

Enquête québécoise sur la vie 10 ans après un diagnostic de cancer, 2024

Méthodologie de l'enquête



Pour tout renseignement concernant l'Institut de la statistique du Québec (ISQ) et les données statistiques dont il dispose, s'adresser à :

Institut de la statistique du Québec
200, chemin Sainte-Foy, 3^e étage
Québec (Québec) G1R 5T4

Téléphone :
418 691-2401
1 800 463-4090 (Canada et États-Unis)

Site Web : statistique.quebec.ca

Ce document est disponible seulement en version électronique.

Dépôt légal
Bibliothèque et Archives nationales du Québec
4^e trimestre 2025
ISBN 978-2-555-02602-5 (en ligne)

© Gouvernement du Québec, Institut de la statistique du Québec, 2025

Toute reproduction autre qu'à des fins de consultation personnelle est interdite sans l'autorisation du gouvernement du Québec.
statistique.quebec.ca/fr/institut/nous-joindre/droits-auteur-permission-reproduction

Novembre 2025

Publication réalisée à l'Institut de la statistique du Québec par :	Leila Nombo et Luc Côté
Sous la coordination de :	Marie-Eve Tremblay
Sous la direction de :	Éric Gagnon
Révision linguistique et édition :	Direction de la diffusion et des communications
Comité de lecture interne :	Jasline Flores, Daniela Gonzalez-Sicilia et Chantale Lecours
Enquête sous la responsabilité de :	Direction des enquêtes de santé Institut de la statistique du Québec
Enquête financée par :	Ministère de la Santé et des Services sociaux du Québec
Pour tout renseignement concernant le contenu de cette publication :	Direction de la méthodologie Institut de la statistique du Québec 200, chemin Sainte-Foy, 3 ^e étage Québec (Québec) G1R 5T4 Téléphone : 418 691-2410 1 800 463-4090 (Canada et États-Unis) Site Web : statistique.quebec.ca

Notice bibliographique suggérée

NOMBO, Leila, et Luc CÔTÉ (2025). *Enquête québécoise sur la vie 10 ans après un diagnostic de cancer 2024. Méthodologie de l'enquête*, [En ligne], Québec, Institut de la statistique du Québec, 27 p. [statistique.quebec.ca/fr/fichier/vie-10-ans-apres-diagnostic-cancer-2024-methodologie.pdf].

Avertissement

Les proportions de ce rapport sont arrondies à une décimale dans les tableaux et à l'unité dans le texte. Exceptionnellement, dans le texte, les proportions inférieures à 5 % et celles qui ne sont pas reportées dans les tableaux sont présentées avec une décimale. Les proportions dont la décimale est ,5 sont arrondies à l'unité inférieure ou supérieure selon la seconde décimale. En raison de l'arrondissement des données, le total ne correspond pas nécessairement à la somme des parties.

Table des matières

Introduction	5
1 Plan d'échantillonnage	6
1.1 Population visée	6
1.2 Base de sondage	7
1.3 Méthode de sélection, taille et répartition de l'échantillon	8
2 Collecte de données	10
2.1 Prétest	10
2.2 Méthodes de collecte	11
2.3 Résultats de la collecte	13
3 Traitement des données	15
3.1 Validation et saisie	15
3.2 Pondération	16
3.3 Non-réponse totale	18
3.4 Non-réponse partielle	18
4 Analyse des données, précision et tests statistiques	20
4.1 Précision des estimations et tests statistiques	20
4.2 Indicateur du niveau de revenu du ménage	21
5 Présentation des résultats	22
6 Portée et limites de l'enquête	23
Annexe	25
Références bibliographiques	26

Introduction

Pour que les données et les résultats issus de l'*Enquête québécoise sur la vie 10 ans après un diagnostic de cancer, 2024* soient utilisés adéquatement, il importe de connaître la méthodologie d'enquête utilisée. En effet, le plan de sondage, les procédures de collecte et le traitement des données sont tous des éléments qui ont une incidence sur les résultats d'une enquête. La connaissance des aspects méthodologiques aide à interpréter correctement les résultats et à en apprécier la qualité, la portée et les limites.

Objectifs de l'enquête

L'*Enquête québécoise sur la vie 10 ans après un diagnostic de cancer, 2024* est réalisée, pour la toute première fois, à la demande du ministère de la Santé et des Services sociaux. L'enquête, à portée provinciale, vise principalement à décrire l'état de santé physique et émotionnelle, les conditions et la qualité de vie ainsi que l'utilisation de soins et de services de santé des personnes 10 ans après qu'elles ont reçu un diagnostic de cancer au Québec.

Structure du rapport méthodologique

Les six sections du présent document portent sur les principaux éléments de la méthodologie de l'enquête. La section 1 présente le plan d'échantillonnage de l'enquête, tandis que la section 2 décrit les stratégies utilisées et les résultats obtenus lors de la collecte de données. La section 3 porte sur le traitement des données ; on y aborde le processus de validation, la méthode de pondération utilisée pour que les résultats puissent être inférés à la population visée de même que l'examen de l'ampleur de la non-réponse partielle et le recours à l'imputation des données pour en atténuer l'effet sur les analyses. La section 4 traite des méthodes utilisées pour l'analyse des données, de l'estimation de la précision et des tests statistiques. Les normes de présentation des résultats sont précisées à la section 5, alors qu'un aperçu de la portée et des limites de l'enquête est offert à la section 6.

Ce document accompagne le rapport de résultats de l'enquête intitulé *Enquête québécoise sur la vie 10 ans après un diagnostic de cancer 2024*, disponible sur le site Web de l'Institut de la statistique du Québec (ISQ).

1 Plan d'échantillonnage

Cette section comprend une description de la population visée et de la base de sondage ainsi que toutes les informations utiles sur la sélection de l'échantillon à partir de cette base.

1.1 Population visée

La population visée par l'*Enquête québécoise sur la vie 10 ans après un diagnostic de cancer, 2024* correspond à l'ensemble des personnes ayant reçu, pour la première fois, un diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014 et étant âgées de 18 ans et plus au moment du diagnostic. Le diagnostic de cancer pouvait être établi en dehors de la province du Québec. Toutefois, seules les personnes toujours vivantes résidant au Québec au moment du diagnostic et au moment de l'enquête sont concernées.

Sont donc exclues :

- les personnes vivant dans un logement collectif institutionnel au moment de l'enquête (hôpital, centre d'hébergement de soins de longue durée, établissement pénitentiaire, centre de réadaptation, etc.) ;
- les personnes décédées au moment de l'enquête ;
- les personnes dont la tumeur n'était pas maligne – siège primaire (code de comportement = 3) –, et n'était pas un carcinome in situ (code de comportement = 2) au moment du diagnostic ;¹
- les personnes vivant hors Québec pour une période prolongée ou définitive, identifiées par la Régie de l'assurance maladie du Québec (RAMQ) ou au moment de l'enquête ;
- les personnes inadmissibles à l'assurance maladie dont la carte est expirée ou les personnes n'ayant pas d'adresse, selon la RAMQ au moment de l'enquête ;
- les personnes qui ont déclaré ne pas avoir reçu de diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014 au moment de l'enquête.

Le nombre total de personnes dans la population visée par l'enquête est estimé à 38 907 personnes². Étant donné que la population visée concerne les personnes de 18 ans et plus au moment du diagnostic en 2013-2014, les personnes étaient par conséquent âgées de 27 ans ou plus au moment de l'enquête, dépendamment de la date de naissance, de la date exacte du diagnostic et de la date de remplissage du questionnaire.

1. Voir plus de détails à la section 1.2 sur la description de la base de sondage.

2. Ce sont les totaux de population, soit les personnes ayant eu un premier diagnostic de cancer entre 2013 et 2014, utilisés pour le calage aux marges des données de l'enquête (section 3.2).

1.2 Base de sondage

Le Registre québécois du cancer (RQC) a été retenu comme fichier source de données nécessaires pour la construction de la base de sondage. Le RQC est alimenté par plusieurs sources d'information : les registres locaux de cancer (les rapports de pathologie, le dossier médical hospitalier et le dossier médical clinique externe), le fichier des hospitalisations (MED-ÉCHO), le fichier des décès, les données provenant des registres des autres provinces (cas hors Québec) et certains fichiers de la RAMQ (Direction générale de la santé publique 2012).

Le RQC contient des renseignements médicaux tels que les codes de topographie (ou siège primaire de la tumeur), d'histologie (description microscopique des cellules ou des tissus de la tumeur) et de comportement (niveau de prolifération de la tumeur), la date de diagnostic de même que des renseignements personnels, comme les noms et prénoms, la date de naissance, le sexe, le numéro d'assurance maladie et la date de décès (MSSS 2019).

La constitution de la base de sondage a nécessité trois grandes étapes. Lors de la première, toutes les personnes se trouvant dans le RQC et respectant les critères de la période de référence visée quant à la date de diagnostic³ et du code de comportement⁴ ont été retenues. À cette étape, on dénombrait 87 283 personnes.

La deuxième étape a permis de vérifier si les personnes retenues après la première étape étaient encore en vie. Pour ce faire, deux sources différentes de données sur les décès ont été utilisées : la date de décès selon le RQC et la date de décès selon le Registre des événements démographiques (RED) de l'ISQ. Au total, ces deux sources ont permis d'identifier 47 856 personnes décédées⁵. Après cette deuxième étape, la base de sondage dénombrait 39 427 personnes.

Une dernière étape a permis d'identifier les personnes qui n'habitent plus au Québec afin de les exclure. Pour ce faire, un jumelage avec le Fichier d'inscription des personnes assurées (FIPA) de la RAMQ a été nécessaire. Après avoir exclu les personnes déménagées selon le FIPA, on a obtenu une base de sondage finale (contenant l'ensemble de la population admissible), laquelle comptait 38 907 personnes au 12 avril 2024.

Lors de l'élaboration de la base de sondage, le jumelage avec le FIPA présente plusieurs avantages. En effet, le FIPA permet l'identification et l'exclusion de la majeure partie des personnes non visées par l'enquête, notamment les personnes déménagées. Il constitue aussi une excellente source d'information pour l'obtention des coordonnées postales. D'autres renseignements nécessaires pour communiquer avec les personnes sélectionnées sont également disponibles au sein de ce fichier, tels que le sexe⁶, la langue de correspondance et les numéros de téléphone. Comme principal inconvénient de l'utilisation du FIPA, on note la qualité de la mise à jour des numéros de téléphone, qui, dans certains cas, sont inexacts et parfois même absents. Ce qui complique du même coup la collecte de données. Une fois la base de sondage constituée, on a constaté que cette situation s'observe fréquemment chez les personnes âgées de 75 ans et plus (13 %) ainsi que chez les personnes atteintes d'un cancer de la prostate (14 %) ou d'un cancer colorectal (12 %). La base de sondage pourrait compter quelques personnes inadmissibles à l'enquête, parce qu'elles n'auraient pas informé la RAMQ d'un déménagement à l'extérieur du Québec, par exemple.

3. Entre le 1^{er} janvier 2013 et le 31 décembre 2014.

4. Seules les personnes dont le code de comportement = 2 (carcinome in situ) ou 3 (tumeur maligne – siège primaire) sont visées.

5. Le RQC avait identifié 43 844 décès jusqu'au 3 février 2022. Le RED en a identifié 4 012 supplémentaires jusqu'au 12 avril 2024.

6. Pour les personnes transgenres ayant obtenu un changement de la mention du sexe figurant à leur acte de naissance auprès du Directeur de l'état civil, le sexe indiqué dans le FIPA est celui indiqué dans l'acte de naissance après le changement.

Dans cette enquête, il est difficile de quantifier de manière précise la couverture nette de la base de sondage, étant donné que la taille de la population visée ne peut être établie de façon exacte. Les meilleures données disponibles pour estimer cette taille sont celles résultant du jumelage du RQC avec le RED et le FIPA pour identifier les personnes toujours vivantes et demeurant au Québec. Toutefois, comme on a constaté lors de diverses enquêtes sur le cancer faites par l'ISQ dans le passé, la base de sondage comprend des personnes inadmissibles à l'enquête, notamment celles qui déclarent n'avoir jamais eu de cancer⁷. Cela contribue à complexifier la quantification précise de la couverture nette de la base de sondage.

1.3 Méthode de sélection, taille et répartition de l'échantillon

L'objectif général de l'enquête est de fournir des estimations fiables sur différents aspects entourant l'état de santé physique et émotionnelle, les conditions et la qualité de vie, et l'utilisation de soins et services de santé à l'échelle du Québec. Pour ce faire, un scénario avec une taille d'échantillon de 20 000 personnes a été retenu. Avec une telle taille, le plan de sondage de cette enquête a donc été construit de manière à⁸ :

- estimer une proportion de 0,5 % et plus avec une bonne précision relative, c'est-à-dire un coefficient de variation (CV)⁹ d'au plus 15 % à l'échelle provinciale ;
- estimer avec un CV d'au plus 15 %, une proportion minimale de 2,6 % parmi les personnes ayant reçu un diagnostic de cancer du sein, de 1,9 % parmi les personnes ayant reçu un diagnostic de cancer de la prostate et de 3,5 % parmi les personnes ayant reçu un diagnostic de cancer colorectal de manière que cette estimation de proportion soit associée à un CV d'au plus 15 %.

Des hypothèses quant aux taux d'admissibilité et de réponse ont été posées afin d'estimer le nombre attendu de répondants et répondantes pour atteindre ces objectifs de précision.

Plus précisément, nous avons présumé que nous aurions :

- un taux de réponse de 57 % au provincial ;
- un taux d'admissibilité à l'enquête de 87 %¹⁰ ;
- un effet de plan provincial de 1,1.

7. Les éditions de 2008 et de 2013 de l'*Enquête québécoise sur la qualité des services de lutte contre le cancer de même que l'Enquête sur l'expérience des patients atteints de cancer et en transition 2016*, toutes trois menées par l'ISQ, avaient estimé entre 6 % et 9 % le taux d'inadmissibilité déclaré, même si les registres administratifs confirmaient la présence du cancer.

8. Dans la mesure où les hypothèses de collecte (taux de réponse et taux d'admissibilité) sont valides.

9. Des renseignements sur le calcul du coefficient de variation sont fournis à la section 4.1.1.

10. Il y a deux grands types d'inadmissibilité qui peuvent survenir lors de l'enquête. Le premier type touche les personnes qui, au moment de la collecte, étaient décédées, étaient en logement collectif institutionnel ou avaient déménagé hors du Québec. On estime que ce taux d'inadmissibilité est de 1 % ou, autrement dit, que le taux d'admissibilité est égal à 99 %. Le deuxième type correspond à la population de personnes déclarant ne pas avoir reçu un diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014. Ce taux d'inadmissibilité est estimé à 12 %. L'hypothèse d'un taux d'admissibilité à l'enquête de 87 % est obtenue par multiplication des taux propres à ces deux types d'inadmissibilité $\{(100 \% - 1 \%) \times (100 \% - 12 \%)$.

Selon ces hypothèses, le nombre attendu de personnes répondantes est estimé à 9 918. Le tableau 1.1 fournit la taille de population et d'échantillon ainsi que le nombre attendu de personnes répondantes selon diverses caractéristiques.

Tableau 1.1

Taille de la population visée, taille de l'échantillon et nombre attendu de personnes répondantes, Québec, 2025

Caractéristiques	Taille de la population visée	Taille de l'échantillon	Nombre attendu de personnes répondantes
	n		
Âge au 27 mai 2024 ¹			
49 ans et moins	2 845	1 473	730
50-59 ans	3 921	2 009	996
60-64 ans	4 337	2 235	1 108
65-69 ans	5 814	2 986	1 481
70-74 ans	6 467	3 328	1 650
75-79 ans	6 887	3 530	1 751
80-84 ans	4 767	2 441	1 210
85 ans et plus	3 869	1 998	991
Siège du cancer ²			
Sein	8 455	4 343	2 154
Prostate	7 200	3 698	1 834
Colorectal	4 623	2 377	1 179
Mélanome	1 699	872	432
Hématopoiétique	3 001	1 546	767
Autres cancers	13 929	7 164	3 553
Sexe ³			
Femme	20 628	10 600	5 257
Homme	18 279	9 400	4 661
Total	38 907	20 000	9 918

1. Il s'agit de la catégorie d'âge selon le FIPA au premier jour de collecte.

2. Il s'agit du siège au moment du diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014.

3. Il s'agit du sexe selon le FIPA.

Source : Institut de la statistique du Québec, *Enquête québécoise sur la vie 10 ans après un diagnostic de cancer, 2024*.

Le tirage de l'échantillon s'est fait selon une stratification de la base de sondage visant à assurer un certain contrôle de la composition de l'échantillon. Les strates ont été formées par le croisement entre le siège du cancer au moment du diagnostic, le sexe et la structure d'âge. Au total, 59 strates ont été construites. Lors de la construction de l'échantillon, aucun objectif précis n'était déterminé pour obtenir des résultats justes et détaillés pour tous les types de sièges. L'échantillon a été sélectionné aléatoirement et indépendamment par strate, proportionnellement à la taille des strates.

Ainsi, ce plan d'échantillonnage ne permet pas de réaliser une analyse statistique pour chaque type de siège pris individuellement. Des estimations sont envisageables uniquement pour certaines sous-populations associées aux cancers les plus prévalents, comme celles des personnes ayant reçu un diagnostic de cancer de la prostate, du sein, colorectal ou de type autre. En revanche, selon ce plan de sondage, on ne peut produire des estimations fiables pour les sous-populations atteintes de cancers moins prévalents, comme le mélanome ou les cancers hématopoiétiques. Par ailleurs, même pour les cancers plus prévalents, les possibilités de croisement avec d'autres variables sont déconseillées, car la précision des estimations serait très faible.

2

Collecte de données

Cette section porte sur les méthodes et les résultats de la collecte de données. On y présente notamment les instruments et le mode de collecte ainsi que les taux de réponse obtenus.

2.1 Prétest

Du 1^{er} au 15 février 2024, un prétest a été effectué auprès de 222 personnes répondantes à partir d'un échantillon de 500 personnes. Le taux de réponse s'est élevé à 47 %. Les entrevues ont été réalisées au moyen d'interviews téléphoniques assistées par ordinateur (ITAO).

Le prétest visait à estimer le temps requis pour répondre aux différentes sections au téléphone. Il avait également comme objectif de vérifier l'intelligibilité du questionnaire (notamment les instructions aux participants ou participantes et au personnel intervieweur) ainsi que la capacité des participants et participantes à fournir l'information demandée, d'identifier des questions à modifier ou à éliminer et de tester la compréhension des questions en français et en anglais. Enfin, de l'information sur certains paramètres était requise (taux de réponse, taux de refus et taux d'admissibilité) afin de peaufiner le plan de sondage de l'enquête et les paramètres servant à la détermination de la taille d'échantillon.

L'analyse des résultats du prétest a conduit à des ajustements du questionnaire. Compte tenu de ces modifications, les données fournies par les répondantes et répondants au prétest n'ont pas été conservées pour les estimations produites en fin d'enquête. Quelques personnes échantillonnées au prétest et à l'enquête, qui n'avaient pas répondu au prétest ni refusé de participer, ont été à nouveau invitées à répondre lors de l'enquête.

2.2 Méthodes de collecte

Instruments de collecte

Le questionnaire de l'*Enquête québécoise sur la vie 10 ans après un diagnostic de cancer 2024* a été élaboré par l'ISQ en collaboration avec les membres du Comité d'orientation de projet (COP). Voici les principales sources utilisées pour l'élaboration du questionnaire :

QLACS, 2005	Quality of Life in Adult Cancer Survivors
EQSLCC, 2013	Enquête québécoise sur la qualité des services de lutte contre le cancer (ISQ)
NHS, 2012	National Health Service England - Living with and beyond cancer (NHS)
EPCT, 2016	Enquête sur l'expérience des patients atteints de cancer en transition, volet québécois
CARES, 1998	Cancer Rehabilitation Evaluating System
NHIS, 2010	National Health Interview Survey - Cancer Control Supplement (CDC)
EQES, 2011-2012	Enquête québécoise sur l'expérience des soins (ISQ)
EQSP, 2020-2021	Enquête québécoise sur la santé de la population, 2020-2021
EQAPS, 2018-2019	Enquête québécoise sur l'activité physique et le sport 2018-2019
LIVESTRONG, 2010	Lance Armstrong Foundation

Si on exclut l'introduction et les remerciements, le questionnaire compte quatre sections. Les thèmes abordés sont les suivants :

Section 1 : État de santé ;

Section 2 : Conditions et qualité de vie ;

Section 3 : Utilisation des soins et services de santé ;

Section 4 : Caractéristiques sociodémographiques et économiques.

La première section inclut à la fois les questions sur l'état de santé physique actuel et les séquelles physiques des personnes de même que les questions sur l'état de santé émotionnelle actuel et les séquelles psychosociales découlant du cancer et de son traitement, 10 ans après son diagnostic. La deuxième section traite des conséquences du cancer et de ses traitements sur les besoins d'aide et de soutien ainsi que sur les habitudes de vie (ex. : la pratique d'activité physique). Elle traite aussi des répercussions du cancer et de ses traitements sur les finances, la vie professionnelle et la vie familiale. La troisième section porte sur les soins et services de santé auxquels les personnes ont eu recours, ainsi que sur les traitements de médecine douce utilisés en raison du cancer et les sources d'information consultées. La quatrième section comporte des questions sur les caractéristiques sociodémographiques des personnes ayant participé à l'enquête.

Période de collecte

La collecte des données a eu lieu du 27 mai 2024 au 24 septembre 2024. Pour faciliter la gestion de la collecte des données, on a scindé l'échantillon de l'enquête en deux lots afin de ne pas envoyer sur le terrain 20 000 lettres en même temps. La collecte du premier lot a débuté le 27 mai 2024, alors que celle du deuxième lot a commencé une semaine plus tard, soit le 3 juin 2024. La durée moyenne des entrevues sur le Web est estimée à environ 21 minutes, tandis que celle des entrevues téléphoniques est évaluée à environ 31 minutes.

Modes de collecte

La collecte de données a été effectuée par ITAO et/ou par interview Web assistée par ordinateur (IWAO). Il s'agit donc d'une collecte multimode. Ces modes de collecte présentent des avantages pour ce qui est de la qualité des données. Premièrement, le libellé des questions, y compris les périodes de référence et les pronoms, est personnalisé automatiquement en fonction de facteurs comme l'âge et le genre de la personne, et des réponses aux questions précédentes. Deuxièmement, des mesures de contrôle isolent les réponses incohérentes ou hors norme, et des instructions apparaissent à l'écran lorsqu'une telle situation se présente. La personne remplissant le questionnaire reçoit une rétroaction immédiate, et une correction de l'incohérence peut être apportée. Troisièmement, le processus fait que les questions qui ne concernent pas une personne sont automatiquement omises.

Stratégies de collecte

Une formation a été donnée à l'équipe d'intervieweurs et d'intervieweuses responsables de la collecte téléphonique de l'enquête. Cette équipe a été formée non seulement pour mener des entrevues, mais aussi pour effectuer des relances et faire le suivi des dossiers. Étant donné le caractère délicat et sensible de l'étude, l'équipe a aussi été particulièrement formée pour mieux s'adapter à diverses situations susceptibles de survenir lors du déroulement de l'interview (détresse, tristesse, anxiété, colère, etc.).

Une lettre de présentation de l'enquête a d'abord été envoyée aux personnes sélectionnées pour leur expliquer les objectifs de l'enquête et les inviter à se rendre sur le site Web pour remplir le questionnaire électronique ou à appeler l'ISQ pour le faire avec un intervieweur ou une intervieweuse. Deux relances postales ont ensuite été envoyées à intervalles d'environ deux semaines aux personnes n'ayant toujours pas répondu. Si le questionnaire n'était pas encore rempli environ six semaines après l'envoi de la lettre initiale, des appels téléphoniques ont été faits pour encourager les personnes n'ayant toujours pas rempli le questionnaire à le faire. Une entrevue téléphonique était proposée aux personnes préférant ne pas répondre en ligne ou n'ayant pas accès à Internet. Cette enquête a reçu un fort volume d'appels entrants à l'ISQ tout au long de la collecte qui s'est terminée le 24 septembre 2024.

2.3 Résultats de la collecte

À l'instar des taux de réponse généralement présentés par l'ISQ, ceux de l'enquête sont pondérés¹¹. Ils tiennent compte, entre autres, de la propension à répondre au questionnaire, qui varie selon diverses caractéristiques (voir section 3.2). Au total, 13 070 personnes ont participé à l'enquête, ce qui correspond à un taux de réponse pondéré de 73 % (tableau 2.1). Ce taux de réponse est beaucoup plus élevé que la cible visée de 57 %. On tire également ce constat selon le groupe d'âge à l'exception des personnes de 85 ans et plus dont le taux de réponse est légèrement inférieur à la cible et égal à 55 %. Le taux de réponse par siège est également supérieur à la cible et se situe entre 70 %, qui correspond à celui du siège colorectal, et 76 %, qui correspond à celui du sein.

Tableau 2.1

Nombre de personnes répondantes, taux de réponse pondéré, taux d'admissibilité pondéré associé à la déclaration de ne pas avoir eu de cancer et taux d'admissibilité pondéré associé au décès, déménagement hors-Québec ou vivant dans un logement collectif institutionnel, selon l'âge, le sexe et le siège de cancer, Québec, 2025

Caractéristiques	Nombre de personnes répondantes	Taux de réponse pondéré	Taux d'admissibilité pondéré	
			Personnes ayant déclaré ne pas avoir eu de cancer	Personnes vivant dans un logement collectif institutionnel ou déménagées hors-Québec ou décédées
			n	%
Âge au 27 mai 2024 ¹				
49 ans et moins	811	59,9	94,0	99,2
50-59 ans	1 288	67,5	95,9	99,5
60-64 ans	1 592	75,3	95,7	99,0
65-69 ans	2 221	79,6	95,0	98,5
70-74 ans	2 483	80,9	94,0	98,1
75-79 ans	2 445	76,6	92,7	97,4
80-84 ans	1 393	66,8	90,8	93,9
85 ans et plus	837	55,2	86,9	85,9
Siège du cancer ²				
Sein	3 188	75,8	99,3	98,1
Prostate	2 459	75,0	92,1	96,6
Colorectal	1 436	69,8	91,2	95,5
Mélanome	539	72,6	86,7	97,9
Hématopoïétique	1 002	72,2	93,4	97,0
Autres cancers	4 446	70,0	92,0	96,8
Sexe ³				
Femme	7 124	73,1	95,0	97,6
Homme	5 946	71,8	91,7	96,3
Total	13 070	72,5	93,5	97,0

1. Il s'agit de la catégorie d'âge selon le FIPA au premier jour de collecte.

2. Il s'agit du siège du cancer au moment du diagnostic reçu entre le 1^{er} janvier 2013 et le 31 décembre 2014.

3. Il s'agit du sexe selon le FIPA.

Source : Institut de la statistique du Québec, *Enquête québécoise sur la vie 10 ans après un diagnostic de cancer*, 2024.

11. Plus de renseignements sur la pondération sont fournis à la section 3.2.

Le refus de certaines personnes de participer à l'enquête a contribué de façon importante à la non-réponse : en effet, 22 % des personnes non répondantes tirent leur statut du fait qu'ils ont refusé d'y participer. Une deuxième raison, soit un mauvais numéro de téléphone après recherche, contribue aussi à la non-réponse. Celle-ci est à l'origine du statut de 21 % des personnes non répondantes.

Environ 7 % des personnes sélectionnées ont déclaré ne pas avoir eu de cancer pendant la période visée par l'enquête (du 1^{er} janvier 2013 au 31 décembre 2014), ce qui représente un taux plus faible que celui prévu. Ce taux varie selon le sexe et l'âge. Plus précisément, on estime à 8 % ce taux d'inadmissibilité chez les hommes et à 5 % chez les femmes. Selon l'âge, il est de 4,9 % chez les 27-59 ans, de 4,7 % chez les 60-69 ans, de 7 % chez les 70-79 ans et de 11 % chez les 80 ans et plus. On a également exclu quelques personnes qui, au moment de la collecte, avaient déménagé dans un logement collectif institutionnel ou à l'extérieur du Québec ou qui étaient décédées. Ces motifs d'exclusion sont les autres sources d'inadmissibilité à l'enquête¹². Lorsqu'on tient compte de ces deux types d'inadmissibilité, on constate que le taux d'admissibilité pondéré à l'enquête de 91 % demeure plus élevé que celui émis comme hypothèse (87 %).

Finalement, environ 63 % des répondants et répondantes ont opté pour le questionnaire en ligne.

12. Le taux d'inadmissibilité de ces autres sources s'élève à 3,0 %.

3

Traitement des données

Dans cette section, on aborde les étapes de traitement des données recueillies. Elle traite de la validation de la qualité des données, de la pondération nécessaire à l'inférence à la population visée ainsi que de l'examen de la non-réponse totale et partielle et des traitements d'imputation effectués.

3.1 Validation et saisie

Une validation du statut de réponse des personnes sélectionnées (répondantes, non répondantes ou inadmissibles) a d'abord été effectuée. Cela a permis de considérer comme répondantes certaines personnes qui n'avaient pas rempli le questionnaire en entier et de rejeter certains questionnaires comportant une importante proportion de questions sans réponse. Des questions posées en début d'entrevue visaient à s'assurer que le questionnaire était bel et bien rempli par la personne sélectionnée. Des validations ont été effectuées en cours de collecte et a posteriori pour comparer les réponses fournies aux renseignements contenus dans la base de sondage, entre autres la date de naissance et le code postal du lieu de résidence. La concordance étant exacte à quelques exceptions près, rien n'indiquait que certains questionnaires devaient être rejetés sur la base de ces comparaisons.

Étant entièrement informatisées, les collectes téléphoniques et sur le Web permettent d'effectuer plusieurs validations de base pendant l'entrevue, notamment une vérification du respect des choix de réponse ou de l'adéquation des sauts de section, pour les blocs de questions ne s'adressant qu'à un sous-groupe de personnes répondantes. Elles permettent aussi une saisie automatique des réponses fournies par les répondants et répondantes. La très grande majorité des questions sont donc ainsi autocodées, c'est-à-dire que l'intervieweur ou intervieweuse ou la personne répondante n'a qu'à cocher le ou les choix de réponse qui s'appliquent : la saisie manuelle de réponse s'en trouve donc réduite au minimum. Une validation a posteriori a été effectuée afin d'assurer la cohérence des réponses fournies par une même personne, ce qui a permis de relever quelques erreurs qui ont été corrigées. Par exemple, certaines réponses fournies à des questions offrant le choix de réponse « autre, précisez » ont été recodées lorsque la réponse correspondait à une catégorie déjà offerte. Lorsque ceci n'était pas possible ou lorsqu'une réponse donnée était trop vague pour être recodée avec certitude, la réponse demeurait dans une catégorie « autres ». Enfin, les réponses ont été colligées de manière à distinguer les personnes ayant omis volontairement de répondre à une question de celles qui n'étaient pas concernées par cette question.

3.2 Pondération

La pondération est essentielle pour la production des résultats de l'enquête. Elle permet de faire des inférences adéquates à la population visée, bien que celle-ci n'ait pas été sondée dans sa totalité. Elle consiste à attribuer un poids statistique à chaque personne répondante. Ce poids correspond au nombre de personnes qu'elle « représente » au sein de la population visée. Il doit tenir compte, entre autres, de la probabilité de sélection de la personne prédéterminée par le plan d'échantillonnage, de l'admissibilité et de la non-réponse à l'enquête. En effet, il est connu que, dans les enquêtes, la probabilité de répondre varie selon plusieurs caractéristiques. Cette enquête ne fait pas exception, et on y note également des différences selon certains profils quant à l'admissibilité à l'enquête. Il est donc nécessaire de tenir compte de ces éléments en les intégrant à la pondération.

De façon plus détaillée, le poids initial de chaque personne faisant partie de l'échantillon de l'enquête a été déterminé en fonction de l'inverse de sa probabilité de sélection. Ensuite, un ajustement a été fait pour tenir compte de l'admissibilité à l'enquête. Le statut d'admissibilité de la plupart des personnes non répondantes étant inconnu, il fallait réduire leur poids afin de refléter le fait qu'une partie d'entre elles étaient probablement inadmissibles. Dans cette enquête, l'ajustement pour l'admissibilité s'est déroulé en deux phases, puisqu'il y a deux types d'inadmissibilité à l'enquête. La première phase visait à faire un ajustement des poids pour tenir compte de l'inadmissibilité causée par le fait que des personnes sont décédées, ont déménagé hors du Québec ou résidaient dans un logement collectif institutionnel au moment de l'enquête. Pour cet ajustement, la méthode du score de propension à répondre a été employée (Eltinge et Yansaneh 1997 ; Haziza et Beaumont 2007). Cette méthode consiste à modéliser le fait d'être admissible ou non à l'enquête et à créer ensuite des classes de pondération correspondant à des groupes d'individus ayant des caractéristiques et des proportions de personnes admissibles semblables. Dans cette enquête, la proportion de personnes inadmissibles varie selon le groupe d'âge et le sexe. Par conséquent, l'ajustement a été fait selon ces deux variables. Un taux d'admissibilité est calculé pour chacune des classes formées à partir de cette méthode, ce qui permet l'ajustement du poids des personnes non répondantes dont le statut du premier type d'admissibilité est inconnu.

La seconde phase visait à faire un ajustement des poids pour tenir compte de l'autre type d'inadmissibilité, soit le fait de déclarer ne pas avoir eu de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014. Pour cet ajustement, la méthode du score de propension à répondre est utilisée à nouveau. À l'instar de la première phase, cette méthode consistait à modéliser le fait d'être admissible ou non à l'enquête et à créer ensuite des classes de pondération correspondant à des groupes d'individus ayant des caractéristiques et des proportions de personnes admissibles semblables¹³. Un taux d'admissibilité est calculé pour chacune des classes formées à partir de cette méthode, ce qui permet l'ajustement des poids des personnes non répondantes dont le statut de l'autre type d'admissibilité est inconnu. La différence entre la somme des poids initiaux des personnes non répondantes dont le statut d'admissibilité est inconnu et la somme de leurs poids ajustés à la baisse à la seconde phase du traitement pour l'inadmissibilité est ensuite redistribuée aux personnes inadmissibles ayant déclaré ne pas avoir eu de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014, selon chacune des classes de pondération formées. Le poids de ces personnes est ainsi ajusté à la hausse. Cette manipulation est nécessaire dans cette enquête, puisqu'il n'y a aucune estimation connue du nombre de personnes de 18 ans et plus au moment du diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014 vivant au Québec et qui auraient déclarées ne pas avoir eu de cancer durant cette période. Ainsi, ces personnes inadmissibles doivent être conservées dans

13. Pour ce deuxième ajustement, les variables (qui sont issues du RQC ou du FIPA ou créées à partir de ces derniers) retenues dans le modèle sont : l'âge en huit catégories, le siège du cancer et la langue de correspondance au FIPA.

le fichier jusqu'à la dernière étape de la pondération, où elles représenteront toutes les personnes de 18 ans et plus au moment du diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014 vivant au Québec et qui seraient inadmissibles à l'enquête, car elles auraient déclaré ne pas avoir eu de cancer durant cette période¹⁴.

L'étape suivante est l'ajustement pour la non-réponse totale. Lors de cette étape, le poids des répondants et répondantes est augmenté afin de représenter également les non-répondants et non-répondantes. Une fois de plus, la méthode du score de propension à répondre est employée. Dans ce cas-ci, elle consiste à modéliser le fait d'avoir répondu ou non à l'enquête selon les renseignements disponibles dans la base de sondage ou créés à partir de cette dernière¹⁵. Des classes composées de personnes ayant des caractéristiques et une propension à répondre semblables, différentes des classes de pondération créées précédemment pour l'ajustement des deux types d'admissibilité, ont ainsi été formées. À l'intérieur de chaque classe, le poids des personnes répondantes a été ajusté par l'inverse du taux de réponse observé à l'enquête.

Par la suite, on a vérifié qu'aucune personne n'avait de poids très élevé comparativement aux poids des individus de même sexe et de même catégorie d'âge¹⁶, afin de s'assurer qu'une personne répondante n'exerce pas une influence indue sur les statistiques produites. Pour ce faire, une méthode appelée « écart-sigma » a été utilisée (Bernier et Nobrega 1998). Lorsque jugé trop élevé, le poids d'une personne est abaissé au poids inférieur le plus près dans le même croisement sexe-catégorie d'âge. Très peu de personnes répondantes ont eu un poids modifié à cette étape.

La dernière étape, soit le calage aux marges¹⁷, consiste à ajuster la pondération afin que la somme des poids des personnes répondantes et celles des personnes inadmissibles aux deux types correspondent bien aux effectifs connus de la population de l'ensemble des personnes adultes ayant reçu un diagnostic de cancer entre le 1^{er} janvier 2013 et le 31 décembre 2014¹⁸, et cela pour :

- chaque catégorie du siège du cancer (six catégories : sein, prostate, colorectal, mélanome, hématopoïétique, autres sièges de cancer) ;
- le croisement du sexe (deux catégories) et de la catégorie d'âge (sept catégories : 27-50 ans, 50-59 ans, 60-64 ans, 65-69 ans, 70-74 ans, 75-79 ans, 80 ans et plus).

Une fois cette dernière étape terminée, les poids de l'ensemble des personnes inadmissibles sont mis de côté. Il ne reste que les poids associés aux personnes répondantes. Ce sont ces poids, qu'on appelle poids populationnels, qui permettent de généraliser les résultats obtenus à partir de l'échantillon de personnes répondantes à l'ensemble de la population visée. La somme des poids populationnels constitue donc une estimation de la taille de la population visée.

14. Et elles ne remplissent pas les autres critères d'exclusion de la population visée.

15. Les variables retenues dans le modèle sont : le siège du cancer, l'indice de défavorisation matérielle et sociale (Azevedo Da Silva et autres 2024), l'âge en huit catégories, le sexe, la langue de correspondance, la présence ou non d'un conjoint et la présence ou non d'un numéro de téléphone.

16. L'âge a été subdivisé en huit catégories : moins de 50 ans, 50-59 ans, 60-64 ans, 65-69 ans, 70-74 ans, 75-79 ans, 80-84 ans, 85 ans et plus.

17. Le *calage aux marges* est une procédure que l'on peut appliquer pour incorporer des données auxiliaires. Cette procédure rajuste les poids d'échantillonnage au moyen de multiplicateurs appelés *facteurs de calage*, lesquels font correspondre les estimations aux totaux connus.

18. Ces effectifs sont dérivés des tailles de population fournies par le Registre québécois du cancer (12 avril 2024).

3.3 Non-réponse totale

Définition

La non-réponse totale survient lorsqu'une personne sélectionnée et admissible à l'enquête ne remplit pas le questionnaire pour diverses raisons. Cette non-réponse peut entraîner des biais dans les estimations si les personnes ayant répondu présentent des caractéristiques différentes de celles ne l'ayant pas fait, et que ces caractéristiques sont liées au sujet de l'enquête. La pondération est ajustée pour la non-réponse totale à l'enquête (section 3.2), ce qui permet de réduire le risque de biais dû à celle-ci. Toutefois, seules les informations contenues dans la base de sondage, donc connues pour l'ensemble de l'échantillon, peuvent être prises en considération pour cet ajustement. Ainsi, malgré l'utilisation de la pondération, des résultats de l'enquête peuvent quand même être biaisés si la non-réponse totale est importante et qu'elle est liée à une ou plusieurs caractéristiques non disponibles dans la base de sondage, et que ces caractéristiques sont fortement corrélées à certains indicateurs.

3.4 Non-réponse partielle

Définition

La non-réponse partielle désigne l'absence de réponse à une question pour certaines personnes ayant rempli le questionnaire due au fait qu'elles refusent de répondre à la question ou qu'elles ne connaissent pas la réponse à cette question. Il est connu qu'une non-réponse partielle importante peut entraîner certains biais dans les estimations, au même titre que la non-réponse totale, s'il s'avère que les personnes non répondantes présentent des caractéristiques différentes de celles des personnes répondantes et que ces caractéristiques sont de surcroît liées au thème étudié. La pondération tient compte de la non-réponse totale, mais n'a pas été ajustée pour la non-réponse partielle.

Taux de non-réponse partielle

Le taux de non-réponse partielle pondéré à une question est défini comme le rapport entre le nombre pondéré de personnes n'ayant pas répondu à celle-ci et le nombre pondéré de personnes admissibles à y répondre. Plus ce taux est élevé, plus le risque de biais induit par la non-réponse partielle est grand. On fait généralement l'hypothèse qu'une non-réponse partielle inférieure à 5 % a une incidence négligeable sur les estimations à l'échelle provinciale, sauf dans la situation où la non-réponse partielle est fortement caractérisée.

La non-réponse partielle étant peu élevée pour l'ensemble des questions et indicateurs de l'enquête, son incidence sur le risque de biais dans les résultats est faible. Les indicateurs ou les questions qui affichent une non-réponse partielle supérieure à 5 % (tableau A1 en annexe) concernent notamment :

- Le niveau de difficulté à avoir des enfants en raison du cancer ou de son traitement ;
- Le fait d'associer au moins un des problèmes de santé physique ou mentale au cancer ou à son traitement ;
- Le fait d'associer au moins un des problèmes de santé physique au cancer ou à son traitement ;
- Le niveau de revenu du ménage.

Pour les quatre indicateurs dont la non-réponse partielle est au-dessus de 5 %, les mesures prises pour en tenir compte sont les suivantes. Comme on l'explique à la section suivante, le revenu du ménage a été imputé pour toutes les personnes répondantes lorsqu'il était manquant. Les trois autres indicateurs et questions du questionnaire ont fait l'objet d'une analyse de la non-réponse partielle plus approfondie afin de déterminer si les personnes ayant omis de répondre présentaient des caractéristiques différentes des autres personnes répondantes et, le cas échéant, l'ampleur du biais potentiel a été estimée. Cette analyse laisse supposer que la non-réponse partielle n'a qu'un effet négligeable sur les estimations pour ces questions. De ce fait, les estimations produites à partir de ces quatre indicateurs sont fiables.

Imputation du revenu du ménage

Le revenu du ménage est une information généralement difficile à obtenir de la part des personnes répondantes dans une enquête populationnelle, et cette enquête ne fait pas exception. Le revenu du ménage est manquant pour 19 % des personnes répondantes. En présence d'une non-réponse partielle aussi élevée, il est souhaitable de procéder à l'imputation de données, de manière à réduire les risques de biais. L'imputation des données manquantes est une solution mise de l'avant pour minimiser l'effet de la non-réponse partielle dans les enquêtes. L'imputation consiste à remplacer la donnée manquante pour une personne par une valeur attribuée sur la base des renseignements disponibles au sujet de cette personne. Une imputation valable requiert la présence de renseignements corrélés aux questions à imputer et qui ne présentent pas (ou presque) eux-mêmes une non-réponse partielle. La variable quantitative indiquant le revenu total du ménage a été imputée au moyen d'arbres de décision et de modèles de régression linéaire générés à partir de renseignements potentiellement liés au revenu tels que l'âge au moment de l'entrevue, l'occupation principale, la composition du ménage, l'indice de défavorisation matérielle et sociale, le profil de parcours de diagnostic et de traitement, etc.

Toutefois, l'imputation de données a pour effet d'augmenter artificiellement le nombre de personnes répondantes pour lesquelles nous avons une donnée de revenu du ménage, sans que l'on en tienne toutefois compte dans l'estimation de la variance selon les méthodes habituelles, ce qui fait que la variance des estimations faisant intervenir le niveau de revenu est sous-estimée. Il est important de tenir compte de cette réalité lors de l'interprétation des résultats. De plus, étant donné que la non-réponse partielle pour le revenu est plus élevée pour certaines sous-populations, il est recommandé d'interpréter avec prudence les tests de différences dont le seuil observé est très près du seuil de signification fixé.

4

Analyse des données, précision et tests statistiques

Cette section porte sur certains aspects de l'analyse, dont l'estimation de la précision et les tests statistiques. On y présente également l'indicateur de niveau de revenu du ménage pouvant être utilisé comme variable de croisement dans l'enquête.

4.1 Précision des estimations et tests statistiques

Dans toute enquête populationnelle, les estimations produites comportent différents types d'erreurs. Il y a tout d'abord l'erreur liée à l'échantillonnage, qui est attribuable au fait que seule une partie de la population visée est sélectionnée pour participer à l'enquête. Il peut aussi y avoir d'autres erreurs comme celles causées par la non-réponse, par une faiblesse de la base de sondage ou encore par des erreurs de mesure associées à une mauvaise compréhension de la question par le répondant, à la saisie ou au traitement de données. Rappelons que la principale source d'erreur non liée à l'échantillonnage est la non-réponse, comme mentionné dans les sections précédentes.

L'erreur due à l'échantillonnage se répercute sur les estimations produites pour lesquelles la précision est influencée par la complexité du plan d'échantillonnage. Il est donc nécessaire de mesurer la précision de chaque estimation avant de procéder à l'analyse et à l'interprétation des résultats qui sont inférés à la population visée. Dans l'enquête, l'estimation de la variance et les tests statistiques sont effectués à l'aide de poids d'autoamorçage, aussi appelés « poids *bootstrap* ». Ainsi, une série de 500 poids d'autoamorçage a été créée afin de tenir compte adéquatement non seulement du plan de sondage complexe, mais aussi de tous les ajustements d'inadmissibilité, de non-réponse et de calage apportés à la pondération dans l'estimation de la variance et les tests statistiques (Rust et Rao 1996). Ces poids doivent être utilisés dans l'estimation de la variance et dans les tests statistiques à l'aide de logiciels tels que SAS et SUDAAN. Le coefficient de variation (CV)¹⁹ a été retenu comme indicateur de précision relative pour les diffusions de résultats de l'ISQ. Plus le CV est petit, plus l'estimation est précise.

Une autre mesure de la précision d'une estimation est l'étendue de l'intervalle de confiance (IC) : plus l'intervalle est court, plus la valeur du paramètre est circonscrite. L'intervalle de confiance associé à une proportion représente la zone d'incertitude liée à l'estimation. Cette zone d'incertitude est étroitement liée au niveau de confiance choisi. Ainsi, un intervalle de confiance pour une proportion à un niveau de confiance de 95 % signifie que si on répétait l'enquête 100 fois et qu'à chaque fois, on estimait la proportion et calculait son intervalle de confiance, 95 des 100 intervalles ainsi créés contiendraient la vraie valeur de la proportion dans la population.

19. Le coefficient de variation est obtenu en divisant l'erreur-type de l'estimation par l'estimation elle-même.

Dans les tableaux de l'enquête, en présence d'un croisement entre deux variables, un test statistique d'indépendance du khi-deux²⁰ est utilisé pour faire une comparaison globale des proportions entre différents sous-groupes (p. ex. les catégories d'âge). En présence d'un écart significatif au seuil de signification fixé²¹, et lorsque la variable d'analyse ou la variable de croisement compte plus de deux catégories, des tests de comparaison des proportions peuvent être menés afin de repérer les écarts les plus importants. Ces tests reposent sur une statistique de Wald construite à partir de la différence de la transformation « logit » des proportions²². Notons qu'un écart important entre deux proportions n'est pas nécessairement statistiquement significatif. Tout dépend de la variabilité associée aux estimations. Des résultats non significatifs peuvent être signalés s'ils offrent un intérêt particulier ou s'ils font exception dans une série de résultats significatifs. Ils peuvent être présentés sous forme de tendance.

4.2 Indicateur du niveau de revenu du ménage

Un indicateur du niveau de revenu a été construit à partir des données du revenu total du ménage (déclarées par les personnes répondantes ou imputées²³ à la suite de la validation) en utilisant l'indicateur de faible revenu basé sur la mesure de faible revenu (MFR) avant impôt.

Rappelons que cette enquête, contrairement aux enquêtes portant spécifiquement sur le sujet, ne comporte qu'une seule question sur le revenu. L'indicateur construit permet de créer des groupes homogènes de personnes selon le niveau de revenu de leur ménage. C'est donc possible d'utiliser cet indicateur comme variable de croisement afin d'examiner les liens avec les variables d'intérêt portant sur l'expérience de traitement, l'état de santé, les conditions et la qualité de vie, et sur l'utilisation de soins et de services de santé des personnes ayant eu des antécédents de cancer. Cependant, comme c'est en général le cas dans les enquêtes auprès des individus, qui comportent souvent une seule question globale sur le revenu total du ménage, on observe une surestimation de la défavorisation des ménages causée par une sous-déclaration des revenus. Pour une estimation plus fiable du taux de faible revenu, le lectorat est invité à consulter les publications de l'ISQ sur le sujet (ISQ 2022).

20. On utilise une version modifiée du test du khi-deux habituel qui tient compte du plan de sondage de l'enquête : il s'agit de l'ajustement de Satterthwaite du test du khi-deux. Plus précisément, c'est la statistique F correspondant à cette correction du test du khi-deux qui est utilisée dans les analyses.

21. Le seuil fixé dans cette étude est de 5 %.

22. La transformation $\text{logit} = \log(p/(1 - p))$ est utilisée comme transformation normalisatrice dans la construction d'intervalles de confiance afin que leurs taux de couverture soient améliorés, en particulier pour de petites proportions (Korn et Graubard 1999).

23. Voir la section 3.4 au sujet de l'imputation du revenu.

5 Présentation des résultats

Dans le cadre de l'enquête, les estimations de proportions publiées par l'ISQ sont arrondies à la décimale près dans les tableaux et à l'unité près dans le texte, à l'exception des proportions inférieures à 5 % pour lesquelles une décimale a été conservée. Les proportions dont la décimale est ,5 sont arrondies à l'unité inférieure ou supérieure selon la seconde décimale. En raison de l'arrondissement, la somme des proportions présentées dans certains tableaux peut différer légèrement de 100 %.

Dans les tableaux, les estimations dont le CV est inférieur ou égal à 15 % sont suffisamment précises pour être présentées sans indication. Les estimations dont le CV est supérieur à 15 %, mais inférieur ou égal à 25 % sont accompagnées d'un astérisque (*) (voir le tableau 5.1) indiquant un niveau de précision passable. Les estimations dont le CV est supérieur à 25 % sont marquées d'un double astérisque (**) pour indiquer une précision faible. Ces dernières ne sont fournies qu'à titre indicatif.

Tableau 5.1

Système de classification des estimations, en termes de précision, en vigueur à l'Institut de la statistique du Québec

Précision de l'estimation (CV)	Signe distinctif	Réserve à mentionner lors de la diffusion
CV ≤ 15 %	(aucun signe)	Estimation diffusée sans mention.
15 % < CV ≤ 25 %	*	Estimation à interpréter avec prudence.
CV > 25 %	**	Estimation à utiliser avec circonspection, fournie à titre indicatif seulement.

Source : Institut de la statistique du Québec.

Dans les tableaux présentant des analyses bivariées, en présence d'une association significative selon le test du khi-deux, des lettres en exposant ajoutées aux statistiques présentées indiquent les paires de catégories d'une variable de croisement pour lesquelles la variable d'analyse diffère significativement. Une même lettre révèle un écart significatif entre deux catégories au seuil de 5 %.

En général, dans le but de faire ressortir les principaux résultats, seules les différences significatives au seuil de 5 % sont mentionnées dans le texte. Il arrive que deux proportions qui semblent différentes ne le soient pas d'un point de vue statistique. On dit dans ce cas qu'il n'y a pas de différence statistiquement significative, ou que l'enquête ne permet pas de détecter de différence entre ces proportions.

Conformément aux nouvelles pratiques de l'ISQ en matière de statistiques sociales, les résultats de l'enquête ont été ventilés selon le genre des personnes plutôt que selon le sexe. Par genre, on entend le genre actuel, qui peut différer du sexe assigné à la naissance ou de celui inscrit dans les documents officiels. Pour des raisons de qualité des estimations et de confidentialité, et compte tenu de la petite taille de la population concernée, la publication de statistiques pour les personnes non binaires n'est pas possible pour cette enquête. Une variable « genre » à deux catégories est donc utilisée, laquelle comprend la répartition des personnes non binaires. Ainsi, la catégorie « Hommes+ » comprend les hommes et certaines personnes non binaires ; la catégorie « Femmes+ » comprend les femmes et certaines personnes non binaires²⁴.

24. [Prise en compte du genre \(quebec.ca\)](https://www.quebec.ca/genre).

6

Portée et limites de l'enquête

Les données des 13 070 personnes répondantes de l'enquête offrent un grand potentiel analytique à l'échelle provinciale pour la population visée²⁵. Tout a été mis en place pour maximiser la qualité, la représentativité et la comparabilité des résultats produits. Un prétest a permis de vérifier l'intelligibilité des questions, et des séances de formation ont été offertes aux intervieweurs et intervieweuses afin d'uniformiser les communications avec les personnes sélectionnées et la collecte des données. De plus, des efforts considérables ont été déployés à la collecte pour obtenir le taux de réponse visé. Finalement, le très bon taux de réponse obtenu (73 %) a dépassé celui attendu à l'échelle du Québec (57 %). Il a même aussi été dépassé par siège de cancer, par sexe et pour la majorité des groupes d'âge (tableau 2.1). Ce très bon taux de réponse obtenu permet de réduire les risques de biais dus à la non-réponse. Il permet aussi d'obtenir des précisions supérieures à celles qui étaient prévues.

Toutes les estimations produites dans le cadre de l'enquête sont pondérées et tiennent compte non seulement du plan de sondage, mais aussi de la non-réponse totale, de manière à assurer la fiabilité de l'inférence à la population visée. De plus, toutes les mesures de précision et les tests statistiques ont été produits en prenant en compte la construction du plan de sondage de l'enquête.

Toutefois, malgré les précautions prises lors de la planification et de la réalisation de l'enquête, l'analyse des données demeure contrainte par certaines limites, comme dans toute enquête populationnelle. Tout d'abord, il est impossible de garantir l'exactitude des réponses fournies par les personnes répondantes. En effet, l'étude se fait 10 ans après que ces personnes ont reçu un diagnostic de cancer, ce qui peut entraîner que certaines des personnes répondantes aient oublié quelques informations demandées dans le questionnaire. En plus des biais de mémoire, il se peut que des biais de modulation des faits, de désirabilité sociale ou de compréhension persistent.

Il est important de tenir compte des mises en garde et des recommandations mentionnées dans les sections précédentes afin d'utiliser adéquatement les données de cette enquête. Dans la section 4.1.1, on explique, entre autres, comment tenir compte du plan de sondage complexe lors de la mesure de la précision des résultats produits. De même, il faut prendre en considération lors de l'interprétation des résultats que l'erreur due à l'imputation du revenu n'est pas prise en compte dans la mesure de précision (section 3.4). Par exemple, les résultats des tests pour lesquels le seuil observé est près du seuil de signification devraient être interprétés avec prudence.

De plus, les données d'observation telles que celles recueillies dans le cadre de l'enquête ne permettent pas d'établir de lien de causalité. Les associations ou les différences significatives observées peuvent laisser croire qu'il existe un tel lien, mais elles ne permettent pas de le confirmer. Elles ne permettent pas non plus d'effectuer des analyses détaillées et très précises pour tous les types de sièges de cancer, puisque la construction du plan de sondage n'en a pas tenu compte.

25. Voir la section 1.1 sur la description de la population visée.

Compte tenu des objectifs spécifiques de l'enquête, les analyses présentées par l'ISQ s'appuient sur des méthodes bivariées. L'interprétation de certains résultats doit donc être faite avec prudence. Une analyse multivariée aurait été appropriée dans certains cas pour contrôler des facteurs confondants. Toutefois, l'approche retenue permet adéquatement d'explorer les données recueillies de façon utile afin de fournir un portrait sur l'état de santé physique et émotionnelle, sur les conditions et la qualité de vie ainsi que sur l'utilisation des soins et services de santé des personnes ayant reçu, en 2013 et 2014, un diagnostic de cancer, alors qu'elles étaient âgées de 18 ans et plus.

Annexe

Tableau A1

Questions ou indicateurs présentant un taux de non-réponse partielle pondéré supérieur à 5 %, Québec, 2025

Nom de la variable ou de l'indicateur	Mesure	Taux de non-réponse partielle pondéré	Effet possible sur l'analyse
SBEP_Q10R_4CAT	Avez-vous eu des difficultés à avoir des enfants en raison du cancer ou de son traitement ?	13,7 %	Négligeable
CONSEQ_PBSANTE	Associer au moins un des problèmes de santé physique ou mentale au cancer ou à son traitement.	5,4 %	Négligeable
CONSEQ_PHYS	Associer un des problèmes de santé physique au cancer ou à son traitement.	5,4 %	Négligeable
REVENU_IMP MFR	Pour les 12 derniers mois, à combien estimez-vous approximativement le revenu total de tous les membres du ménage provenant de toutes les sources, avant impôt et autres retenues ?	19,1 %	Sous-estimation de la variance due à l'imputation effectuée

Source : Institut de la statistique du Québec, *Enquête québécoise sur la vie 10 ans après un diagnostic de cancer, 2024*.

Références bibliographiques

AZEVEDO DA SILVA, M., et autres (2024). *Indice de défavorisation matérielle et sociale 2021 : transfert des connaissances : guide d'utilisation*, [En ligne], Québec, Institut national de santé publique du Québec, 19 p. [www.inspq.qc.ca/sites/default/files/2024-04/3476-indice-defavorisation-materielle-sociale-2021.pdf] (Consulté le 2 juillet 2025).

BERNIER, J., et K. NOBREGA (1998). « Outlier detection in asymmetric samples: A comparison of an inter-quartile range method and a variation of a sigma-gap method », *Congrès annuel de la société Statistique du Canada*, [En ligne], p. 137-141. [ssc.ca/sites/default/files/survey/documents/SSC1998_J_Bernier.pdf] (Consulté le 2 juillet 2025).

DIRECTION GÉNÉRALE DE LA SANTÉ PUBLIQUE, D. Q. D. C. (2012). « Aperçu du système d'information du RQC », dans *Registre québécois de cancer. Cadre normatif : Partie 1 : Consignes à la déclaration et dictionnaire de données*, [En ligne], Québec, Ministère de la Santé et des Services sociaux, p. 1-4. [publications.msss.gouv.qc.ca/msss/fichiers/2012/12-902-04W.pdf] (Consulté le 7 juillet 2023).

ELTINGE, J. L., et S. YANSANEH (1997). « Méthodes diagnostiques pour la construction de cellules de correction pour la non-réponse, avec application à la non-réponse aux questions sur le revenu de la U.S. Consumer Expenditure Survey », *Techniques d'enquête*, [En ligne], produit n° 12-001-X19970013103 au catalogue de Statistique Canada, vol. 23, n° 1, juin, p. 37-45. [www150.statcan.gc.ca/n1/fr/pub/12-001-x/1997001/article/3103-fra.pdf?st=YRiuivNK] (Consulté le 2 juillet 2025).

HAZIZA, D., et J. F. BEAUMONT (2007). « On the Construction of Imputation Classes in Surveys », *International Statistical Review*, [En ligne], vol. 75, n° 1, avril, p. 25-43. doi : [10.1111/j.1751-5823.2006.00002.x](https://doi.org/10.1111/j.1751-5823.2006.00002.x). (Consulté le 2 juillet 2025).

INSTITUT DE LA STATISTIQUE DU QUÉBEC (2022). *Revenu et faible revenu au Québec en 2019 : les plus récentes données et les tendances depuis 25 ans*, [En ligne]. [statistique.quebec.ca/fr/document/revenu-et-faible-revenu-quebec-2019-donnees-et-tendances-depuis-25-ans] (Consulté le 31 juillet 2025).

KORN, E. L., et B. I. GRAUBARD (1999). *Analysis of health surveys*, New York, John Wiley and Sons, 382 p.

MINISTÈRE DE LA SANTÉ ET DES SERVICES SOCIAUX (2019). *Registre québécois du cancer CADRE NORMATIF Partie 1 : Présentation du système d'information et description des éléments*, [En ligne], Québec, MSSS, 173 p. [publications.msss.gouv.qc.ca/msss/fichiers/2019/19-902-03W.pdf] (Consulté le 14 septembre 2023).

RUST, K. F., et J. N. K. RAO (1996). « Variance estimation for complex surveys using replication techniques », *Statistical Methods in Medical Research*, [En ligne], vol. 5, n° 3, septembre, p. 283-310. doi : [10.1177/096228029600500305](https://doi.org/10.1177/096228029600500305). (Consulté le 2 juillet 2025).

« Une organisation
statistique performante
au service d'une société
québécoise en évolution »

statistique.quebec.ca