des sections 3 et 4 respectivement, suivis par une brève conclusion.

¹Mylène Lavigne, Statistique Canada, 100 promenade Tunney's Pasture, Ottawa, Ontario, Canada, K1A 0T6 (mylene.lavigne@statcan.gc.ca); Christian Nadeau, Statistique Canada, 100 promenade Tunney's Pasture, Ottawa, Ontario, Canada, K1A 0T6 (christian.nadeau@statcan.gc.ca).

2. Principes et objectifs

Les fichiers de données administratives considérées pour une utilisation à des fins statistiques doivent préalablement faire l'objet d'une entente d'acquisition de données entre Statistique Canada et leur fournisseur². À cet effet, on supposera que l'acquéreur³ a déjà pris connaissance de l'existence du fichier de données administratives considéré. L'acquéreur pourrait également en connaître certains aspects et avoir identifié au moins sommairement les utilisations qu'il compte en faire. Avant de finaliser l'entente d'acquisition de données, il procédera à une évaluation plus approfondie qu'il effectuera au cours des deux premières phases du processus d'utilisation et d'acquisition illustré à la figure 2-1.

Figure 2-1
Modèle d'acquisition et d'utilisation des données administratives



Le processus d'acquisition et d'utilisation des données administratives se déroule en quatre phases débutant après la découverte du fichier, soit les phases « d'exploration », « d'acquisition initiale », « d'acquisition » et « d'utilisation ». À la phase d'exploration, l'acquéreur contacte le fournisseur du fichier. Il se familiarisera avec le contenu, les concepts, les caractéristiques et le contexte de création du fichier par le biais de communications avec le fournisseur, de consultations de la documentation et des métadonnées, mais sans nécessairement avoir accès au fichier. À défaut d'avoir été identifiées au préalable de façon précise, les utilisations devraient l'être au cours de cette phase, mais pourraient encore évoluer par la suite. L'information recueillie au cours de cette phase fait l'objet d'une évaluation qui permettra à l'acquéreur de juger la pertinence de poursuivre l'évaluation afin d'en arriver à une entente d'acquisition de données pour les utilisations visées.

Lors de la phase « d'acquisition initiale », l'acquéreur peut négocier une entente avec le fournisseur afin d'obtenir une version des données lui permettant d'évaluer essentiellement (mais pas exclusivement) l'exactitude et la cohérence des données. Suite à cette évaluation, la pertinence de conclure une entente d'acquisition des données devrait être établie.

À la phase « d'acquisition », l'acquéreur négocie une entente d'acquisition afin d'obtenir une version du fichier pouvant être utilisée aux fins visées. Les évaluations effectuées lors des phases précédentes pourraient avoir mené à l'identification de modifications à apporter au fichier (format, contenu, etc.) et celui-ci devrait être validé avant de passer à la phase d'utilisation.

La dernière phase, celle de « l'utilisation », inclut la mise en œuvre des utilisations visées du fichier menant à la production de renseignements statistiques. Elle comportera également les étapes de conception, d'élaboration et de mise à l'essai des processus de production impliquant le fichier, la validation et l'évaluation des produits pour lesquels les données administratives ont été utilisées, de même que l'évaluation a posteriori des données administratives utilisées en intrant. Dans le cas de processus statistiques répétés et tel qu'illustré à la figure 2-1, les phases d'acquisition et d'utilisation pourraient être répétées avec les versions subséquentes des données administratives sans que les phases d'exploration et d'acquisition initiale ne soient nécessairement répétées.

Bien que des évaluations soient nécessaires tout au long du processus d'acquisition et d'utilisation des données administratives, le cadre d'évaluation présenté ici ne porte que sur les évaluations associées aux phases

² Le « fournisseur » est défini comme l'organisation qui fournit les données administratives.

³ L' « acquéreur » est défini comme la personne à Statistique Canada responsable de négocier l'entente d'acquisition de données avec le fournisseur.

d'exploration et d'acquisition initiale. L'objectif principal de la démarche d'évaluation présentée à la section 3 est donc d'évaluer l'adéquation de fichiers de données administratives en vue de leur utilisation comme intrants à des processus statistiques préalablement définis, et non d'évaluer les produits statistiques issus de l'utilisation de données administratives.

L'évaluation de fichiers de données administratives dans le cadre d'un tel objectif dépend en grande partie des utilisations qui en seront faites. Lavallée (2007) décrit sept types d'utilisation: la création et le maintien de bases de sondage, la construction de plans de sondage, la substitution partielle à la collecte directe, la vérification et l'imputation, l'estimation indirecte (et la pondération), la totalisation directe et l'évaluation d'enquêtes (la validation des données). Le type d'utilisation pourrait avoir un impact direct sur le choix des différents éléments d'évaluation et sur l'importance relative qui leur est accordée. À titre d'exemple, l'élément « couverture de la population » pourrait avoir plus d'importance lorsque l'utilisation visée est le développement d'une base de sondage.

L'adoption des six dimensions de la qualité définies dans le cadre d'assurance de la qualité de Statistique Canada (Statistique Canada, 2002) a été privilégiée pour l'élaboration du cadre d'évaluation des données administratives. Les éléments d'évaluation présentés à la section 3 ont été identifiés avec le souci de couvrir ces six dimensions. Le tableau 2-1 présente une interprétation de ces dimensions dans le contexte de l'utilisation des données administratives.

Tableau 2-1

Dimensions de la qualité appliquées aux données administratives

<p>La pertinence des données administratives correspond à la mesure dans laquelle les besoins réels de l'organisme sont satisfait relativement à l'utilisation visée. Pour être qualifiées de pertinentes, les données administratives doivent par exemple répondre aux motivations de leur acquisition (réduction des coûts, réduction du fardeau de réponse, ...), ou contenir l'information recherchée au sujet des unités statistiques cibles pour la période référence visée.</p>
<p>L'exactitude des données administratives correspond à la mesure dans laquelle l'information décrit bien le phénomène qu'elle doit mesurer. On l'associe aux erreurs pouvant survenir pendant leur production, par exemple les erreurs de déclaration, de non-déclaration, de couverture, d'interprétation, de saisie et de traitement.</p>
<p>L'actualité des données administratives correspond au délai entre le point de référence (ou la fin de la période de référence) auquel se rapporte l'information et la date à laquelle les données sont disponibles pour l'organisme statistique. Il doit souvent y avoir un compromis entre l'actualité et l'exactitude de l'information. Une version préliminaire d'un fichier de données administratives pourrait par exemple comporter certaines lacunes du point de vue de la couverture de l'ensemble des unités ou de la vérification de certaines variables. L'actualité influera sur la pertinence.</p>
<p>L'accessibilité des données administratives correspond à la facilité avec laquelle l'organisme statistique peut se les procurer dans leur intégralité. Il s'agit notamment de la facilité avec laquelle on peut constater que l'information existe, de même que le caractère approprié de la présentation de l'information ou du média au moyen duquel on peut avoir accès aux données. Dans certains cas, le coût peut également constituer un aspect de l'accessibilité.</p>
<p>L'intelligibilité des données administratives correspond à la disponibilité de renseignements et de métadonnées nécessaires à l'interprétation et à l'utilisation appropriée de ces données. Il s'agit en général de renseignements sur les variables, les classifications et les concepts sous-jacents utilisés, sur les méthodes de collecte et de traitement des données et sur les indicateurs de l'exactitude des données administratives.</p>
<p>La cohérence des données administratives correspond à la mesure dans laquelle celles-ci peuvent être jumelées à d'autres données de l'organisme statistique dans un vaste cadre analytique au fil du temps. L'utilisation de concepts, de classifications et de populations cibles types favorise la cohérence.</p>

Étant donné la grande variété d'utilisations possibles, il était important de développer un outil flexible pouvant s'adapter aux différents besoins. La démarche suggérée repose sur l'analyse d'un ensemble d'éléments dont l'importance varie selon le contexte d'acquisition plutôt que sur l'application de règles de décision strictes à partir de résultats empiriques. Dans ce contexte, l'acquéreur est responsable de déterminer l'importance relative attribuée aux divers éléments sur lesquels il basera sa décision et l'ordre dans lequel ils seront évalués.

Enfin, la démarche vise également à documenter et stocker l'information sur la qualité de la source de données administratives. L'information concernant l'évaluation constitue une source de métadonnées importante qui sera utile pour d'autres utilisateurs potentiels ou si d'autres utilisations sont considérées dans le futur. L'information pourrait également servir de source de métadonnées pour les produits statistiques les utilisant.

3. Démarche d'évaluation en vue d'une utilisation

En tenant compte des objectifs et des principes énoncés à la section précédente, la démarche d'évaluation présentée consiste à collecter l'information nécessaire et à l'analyser afin d'évaluer l'adéquation du fichier de données administratives en tant qu'intrant pour une utilisation donnée. Pour ce faire, la démarche d'évaluation proposée se divise en quatre parties : la description de l'initiative d'acquisition visée, l'évaluation à la phase d'exploration, l'évaluation à la phase d'acquisition initiale et finalement l'analyse globale et la prise de décision.

Le tableau 3-1 résume la démarche d'évaluation en présentant les éléments évalués selon les thèmes dans chacune des parties de celui-ci. Lors de l'**initiative d'acquisition** l'acquéreur documente l'initiative d'acquisition, identifie et décrit le fichier de données administratives visé, et ensuite précise les objectifs et les besoins relatifs à cette acquisition. Cette première partie de la démarche d'évaluation est d'une grande importance puisque le fichier de données administratives devrait être évalué en fonction des objectifs et utilisations décrites à cette phase. Toutefois, certains éléments d'information recherchés pourraient se préciser lors de l'évaluation des phases d'exploration et d'acquisition initiale.

Tableau 3-1
Démarche d'évaluation

	Thème	Élément
Description de l'initiative d'acquisition visée	Identification de l'initiative d'acquisition	Identification de l'initiative d'acquisition et identification du responsable de l'initiative d'acquisition
	Information générale sur le fichier de données administratives	Nom, description, renseignements sur le fournisseur et renseignements sur d'autres organismes liés
	Objectifs et besoins relatifs à l'initiative d'acquisition	Type d'utilisation, motivations, objectifs, étapes pour atteindre les objectifs et besoins concernant la population visée, la période de référence visée, l'échéancier et le contenu
Évaluation à la phase d'exploration	Environnement institutionnel	Modalités d'accès et fiabilité
	Caractéristiques du fichier de données administratives	Population (D1, D6), période de référence (D1, D6), dates de réception (D1, D3), contenu (D1, D6), processus de collecte (D1, D2), format du fichier (D4), format des variables (D4), traitement (D1, D2), métadonnées (D5)
	Incidence de l'utilisation du fichier de données administratives	Coûts, ressources humaines et fardeau de réponse
Évaluation à la phase d'acquisition initiale	Évaluation par type d'erreur	Erreurs de couverture, de non-réponse, de déclaration, de mesure et de traitement (D1, D2)
	Validation de données	Validation des différentes statistiques produites à partir du fichier de données administratives (D2)
Analyse globale et prise de décision	Sommaire de l'information recueillie	Note : Les six dimensions de la qualité : D1 = pertinence, D2 = exactitude, D3 = actualité, D4 = accessibilité, D5 = intelligibilité et D6 = cohérence.

L'évaluation à la **phase d'exploration** consiste à évaluer certaines caractéristiques de la source de données administratives ne nécessitant pas l'accès aux données. L'évaluation est effectuée à partir d'informations pouvant être obtenues du fournisseur via des rencontres, des discussions, le partage d'information, de documentation ou autres. Elle met l'accent sur les dimensions de pertinence, d'actualité, d'accessibilité et de cohérence des données. L'évaluation à la phase d'exploration porte sur l'environnement institutionnel⁴, les caractéristiques du fichier de données administratives en lien avec l'utilisation visée et l'incidence de l'utilisation du fichier de données administratives sur certains aspects du processus statistique.

L'évaluation à la **phase d'acquisition initiale** met quant à elle l'emphasis sur l'exactitude et la cohérence des données. Elle consiste à évaluer certains types d'erreurs en calculant différents indicateurs d'exactitude à partir d'une version de la source de données et à la valider en la comparant à d'autres sources connues. Une version

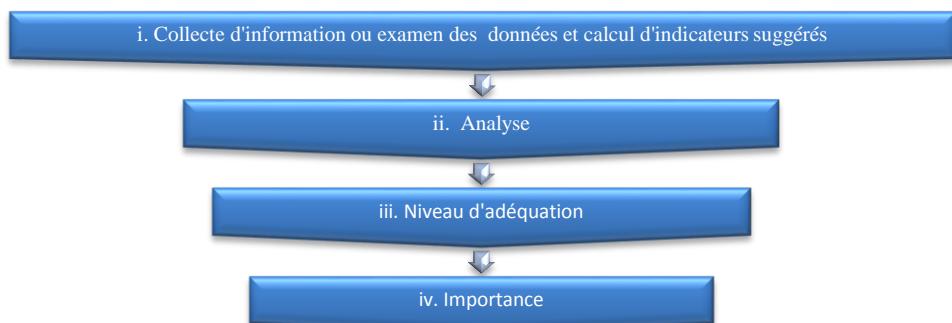
⁴ L'évaluation de l'environnement institutionnel tient compte des facteurs organisationnels ou institutionnels qui peuvent avoir un impact sur la capacité du fournisseur à fournir des données qui rencontrent les critères de qualité requis. Ceci inclut une évaluation des standards de qualité du fournisseur et des procédures pour en assurer le maintien.

préliminaire du fichier de données administratives (échantillon ou autres) peut être utilisée pour cette évaluation. Celle-ci permet aussi, si nécessaire, une mise à jour de certains résultats de l'évaluation à la phase d'exploration. Elle est complémentaire à la première phase d'évaluation et permet d'évaluer la pertinence d'acquérir ou non la source de données administratives évaluées.

Pour ces deux phases d'évaluation, les éléments sont évalués indépendamment les uns des autres en fonction de l'utilisation visée. Pour ce faire, il est proposé de procéder à l'évaluation de chaque élément des phases d'exploration et d'acquisition initiale en quatre étapes. La figure 3-1 illustre les quatre étapes d'évaluation des phases d'exploration et d'acquisition initiale.

Figure 3-1

Étapes de l'évaluation pour chaque élément et en fonction de l'utilisation visée aux phases d'exploration et d'acquisition initiale



Pour chaque phase, la première étape consiste à collecter et consigner l'information recherchée sur l'élément évalué. La deuxième étape est une analyse de l'adéquation du fichier de données administratives à l'utilisation visée en ne considérant que l'élément d'information traité. Cette analyse devrait tenir compte de l'information recueillie à la première étape et de la description de l'initiative d'acquisition. La troisième étape consiste à évaluer le niveau d'adéquation du fichier de données administratives (élevé, modéré ou faible) pour cet élément qui est déduit de l'analyse consignée à la deuxième étape. Finalement, la quatrième étape consiste à attribuer à l'élément d'information évalué une importance relative (obligatoire, important, peu important ou sans importance) dans le cadre de l'utilisation visée. Cette importance relative sera prise en compte lors de l'analyse globale et la prise de décision. Toutefois, dans certains cas, la décision de ne pas poursuivre une entente d'acquisition pourrait même être prise suite à la phase d'exploration, si des lacunes importantes sont notées.

L'analyse globale et la prise de décision constituent la dernière partie de la démarche d'évaluation. Le cheminement proposé pour la compilation et l'analyse vise à dégager, de façon simple et rapide, l'essentiel de l'information collectée et consignée par le biais des évaluations aux phases d'exploration et d'acquisition initiale pour en arriver à la prise de décision. Elle tient compte de l'adéquation du fichier de données administratives aux besoins et de l'importance accordée à chacun des éléments recueillis. Ce cheminement vise davantage à guider de façon systématique le personnel responsable de la décision à travers l'information collectée qu'à fournir des règles absolues de prise de décision.

La démarche d'évaluation présentée suppose que le fichier de données administratives n'est pas accessible au début du processus d'acquisition. Dans un tel cas, les évaluations des phases d'exploration et d'acquisition initiale sont effectuées séquentiellement. Advenant l'accessibilité du fichier au début du processus d'acquisition, les deux phases peuvent être effectuées parallèlement sans problème. Il est important de noter que la démarche d'évaluation offre une certaine flexibilité en ce qui a trait à l'ordre d'évaluation des éléments. Celui-ci peut être fonction de l'importance des éléments évalués et du niveau d'effort requis. On pourrait chercher à minimiser l'effort requis pour rejeter les sources initialement perçues inadéquates en évaluant d'abord les éléments les plus importants et requérant le moins d'effort. On pourrait favoriser l'exhaustivité de l'évaluation seulement pour les sources semblant adéquates de prime abord. Pour certaines utilisations, il pourrait même s'avérer que certains éléments ne sont pas évalués.

4. Outil d'évaluation

L'outil permettant la mise en œuvre de la démarche d'évaluation présentée est constitué d'un guide (Statistique Canada, 2015a), d'un questionnaire (Statistique Canada, 2015b), d'une grille d'évaluation et d'une grille de recommandations. Le guide inclut une description de la démarche d'évaluation en définissant les erreurs, les concepts, l'information recherchée, ainsi que les niveaux d'adéquation et l'importance de chacun des éléments. Il décrit aussi certaines méthodes d'évaluation et certains indicateurs afin d'aider l'acquéreur lors de l'évaluation du fichier de données administrative à la phase d'acquisition initiale. De plus, des instructions sont données concernant le questionnaire et les grilles d'évaluation et de recommandation.

Le questionnaire est le véhicule choisi pour collecter et consigner l'information relative aux trois premières parties de la démarche d'évaluation d'un fichier de données administratives. En plus de répondre à l'objectif principal d'évaluer l'adéquation de fichiers de données administratifs, il permet aussi de répondre à l'objectif de documenter et stocker l'information sur la qualité des données. Le questionnaire est composé de huit parties. Les parties A, B et C du questionnaire visent à collecter l'information nécessaire à la description de l'initiative d'acquisition visée. Les parties D à H du questionnaire servent à l'évaluer les différents éléments des phases d'exploration et d'acquisition initiale et visent à recueillir les éléments d'information nécessaires à l'évaluation de l'adéquation du fichier de données administratives à l'utilisation visée. Pour chaque élément d'évaluation des phases d'exploration et d'acquisition initiale, le questionnaire a le même format, soient les quatre étapes illustrées à la figure 3-1 de la section précédente. Par exemple, pour l'élément « Population », le questionnaire ressemble à celui illustré à la figure 4-1.

Figure 4-1

Exemple du questionnaire

Dimension de la qualité : Pertinence et cohérence			
i. Décrire brièvement la population du fichier de données administratives.			
ii. Déterminer si la population du fichier décrite en (i) répond aux besoins énoncés à la partie C.5. Si elle n'y correspond pas exactement, décrire l'ampleur des écarts et leurs répercussions sur l'utilisation visée. Identifier les moyens pouvant être utilisés pour atténuer de telles répercussions, ou noter l'absence de moyen.			
iii. Étant donné l'analyse faite en (ii), évaluer le niveau d'adéquation du fichier de données administratives à l'utilisation visée, en ce qui a trait à la population .			
<i>Niveau d'adéquation</i>			
<i>Élevé</i>			
<i>O</i>			
<i>Modéré</i>			
<i>O</i>			
<i>Faible</i>			
<i>O</i>			
iv. Quelle importance doit être accordée à l'élément « population » dans le cadre de l'utilisation visée?			
<i>Importance</i>			
<i>Obligatoire</i>			
<i>O</i>			
<i>Important</i>			
<i>O</i>			
<i>Peu important</i>			
<i>O</i>			
<i>Sans importance</i>			
<i>O</i>			

Les grilles d'évaluation et de recommandation utilisent des chiffriels qui permettent de compiler et d'analyser globalement l'information recueillie lors de l'évaluation aux phases d'exploration et d'acquisition initiale. Si les phases d'évaluation se font séquentiellement, cette étape se fera à deux reprises soit après la phase d'exploration et ensuite après la phase d'acquisition initiale. Par contre, s'il s'avérait que les deux phases se font parallèlement, cette deuxième étape ne se ferait qu'une seule fois après avoir rempli le questionnaire. Un exemple de la grille d'évaluation est illustré à la figure 4-2. Le contenu de la grille d'évaluation peut être trié par ordre croissant d'importance ou filtré afin de permettre une bonne visualisation du niveau d'adéquation aux besoins selon l'importance accordée aux éléments. La grille de recommandation permet de consigner et de justifier la décision de poursuivre le processus d'acquisition après une analyse de la grille d'évaluation et de l'information pertinente documentée lors de l'évaluation aux phases d'exploration et d'acquisition initiale du questionnaire. Un exemple de la grille de recommandation est illustré à la figure 4-3.

Figure 4-2
Exemple de la grille d'évaluation

Phase	Élément	Importance	Adéquation
Exploration	D.1 Modalité d'accès	1-Obligatoire	1-Élevé
	D.2 Fiabilité	3-Peu important	2-Modéré
	E.1 Population	2-Important	3-Faible
	E.2 Période de référence	3-Peu important	3-Faible
	E.3 Dates de réception potentielles	2-Important	1-Élevé
	E.4 Contenu	2-Important	2-Modéré
	E.5 Processus de collecte	4-Sans importance	Non évalué
	E.6 Format du FDA	3-Peu important	3-Faible
	E.7 Format des variables	3-Peu important	2-Modéré
	E.8 Traitement	2-Important	1-Élevé
	E.9 Métadonnées	4-Sans importance	2-Modéré
	F.1 Coûts	2-Important	3-Négatifs
	F.2 Ressources humaines	3-Peu important	2-Neutres
	F.3 Fardeau de réponse	2-Important	1-Positifs
Acquisition initiale	G.1 Erreurs de couverture	1-Obligatoire	1-Élevé
	G.2 Erreurs de non-réponse	2-Important	2-Modéré
	G.3 Erreurs de déclaration, de mesures et de traitements	3-Peu important	2-Modéré
	H.1 Validation basée sur la comparaison de statistiques	2-Important	1-Élevé

Figure 4-3
Exemple de recommandation

Question 1: Énoncé de la recommandation	Recommandation de poursuivre	
	Oui	O
Question 2: Justification basée sur l'analyse des résultats de la grille d'évaluation, de l'information relative aux éléments obligatoires et importants.	Non	O

5. Conclusion

Différentes méthodes sont utilisées afin d'évaluer la qualité des données administratives. Toutefois, il est devenu opportun de développer un cadre d'évaluation dédié aux données administratives visant à soutenir l'ensemble des programmes statistiques dans leurs efforts d'acquisition. Le cadre d'évaluation des données administratives en cours d'élaboration à Statistique Canada vise essentiellement l'évaluation de l'adéquation de fichiers de données administratives comme intrant à des utilisations précises dans des processus statistiques préalablement identifiés.

Une démarche d'évaluation séquentielle et un outil d'évaluation sont proposés afin de juger du potentiel d'adéquation d'un fichier de données administratives. L'approche proposée favorise une utilisation judicieuse des ressources dédiées à l'analyse en permettant d'identifier les éléments clefs à évaluer. Les évaluations proposées sont d'abord effectuées lors de l'exploration d'un nouveau fichier de données administratives, en utilisant les métadonnées et donc avant même d'avoir accès au fichier, puis dans le cadre de l'acquisition d'une version des données permettant une évaluation plus approfondie. Plusieurs éléments sont évalués afin de juger de l'adéquation du fichier à l'utilisation visée. Chacun d'entre eux fait l'objet d'une collecte d'information, d'une analyse, de l'évaluation du niveau d'adéquation et de l'importance reliés à chaque élément, et ce par le biais d'un questionnaire appuyé d'un guide d'évaluation. La démarche guide ensuite l'acquéreur dans une analyse simple faisant intervenir

plusieurs dimensions et faisant ressortir les caractéristiques les plus importantes menant à la décision. Bien qu'ils apparaissent dans le questionnaire selon un ordre séquentiel, les éléments d'évaluation peuvent être utilisés dans n'importe quel ordre conférant ainsi une grande flexibilité à l'outil, lui permettant de s'adapter aux différentes situations d'acquisitions de données.

La démarche et l'outil proposés seront appelés à être modifiés et améliorés suite aux rétroactions reçues lors de leur utilisation. Par exemple, une amélioration potentielle est la conversion de l'outil en format électronique. De plus, le cadre pourrait s'étendre afin d'intégrer l'évaluation aux phases d'acquisition, d'utilisation et d'utilisation répétée.

Remerciements

Les auteurs aimeraient remercier Hélène Bérard, Martin Lessard, Laurie Reedman et Julie Trépanier pour leur participation aux fructueuses discussions tenues au cours de ce projet. Ils remercient également Hélène Bérard et François Brisebois des commentaires constructifs lors de la révision de cet article.

Bibliographie

- Australian Bureau of Statistics (2009), « ABS Data Quality Framework », Australia : Australian Bureau of Statistics.
- Berning, M., Iwig, W., Marck, P. et Prell, M. (2013), « Quality Assessment Tool for Administrative data », Document de travail, Federal Committee on Statistical Methodology, United State.
- Brackstone, G. (1987), « Utilisation des dossiers administratifs à des fins statistiques », *Techniques d'enquête*, vol. 13, n° 1, p. 35-51.
- Daas, P., Ossen, S., Vis-Visschers, R. and Arends-Tóth, J. (2009), « Checklist for the Quality Evaluation of Administrative Data Sources », Discussion paper 09042, The Hague/Heerlen, Netherlands: Statistics Netherlands.
- Lavallée, P. (2007), « Utilisation des données administratives dans le cadre des statistiques sociales : Portraits actuel et futur », rapport non publié, Ottawa, Canada: Statistique Canada.
- Statistique Canada. (2002), « Le cadre de l'assurance de la qualité de Statistique Canada », No 12-585-NIF au catalogue, Ottawa, Canada : Statistique Canada.
- Statistique Canada. (2015a), « Guide d'évaluation des données administratives : Phases d'exploration et d'acquisition initiale », Document non publié, Ottawa, Canada : Statistique Canada.
- Statistique Canada. (2015b), « Questionnaire d'évaluation du processus d'acquisition de données administratives », Document non publié, Ottawa, Canada : Statistique Canada.