

Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022

Rapport d'analyse de comparaison

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 2024
ISBN 978-2-550-99092-5 (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

Mars 2025

Publication réalisée à l'Institut de la statistique du Québec par :	Mbuyi Kelelekela et Daniela Aranibar Zeballos
Avec la collaboration de :	Valeriu Dumitru, France Lapointe
Sous la coordination de :	Marie-Eve Tremblay
Révision linguistique et édition :	Direction de la diffusion et des communications
Comité de relecture interne :	Marie-Eve Tremblay, France Lapointe, Amélie Ducharme, Nathalie Audet, Éric Gagnon et Bertrand Perron
Comité de relecture externe :	Jacinthe Giguère et Mathieu Gerbeau Ministère de l'Éducation
Étude financée par :	Ministère de l'Éducation
Étude sous la responsabilité de :	Direction des enquêtes et des indicateurs sociaux Institut de la statistique du Québec
Pour tout renseignement concernant le contenu de ce document, s'adresser à :	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 pour le rapport

KELELEKELA, Mbuyi, et Daniela ARANIBAR ZEBALLOS (2025). *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022. Rapport d'analyse de comparaison*, [En ligne], Québec, Institut de la statistique du Québec, 58 p. [statistique.quebec.ca/fr/fichier/rapport-evaluation-maternelle-4-ans-temps-plein-2022.pdf].

Avertissement

En raison de l'arrondissement des données, le total ne correspond pas nécessairement à la somme des parties.

Table des matières

Introduction	6
1 Aspects méthodologiques généraux de l'Étude pour l'évaluation de la maternelle 4 ans à temps plein	9
1.1 Plan d'échantillonnage	9
1.1.1 Population visée	9
1.1.2 Base de sondage	11
1.1.3 Hypothèses d'échantillonnage, taille et répartition de l'échantillon	12
1.2 Collecte des données	14
1.2.1 Prétest	14
1.2.2 Consentements	14
1.2.3 Collecte des données de l'enquête principale	15
1.2.4 Résultats de collecte	18
1.3 Traitement des données	18
1.3.1 Validation et saisie	18
1.3.2 Pondération de description	19
1.3.3 Non-réponse partielle	21
1.4 Précisions des estimations et tests statistiques	24
1.4.1 Précisions des estimations	24
1.4.2 Tests statistiques	24
1.5 Présentation des résultats du recueil statistique	25
2 Aspects méthodologiques propres à l'étude comparative	26
2.1 Plan d'échantillonnage pour la comparaison	26
2.1.1 Population visée pour la comparaison	26
2.1.2 Taille et répartition de l'échantillon et des répondants pour la comparaison	28
2.2 Pondération de comparaison	29
2.2.1 Stratégie de pondération pour la comparaison	29
2.2.2 Particularités de la pondération de comparaison	29
2.3 Mesures du développement des enfants et de la transition à la maternelle 5 ans	31
2.3.1 Vulnérabilité par domaine de développement et vulnérabilité globale	31
2.3.2 Transition à la maternelle 5 ans	32

2.4	Démarche de modélisation et présentation des résultats	33
2.4.1	Hypothèse de l'analyse comparative	33
2.4.2	Variables de modélisation	33
2.4.3	Étapes de modélisation	34
3	Résultats des analyses comparatives de la fréquentation de la maternelle 4 ans	38
3.1	Analyses descriptives (recueil statistique) et analyses comparatives (section 3.2)	38
3.2	Résultats des analyses comparatives.	40
3.2.1	Vulnérabilité et fréquentation de la maternelle 4 ans selon les analyses comparatives.	41
3.2.2	Transition à la maternelle 5 ans selon le parent et fréquentation de la maternelle 4 ans selon les analyses comparatives.	44
	Conclusion	45
	Annexe 1 Différences méthodologiques entre l'EQPPEM 2022 et l'EEM4 2022	50
	Annexe 2 Plan d'échantillonnage de l'EEM4 2022 – global	53
	Annexe 3 Plan d'échantillonnage de l'EEM4 2022 – par sous-population	54
	Annexe 4 Modalités des indicateurs retenus dans la modélisation	56
	Bibliographie	57

Introduction

Le programme préscolaire de maternelle 4 ans à temps plein a été implanté en milieux défavorisés lors de l'année scolaire 2013-2014 pour les enfants âgés de 4 ans. Il a pour objectif de développer le plein potentiel des enfants pour un passage réussi vers la maternelle 5 ans et vise à favoriser l'apprentissage et le développement global de l'enfant par le jeu ainsi qu'à renforcer la collaboration entre l'école et les parents.

La participation au programme de maternelle 4 ans à temps plein est facultative. Par ailleurs, il existe d'autres programmes préscolaires publics offerts aux enfants visés par la maternelle 4 ans à temps plein. Par exemple, la maternelle 4 ans à mi-temps et le programme Passe-Partout sont offerts depuis 1973 (Ministère de l'Éducation et de l'Enseignement supérieur 2017).

Lors de son implantation, le programme de maternelle 4 ans à temps plein comptait 56 classes (Ministère de l'Éducation et de l'Enseignement supérieur 2017). L'admissibilité des enfants au programme reposait initialement sur l'indice de milieu socio-économique (IMSE)¹ de leur adresse de résidence. Comme visé à l'origine par la mesure d'implantation de la maternelle 4 ans à temps plein en milieu défavorisé, la majorité des élèves fréquentant le programme en 2013-2014 venait donc d'un milieu dont l'IMSE était de rang 9 ou 10 (Ministère de l'Éducation et de l'Enseignement supérieur 2017).

En 2020, suivant l'adoption du projet de loi n° 5, Loi modifiant la Loi sur l'instruction publique et d'autres dispositions à l'égard des services de l'éducation préscolaire destinés aux élèves âgés de 4 ans, le programme de maternelle 4 ans à temps plein a été plus vastement déployé. L'objectif était de le rendre accessible à tout enfant âgé de 4 ans, indépendamment du milieu économique dans lequel il vit (Ministère de l'Éducation 2023). Le projet de loi permet également aux établissements d'enseignement privés d'offrir la maternelle 4 ans à temps plein, sous réserve de l'autorisation du ministre de l'Éducation.

En 2021-2022, le déploiement du programme a porté à 1 345 le nombre de classes de maternelle 4 ans à temps plein, qui comptaient plus de 15 000 élèves (Ministère de l'Éducation 2023). Ces classes se trouvent dans des milieux socio-économiques variés.

Étant donné l'évolution du programme, le ministère de l'Éducation (MEQ) souhaite l'évaluer une seconde fois à l'aide, entre autres, d'un projet d'enquête auprès des parents et du personnel enseignant. Il a donc mandaté l'Institut de la statistique du Québec (ISQ) pour mener *l'Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022* (EEM4), qui permettra de fournir des données pour l'évaluation du programme de maternelle 4 ans à temps plein. Ce projet a été intégré à deux autres

La première évaluation du programme de maternelle 4 ans à temps plein

Le programme de maternelle 4 ans à temps plein a été évalué une première fois en 2015. Dans le cadre de cette évaluation, l'Institut de la statistique du Québec a été mandaté pour mener une enquête auprès des parents dont l'enfant avait fréquenté la maternelle 4 ans à temps plein en 2013-2014 et des enseignants et enseignantes de ces enfants lors de l'année scolaire suivante, afin d'évaluer leur intégration en maternelle 5 ans. Le ministère de l'Éducation, de son côté, a recueilli des informations pour l'évaluation à l'aide d'entrevues avec des directions d'école, des conseillers et conseillères pédagogiques ainsi que des enseignants et enseignantes du programme de maternelle 4 ans à temps plein.

1. Pour plus d'informations sur l'IMSE, consulter la page [Indice de milieu socio-économique \(IMSE\) | Ministère de l'Éducation \(gouv.qc.ca\)](#).

enquêtes de l'ISQ ayant en commun avec l'EEM4 certains objectifs, soit *l'Enquête québécoise sur le développement des enfants à la maternelle 2022* (EQDEM 2022)², menée auprès des enseignantes et enseignants des enfants inscrits à la maternelle 5 ans au Québec en 2021-2022, et *l'Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2022* (EQPPEM 2022)³, menée auprès des parents.

Puisque l'EEM4 vise à décrire le lien entre le développement de l'enfant et la fréquentation ou non de la maternelle 4 ans à temps plein, elle a été conduite à la suite de l'EQDEM, dont l'objectif est de mesurer le niveau de développement relatif des enfants à la maternelle 5 ans dans cinq domaines de développement.

La collecte de données de l'EEM4 a été réalisée conjointement avec celle de l'EQPPEM, puisque cette dernière vise les mêmes personnes répondantes, soit les parents des enfants qui ont fréquenté la maternelle 5 ans en 2021-2022, et qu'elle recueille des données utiles à l'EEM4 telles que des informations sur le parcours préscolaire de ces enfants et sur les caractéristiques de leur famille. Lors de la collecte, les parents des enfants sélectionnés pour participer à l'EEM4 devaient aussi répondre à quelques questions sur la transition de leur enfant à la maternelle 5 ans ainsi que sur leur connaissance du programme de maternelle 4 ans à temps plein et leur satisfaction envers celui-ci.

Les informations recueillies lors de la collecte de l'EQPPEM 2022 ont été jumelées à celles de l'EQDEM 2022. Les données tirées de ces deux enquêtes ont permis de répondre aux objectifs de l'EEM4, décrits ci-contre.

Objectifs de l'EEM4

L'objectif principal de l'EEM4 est de fournir des informations statistiques fiables qui contribuent à l'évaluation du programme de maternelle 4 ans à temps plein du MEQ. L'étude permet notamment de brosser un portrait des enfants inscrits à la maternelle 5 ans en 2021-2022 qui ont fréquenté ou non la maternelle 4 ans à temps plein en 2020-2021 dans les écoles qui avaient offert la maternelle 4 ans l'année précédente ainsi que de comparer certains aspects de ces deux sous-populations.

Plus précisément, les objectifs de l'étude sont les suivants :

1. Décrire le lien entre le niveau de développement des enfants inscrits à la maternelle 5 ans en 2021-2022 mesuré dans l'EQDEM 2022 à l'aide de l'Instrument de mesure du développement de la petite enfance (IMDPE) et la fréquentation de la maternelle 4 ans à temps plein en 2020-2021 en tenant compte, notamment, du parcours préscolaire et des caractéristiques des enfants et des familles ;
2. Décrire le lien entre la transition à l'école des enfants et la fréquentation de la maternelle 4 ans à temps plein en tenant compte, notamment, du parcours préscolaire et des caractéristiques des enfants et des familles ;
3. Décrire les motifs de fréquentation ou non de la maternelle 4 ans à temps plein, la connaissance du programme par les parents dont les enfants n'y participent pas ainsi que la satisfaction des parents dont les enfants y participent.

En complément de l'étude menée par l'ISQ au printemps 2022, le MEQ a réalisé des sondages (questionnaires et groupes de discussion) auprès d'enseignants et d'enseignantes de maternelle 4 ans à temps plein, de directions d'école et de conseillers et conseillères pédagogiques, afin de faire état de la mise en œuvre de la maternelle 4 ans à temps plein et les perceptions qu'en a le personnel d'intervention en milieu scolaire. Les résultats sur les perceptions du personnel scolaire sont présentés dans un premier rapport⁴ publié par le MEQ en 2023.

2. Pour plus de détails sur la méthodologie de l'EQDEM 2022, consulter le [rapport méthodologique de l'enquête](#).

3. Pour plus de détails sur la méthodologie de l'EQPPEM 2022, consulter le [rapport méthodologique de l'enquête](#).

4. Les informations recueillies par le MEQ sur la perception du personnel scolaire au sujet du programme de maternelle 4 ans à temps plein peuvent être consultées dans le rapport suivant : [Évaluation de la maternelle 4 ans - Rapport d'étape 2023 \(quebec.ca\)](#).

Contenu du rapport

Le présent rapport résume les méthodes et les résultats des analyses statistiques effectuées pour répondre aux objectifs de l'étude.

Dans le cadre de l'EEM4, un recueil statistique fournissant des résultats en lien avec le troisième objectif de l'étude, énoncé plus haut, a été diffusé sur le site Web de l'ISQ⁵. La première partie du présent rapport décrit les principaux paramètres de l'EEM4 ainsi que la méthodologie employée pour produire les résultats fournis dans ce recueil. La deuxième partie décrit les aspects méthodologiques propres à l'analyse de comparaison des enfants qui ont fréquenté la maternelle 4 ans à temps plein avec ceux qui ne l'ont pas fréquentée. Enfin, la troisième partie

présente les résultats de cette analyse de comparaison en ce qui a trait au développement des enfants et à leur transition à la maternelle 5 ans selon leur fréquentation de la maternelle 4 ans.

Des analyses incluant la fréquentation de la maternelle 4 ans ont aussi été réalisées avec l'ensemble des données de l'EQPPEM et les résultats sont présentés dans le rapport de l'EQPPEM. Les résultats peuvent être différents de ceux présentés dans ce rapport, qui sont basés entre autres sur les données concernant un échantillon d'enfants plus restreint. L'annexe I explique les différences méthodologiques entre les deux analyses.

5. Le recueil statistique de l'EEM4 peut être consulté en ligne : (insérer hyperlien).

1 Aspects méthodologiques généraux de l'Étude pour l'évaluation de la maternelle 4 ans à temps plein

Pour utiliser et interpréter adéquatement les résultats de l'EEM4, il est nécessaire de connaître la méthodologie employée pour les produire. Celle-ci s'avère être un outil important qui permet aux utilisateurs de bien comprendre ces résultats.

Sont ici décrits successivement les éléments du plan d'échantillonnage tels que la population visée, la base de sondage et l'échantillonnage, les stratégies de collecte de données et le traitement des données ainsi que les normes de présentation des résultats.

1.1 Plan d'échantillonnage

La présente section décrit les paramètres méthodologiques de la construction des échantillons utiles à l'étude.

1.1.1 Population visée

L'EEM4 vise l'ensemble des enfants fréquentant la maternelle 5 ans en 2021-2022 dans les écoles francophones et anglophones, publiques et privées (subventionnées ou non) du Québec qui comptaient au moins trois enfants ayant fréquenté la maternelle 4 ans à temps plein en 2020-2021.

Notons que le critère du nombre minimal d'enfants a été établi en vue d'améliorer l'efficacité du plan d'échantillonnage de l'enquête quant à la précision des estimations à produire et à la puissance statistique des tests à effectuer. Notons que cette population visée est donc une sous-population de celle de l'EQDEM 2022.

Le statut de fréquentation de la maternelle 4 ans à temps plein a été établi à partir de l'information qu'a le MEQ au sujet de l'enfant dans ses fichiers administratifs.

Étant donné que l'EEM4 est une étude cas-témoins, dont l'un des objectifs est la comparaison des enfants qui ont fréquenté la maternelle 4 ans avec ceux qui ne l'ont pas fréquentée, en ce qui a trait notamment à leur développement lors du passage à la maternelle 5 ans, la population visée est divisée en deux sous-populations cibles :

- Les « enfants-maternelle 4 ans » : enfants de la population visée qui ont fréquenté la maternelle 4 ans à temps plein⁶ en 2020-2021 ;
- Les « enfants-témoins » : enfants de la population visée qui n'ont pas fréquenté la maternelle 4 ans à temps plein en 2020-2021.

Quelques exclusions s'appliquent à cette définition de la population visée :

- Exclusion des enfants exclus de la population visée par l'EQDEM⁷

6. Pour répondre aux objectifs de l'étude, les enfants qui ont fréquenté la maternelle 4 ans à temps partiel ne sont pas dans la sous-population des enfants-maternelle 4 ans, mais plutôt dans la sous-population des enfants-témoins.

7. À titre d'information, considérant l'ensemble des exclusions, la population visée par l'EQDEM représente 98 % de l'ensemble des enfants de maternelle 5 ans en 2021-2022 selon une compilation de la Direction de la méthodologie à partir des renseignements du fichier transmis par le MEQ en mai 2022.

Du fait que l'EEM4 utilise l'information provenant de l'EQDEM, les critères d'exclusion de la population visée par l'EQDEM s'appliquent aussi à sa population visée :

- Les enfants fréquentant une école faisant partie des commissions scolaires crie et Kativik ou située dans les régions sociosanitaires du Nunavik (17) et des Terres-Cries-de-la-Baie-James(18) ainsi que les enfants fréquentant des établissements relevant du gouvernement fédéral dans les réserves autochtones. Ces élèves ont été exclus pour les raisons suivantes : les difficultés à les joindre, les modifications qu'il aurait fallu apporter au questionnaire pour tenir compte de leur contexte de vie ou de leur langue et les modalités de fréquentation de la maternelle 5 ans qui peuvent y être différentes ;
- Tous les élèves fréquentant une école spéciale⁸ ;
- Les enfants qui, selon les fichiers administratifs du MEQ ou selon les renseignements obtenus de l'école, ont un code d'élèves handicapés ou en difficulté d'adaptation ou d'apprentissage (EHDAA)⁹ faisant partie d'une classe spéciale composée de plus de 50 % d'EHDAA. Toutefois, les EHDAA dans une classe ordinaire où 50 % des élèves ou moins ont un code EHDAA participent à l'EQDEM et sont inclus dans l'étude, mais exclus des résultats présentés dans ce rapport et dans le recueil statistique.

- Exclusion des élèves en situation d'admission exceptionnelle – enfants trop jeunes ou trop âgés : ces enfants n'auraient pas pu fréquenter la maternelle 4 ans à temps plein en 2020-2021 puisqu'ils étaient alors trop jeunes ou déjà à la maternelle 5 ans.

Au total, 47 289 enfants inscrits à la maternelle 5 ans en 2021-2022 sont visés par l'étude ; ils sont répartis dans 2 982 classes de 1 009 écoles, dont 24 écoles privées et 985 écoles relevant des 70 centres de services scolaires ou commissions scolaires.

De ces enfants, 10 870 ont fréquenté la maternelle 4 ans à temps plein en 2020-2021, ce qui correspond à un peu plus de 93 % des enfants ayant fréquenté la maternelle 4 ans à temps plein en 2020-2021¹⁰. La sous-population des enfants-témoins compte quant à elle 36 419 enfants de maternelle 5 ans qui n'ont pas fréquenté de maternelle 4 ans à temps plein en 2020-2021.

L'EEM4 porte ainsi sur une population correspondant à un peu plus de la moitié des enfants de maternelle 5 ans visés par l'EQDEM 2022, comme l'indique le tableau 1.1.

Tableau 1.1

Effectifs de population selon l'enquête : EQDEM ou EEM4

	EQDEM	EEM4
Nombre d'écoles admissibles	1 941	1 009
Nombre d'enfants admissibles	87 264	47 289
Nombre d'enfants-maternelle 4 ans admissibles ¹	-	10 870
Nombre d'enfants-témoins admissibles ¹	-	36 419
Nombre de classes dans les écoles participantes	5 373	2 982

1. Information non pertinente pour l'EQDEM.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

8. Comprend, entre autres, les écoles des hôpitaux et des centres de réadaptation. Environ 450 enfants fréquentaient ces établissements en 2021-2022.
9. Pour la liste des codes de difficulté accompagnés des définitions, voir MELS (2007). À noter que le système scolaire utilise généralement la terminologie EHDAA pour décrire les enfants handicapés ou ayant des difficultés d'adaptation ou d'apprentissage.
10. Le critère d'un minimum de trois enfants ayant fréquenté la maternelle 4 ans à temps plein par école en 2020-2021 utilisé pour définir la population visée entraîne une sous-couverture d'un peu plus de 6 % des enfants ayant fréquenté la maternelle 4 ans à temps plein.

1.1.2 Base de sondage

La base de sondage utilisée pour l'EEM4 a été construite à partir de l'information provenant des systèmes administratifs du MEQ. Les données transmises à l'ISQ comprennent, entre autres, divers renseignements nominaux sur les centres de services scolaires ou les commissions scolaires, les écoles¹¹ et les élèves¹². Comme les renseignements contenus dans les fichiers du MEQ sont mis à jour de façon continue pendant l'année scolaire, deux extractions des banques administratives ont été faites. Les renseignements de l'extraction de décembre 2021 ont servi à élaborer la base de sondage de l'étude pour l'échantillonnage. Les renseignements recueillis lors de la collecte de l'EQDEM auprès du personnel enseignant ainsi que l'extraction de mai 2022 ont servi à l'établissement définitif de la taille de la population visée et permis l'évaluation de la couverture de la base de sondage.

Cette base correspond à la liste des enfants inscrits à la maternelle 5 ans en 2021-2022 dans une école comptant au moins trois enfants qui ont fréquenté la maternelle 4 ans à temps plein en 2020-2021; seules les inscriptions présentes dans le fichier du MEQ en décembre 2021 apparaissent dans la base de sondage. Il faut mentionner que les élèves en situation d'admission exceptionnelle n'ont pas été exclus de cette liste avant la sélection des échantillons. Cette base de sondage comporte de la surcouverture puisque certains élèves présents dans le fichier du MEQ de décembre 2021 n'y figuraient plus en mai 2022 et n'avaient pas été retrouvés lors de la collecte de données de l'EQDEM. Enfin, certains élèves de la base de sondage fréquentaient dans les faits, en 2021-2022, une école qui n'était pas admissible à l'étude ou encore se trouvaient dans une classe spéciale composée de plus de 50 % d'EHDA.

La base de sondage utilisée contient également de la sous-couverture : certains élèves n'apparaissaient pas dans le fichier du MEQ en décembre 2021, mais étaient

présents dans la mise à jour de mai 2022 ou ont été ajoutés par le personnel enseignant lors de la collecte de données de l'EQDEM.

Malgré cela, l'utilisation de l'information provenant du MEQ dans la construction de la base de sondage garantit une bonne couverture de la population visée et des renseignements de qualité sur les écoles et les enfants de maternelle 5 ans. En effet, le fichier du MEQ de décembre 2021 a permis d'obtenir une base de sondage comportant une surcouverture¹³ d'un peu plus de 3 % et une sous-couverture¹⁴ de moins de 7 % de la population visée estimée. La base de sondage est donc suffisante et adéquate pour représenter la population visée. Par ailleurs, la comparaison des renseignements du fichier du MEQ de décembre 2021 concernant chaque enfant avec celle de la mise à jour de mai 2022 a montré des différences marginales. Les coordonnées pour joindre les parents des enfants de l'échantillon étaient rarement manquantes dans la base de sondage.

Enfin, les informations sur la fréquentation de la maternelle 4 ans à temps plein en 2020-2021 dans une école publique ou privée proviennent aussi du fichier administratif du MEQ de décembre 2021, mis à jour avec celui de mai 2022. Pour plus de 99,9 % des enfants de maternelle 5 ans présents dans les deux fichiers, l'information du MEQ était concordante. Cela permet de conclure à une bonne classification des enfants en fonction de la fréquentation ou non de la maternelle 4 ans à temps plein.

11. Les coordonnées des centres de services scolaires, des commissions scolaires et des écoles, la langue d'enseignement de l'école, le réseau d'enseignement, etc.

12. Le nom, le prénom, l'âge, le sexe, la langue maternelle, le code postal, le code d'admissibilité exceptionnelle, etc.

13. La surcouverture correspond au rapport du nombre d'enfants de la base de sondage initiale de décembre 2021 inadmissibles à l'étude au nombre d'enfants de la population visée comme établi à la fin de la collecte de l'EQDEM et après la mise à jour du fichier du MEQ en mai 2022.

14. La sous-couverture correspond au rapport du nombre d'enfants qui se sont ajoutés à la base de sondage initiale lors de sa mise à jour au nombre d'enfants de la population visée comme établi à la fin de la collecte de l'EQDEM et après la mise à jour du fichier du MEQ en mai 2022.

1.1.3 Hypothèses d'échantillonnage, taille et répartition de l'échantillon

Comme l'EEM4 est une étude cas-témoins, elle nécessite d'échantillonner les enfants dans les écoles ayant possiblement offert le programme l'année précédente, c'est-à-dire celles comptant au moins trois enfants qui ont fréquenté la maternelle 4 ans à temps plein. Par conséquent, la réalisation du projet implique la sélection d'un échantillon de « cas » (enfants-maternelle 4 ans) parmi les enfants de la base de sondage ayant fréquenté la maternelle 4 ans à temps plein durant l'année 2020-2021 et d'un échantillon de « témoins » (enfants-témoins) parmi ceux qui ne l'ont pas fréquentée.

Hypothèses d'échantillonnage

La taille d'échantillon a été déterminée en vue d'assurer une bonne puissance statistique dans les deux groupes d'enfants, aussi bien pour leur description que pour leur comparaison. Les objectifs de comparaison des enfants qui ont fréquenté la maternelle 4 ans à temps plein avec les enfants-témoins, notamment en ce qui a trait au niveau de leur développement lors du passage à la maternelle 5 ans, se trouvent au cœur des visées de l'étude. La puissance statistique de ces analyses comparatives a donc été prise en considération dans l'établissement de la taille de l'échantillon nécessaire à l'étude. Les proportions considérées sont celles de l'édition précédente de l'EQDEM. La proportion d'enfants dits vulnérables¹⁵ dans chacun des cinq domaines de développement mesurés par l'EQDEM était d'environ 11 %¹⁶ en 2017. Une mesure composite du développement des enfants, c'est-à-dire la mesure de la vulnérabilité dans au moins un domaine de développement, est également prise en compte.

La taille totale des deux échantillons à sélectionner est déterminée de manière qu'un test statistique au seuil 5 % ait une puissance d'au moins 80 % pour détecter une différence assez petite entre la proportion d'enfants vulnérables chez les enfants-maternelle 4 ans et celle chez les enfants-témoins. À partir de ces paramètres, il a été établi que 4 015 enfants répondants *efficaces*¹⁷ (2 000 enfants ayant fréquenté la maternelle 4 ans à temps plein et 2 015 enfants-témoins) seraient nécessaires à

l'étude pour qu'un écart de 2,8 % entre les deux groupes d'enfants soit détecté dans les proportions d'enfants vulnérables dans chaque domaine de développement et qu'un écart de 4 % le soit dans les proportions d'enfants vulnérables dans au moins un domaine.

Les cibles de nombre de répondants *efficaces* étant fixées, des hypothèses relatives à l'effet du plan de sondage, au taux d'admissibilité ainsi qu'aux taux de réponse du personnel enseignant et des parents déterminent la taille des deux échantillons devant être sélectionnés dans la population visée par l'étude (tableau 1.1).

Taille et répartition de l'échantillon

Avant de fournir des informations détaillées sur la taille et la répartition de l'échantillon de l'EEM4, il faut rappeler que cette enquête fait suite à l'EQDEM 2022 et qu'elle compose avec l'échantillon de l'EQPEM 2022.

Tout d'abord, le plan de sondage doit tenir compte de l'effet de la combinaison de l'EQPEM 2022 et de l'EQDEM 2022. En effet, pour qu'un enfant puisse être considéré comme répondant à l'EEM4, il faut que le jumelage de la réponse des parents à l'EEM4 soit possible avec la réponse de l'enseignante ou l'enseignant à l'EQDEM. Bien que cette dernière soit une enquête par recensement à caractère obligatoire menée auprès du personnel enseignant des écoles des centres de services scolaires et des commissions scolaires ainsi que des écoles privées, elle comporte tout de même un certain niveau de non-réponse globale. La taille de l'échantillon de départ pour l'EEM4 a donc été établie en fonction des pertes associées à la non-réponse à l'EQDEM.

Ensuite, la taille de l'échantillon et la répartition des enfants de l'EEM4 sont déterminées de manière à tenir compte de l'efficacité d'une étude cas-témoins. Les enfants retenus pour l'EEM4 l'étaient systématiquement pour l'EQPEM. Le plan d'échantillonnage de l'étude est donc un plan stratifié à deux niveaux.

Tout d'abord, on a accordé à chaque école de la base de sondage un statut par rapport à l'EEM4. Une école comptant au moins trois enfants ayant fréquenté la

15. La notion de vulnérabilité est abordée à la section 2.3.

16. Pour plus de détails sur les résultats de l'EQDEM 2017, consulter le document [Enquête québécoise sur le développement des enfants à la maternelle 2017.pdf](#) (Simard et autres 2018).

17. C'est-à-dire le nombre de répondants qu'un échantillon aléatoire simple fournirait.

maternelle 4 ans à temps plein l'année précédente était automatiquement considérée comme visée par l'EEM4. À l'inverse, une école comptant moins de trois enfants ayant fréquenté la maternelle 4 ans à temps plein était considérée comme non visée. Ensuite, pour les enfants des écoles visées par l'EEM4, 396 strates principales ont été construites selon l'école ou le regroupement d'écoles¹⁸ fréquentées par les enfants de maternelle 5 ans en 2021-2022 pour assurer le meilleur « pairage » des enfants ayant fréquenté la maternelle 4 ans à temps plein avec ceux ne l'ayant pas fréquentée à temps plein. Finalement, puisque le tirage des enfants-maternelle 4 ans est fait indépendamment du tirage des enfants-témoins, chacune des 396 strates principales des enfants visés par l'EEM4 est subdivisée en deux (enfants-maternelle 4 ans et enfants-témoins) pour former ainsi 792 strates secondaires.

La répartition de l'échantillon des enfants-témoins entre les strates est choisie de manière que l'atteinte de l'objectif de comparaison avec les enfants-maternelle 4 ans soit optimisée et qu'une description fiable de la sous-population des enfants qui n'ont pas fréquenté la maternelle 4 ans à temps plein en 2021-2022 soit obtenue. Ce choix optimal pour l'EEM4 rend toutefois l'échantillon des enfants-témoins moins efficace pour la description de cette sous-population des enfants de maternelle 5 ans, c'est-à-dire qu'il y a un effet de plan induit par cette sélection qui diminue la précision des estimations de description de cette sous-population. Il faut par ailleurs mentionner que le plan d'échantillonnage de l'EEM4 n'est pas conçu pour répondre à des objectifs régionaux.

Afin que la cible de répondants efficaces déterminée soit atteinte, la taille d'échantillon sélectionné pour la collecte a été ajustée de manière qu'elle tienne compte du taux d'admissibilité et des taux de réponse attendus pour les questionnaires destinés au personnel enseignant (EQDEM) et aux parents (EQPPEM et EEM4). Un échantillon total de 7 070 enfants a donc été sélectionné pour l'EEM4.

Les hypothèses retenues pour les différents facteurs d'ajustement de la taille d'échantillon, selon le type de participation à l'enquête sont présentées dans le tableau 1.2.

Pour répondre au mieux aux objectifs de comparaison avec les enfants-maternelle 4 ans, près de 85 % de l'échantillon des quelque 3 550 enfants-témoins nécessaires à l'étude est réparti proportionnellement au nombre d'enfants-maternelle 4 ans dans la strate secondaire ; le 15 % restant est réparti proportionnellement au nombre d'enfants-témoins dans la strate secondaire. En général, on sélectionne donc plus d'enfants-témoins là où les enfants-maternelle 4 ans sont proportionnellement plus nombreux que là où il y en a moins. L'échantillon des enfants-maternelle 4 ans est quant à lui réparti entre les strates secondaires de manière proportionnelle.

Avant le tirage de l'échantillon, les enfants de chacune des 792 strates ont été triés selon leur IMSE. Ce tri permet d'assurer un certain contrôle de la composition de l'échantillon selon le niveau socio-économique des enfants.

Tableau 1.2

Hypothèses retenues pour le calcul de la taille des échantillons de l'EEM4

Échantillon	Effet de plan	Taux d'admissibilité attendu	Taux de réponse attendu		Nombre de répondants efficaces visé	Taille de l'échantillon sélectionné
			(Enseignants)	(Parents)		
Enfants-maternelle 4 ans	1,00	99 %	90 %	67 %	2 000	3 518
Enfants-témoins	1,45	99 %	90 %	67 %	2 015	3 552
Total	1,37	99 %	90 %	67 %	4 015	7 070

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

18. Certaines écoles ont été regroupées pour avoir un nombre suffisant d'enfants, surtout témoins, dans la strate. Le regroupement d'écoles tient compte des indices de milieu socio-économique (IMSE) similaires, des centres de services scolaires et parfois de la langue d'enseignement. Dans la mesure du possible, le regroupement se fait à l'intérieur d'un même centre de services scolaire ou d'une même commission scolaire.

1.2 Collecte des données

Étant donné que l'EEM4 et l'EQPPEM recherchent en partie une même information auprès des mêmes parents pour les mêmes enfants qui ont fréquenté la maternelle 5 ans en 2021-2022, la collecte de l'EEM4 a été jumelée à celle de l'EQPPEM. Les questions propres à l'EEM4 ont donc été intégrées au questionnaire¹⁹ de l'EQPPEM.

Définitions

Le terme **répondant** fait référence à l'enfant de maternelle 5 ans pour qui un questionnaire a été rempli par le parent.

Le terme **personne répondante** désigne le parent qui a rempli le questionnaire d'enquête de l'enfant.

1.2.1 Prétest

Un prétest téléphonique a été effectué du 12 novembre au 8 décembre 2021 pour l'EQPPEM et l'EEM4. L'échantillon du prétest était composé de 500 enfants, dont 200 enfants visés par l'EEM4, également répartis entre les enfants-maternelle 4 ans et les enfants-témoins pour assurer une bonne représentation de l'EEM4 au prétest. Au total, 268 questionnaires ont été remplis pour le prétest. Parmi ceux-ci, 55 concernaient les enfants-maternelle 4 ans et 44 les enfants-témoins, pour un total de 99 questionnaires de l'EEM4. Le taux de réponse global au prétest était ainsi de 53,7 % pour les enfants-maternelle 4 ans et de 49,5 % pour les enfants-témoins.

Le prétest visait entre autres à vérifier la fluidité du questionnaire, à valider la compréhension des questions ainsi qu'à évaluer la durée de l'entrevue et de certaines sections.

Puisque la durée anticipée du questionnaire était plus longue que prévu, la section sociodémographique a été exclue du prétest et l'échantillon de l'EQPPEM a été divisé en deux groupes selon que les enfants étaient sélectionnées pour l'EQPPEM seulement ou pour l'EQPPEM et l'EEM4. Les personnes répondantes ne passaient donc pas par l'ensemble des sections du questionnaire, ce qui a permis d'alléger le fardeau de réponse durant le prétest, de mettre l'accent sur les questions à l'étude et d'ainsi maximiser le nombre de questionnaires remplis.

À la suite du prétest, quelques modifications ont été apportées au questionnaire de l'enquête : certaines questions ont été reformulées ou supprimées en vue de ramener à environ 30 minutes la durée d'administration du questionnaire. Les 99 questionnaires de l'EEM4 remplis lors du prétest n'ont pas été conservés dans le fichier de données de l'enquête principale.

1.2.2 Consentements

Rappelons que l'un des objectifs de l'EEM4 est d'analyser le lien entre le développement des enfants à la maternelle 5 ans et la fréquentation de la maternelle 4 ans à temps plein, en tenant compte de certaines caractéristiques des enfants ou des familles. Pour ce faire, il est nécessaire de jumeler les données recueillies auprès des parents avec celles fournies par le personnel enseignant de maternelle 5 ans dans le cadre de l'EQDEM 2022.

Les lettres d'information et la documentation envoyées aux centres de services scolaires, aux commissions scolaires et aux écoles dans le cadre de l'EQDEM 2022 présentaient l'EQPPEM 2022 menée auprès de parents d'enfants visés par l'EQDEM en mentionnant notamment qu'on y trouvait des questions sur le programme de maternelle 4 ans.

En effet, l'envoi adressé aux centres de services scolaires, aux commissions scolaires ainsi qu'aux écoles privées dans le cadre de l'EQDEM incluait une lettre expliquant le contexte général de l'EQDEM 2022 et un formulaire de consentement pour les parents des enfants visés. Ce formulaire de consentement autorisait l'ISQ à transmettre les données aux partenaires de l'EQDEM ainsi

19. Pour plus de détails, consulter le [Questionnaire-eqppem 2022.pdf](#).

qu'à l'Université McMaster²⁰ (Ontario) et à jumeler les renseignements recueillis dans le cadre de l'EQDEM avec ceux fournis par les parents dans le cadre de l'EQPPEM, qui incluait l'EEM4, ou avec certaines données administratives liées notamment à l'éducation, à la santé et aux services sociaux. Toutefois, les informations recueillies dans la section du questionnaire propre à l'EEM4 n'ont été transmises qu'au MEQ.

Dans le cadre de l'EQPPEM 2022, les parents ont aussi été informés dès le début du questionnaire du jumelage des données qu'ils fourniraient avec celles recueillies dans le cadre de l'EQDEM auprès de l'enseignante ou de l'enseignant de leur enfant fréquentant la maternelle 5 ans.

Enfin, mentionnons que le partage des données avec les partenaires et avec l'Université McMaster exclut les informations nominatives.

Si au moins un consentement était refusé (que ce soit pour le jumelage ou pour le partage des données), on attribuait aux questionnaires des enfants touchés le statut de non-répondant à l'EQPPEM et à l'EEM4.

1.2.3 Collecte des données de l'enquête principale

Le niveau de développement, mesuré par la vulnérabilité relative²¹ des enfants, provient de l'EQDEM, qui collecte l'information utile à la production d'indicateurs pour l'ensemble des enfants de maternelle 5 ans du Québec. Dans le cadre de l'enquête auprès des parents, les personnes répondantes sont interrogées sur les caractéristiques des enfants, de leur famille et de leur parcours préscolaire. Les questions propres à l'EEM4, posées exclusivement aux parents des enfants sélectionnés pour l'EEM4, sont inspirées de l'*Étude sur l'évaluation du programme Passe-Partout 2017* menée par l'ISQ à la demande du MEQ.

La collecte de l'EEM4 s'est effectuée à la suite de la collecte EQDEM 2022, en même temps que celle de l'EQPPEM 2022. Elle s'est échelonnée sur près de 20 semaines. La

collecte téléphonique a pris fin le 31 juillet 2022 et la collecte Web, le 22 août 2022 ; les deux avaient commencé le 11 avril 2022.

Méthode de collecte

La collecte de données de l'EQPPEM, qui inclut celle de l'EEM4, a été menée auprès des parents une fois la collecte de l'EQDEM pour leur enfant menée auprès de son enseignante ou enseignant. Notons que l'on incluait les parents biologiques ou adoptifs ainsi que les tuteurs et tutrices parmi les personnes clés potentielles pour recueillir l'information concernant l'enfant échantillonné.

Il s'agissait d'une collecte multimode. Les parents avaient ainsi la possibilité de répondre au questionnaire informatisé autoadministré à partir d'une page Web hébergée sur le site sécurisé de l'ISQ (mode IWAO – interview Web assistée par ordinateur) ou encore de prendre part à une entrevue téléphonique assistée par ordinateur menée par une intervieweuse ou un intervieweur (mode ITAO – interview téléphonique assistée par ordinateur). Comme dans les autres enquêtes multimodes menées à l'ISQ, le mode Web était d'abord offert aux parents. Le mode téléphonique était utilisé dans le cas d'enfants d'une même fratrie, lorsque les parents signalaient à l'ISQ ne pas vouloir remplir le questionnaire en ligne, s'ils étaient dans l'impossibilité de le faire ou après plusieurs relances téléphoniques infructueuses effectuées auprès des personnes non-répondantes. Mentionnons qu'avec cette approche, environ 9 personnes répondantes sur 10 (88 %) ont rempli le questionnaire en ligne.

Stratégies de collecte

La collecte pour l'EQPPEM 2022 et l'EEM4 s'est effectuée en plusieurs étapes : identification des enfants de l'échantillon dont le questionnaire à l'EQDEM 2022 a été rempli par l'enseignante ou l'enseignant, envoi d'une lettre d'information aux parents des enfants ciblés pour les inviter à participer à l'EQPPEM 2022 et les inciter à répondre au questionnaire Web en utilisant l'identifiant qui leur a été attribué par l'ISQ. Par la suite, chaque

20. L'Offord Center for Child Studies (OCCS) à l'Université McMaster est propriétaire de l'IMDPE. Il effectue le calcul des scores à l'aide d'une syntaxe protégée par droit d'auteur.

21. Pour plus d'informations sur le concept de vulnérabilité, les domaines de développement et les indicateurs qui y sont associés, consulter le rapport [Enquete-quebecoise-developpement-enfants-maternelle-2022.pdf](#).

parent qui n'avait pas rempli le questionnaire sur le Web a reçu une lettre de relance lui réitérant l'invitation à remplir le questionnaire Web. Deux ou trois semaines plus tard, une deuxième lettre de relance a été envoyée à toutes les personnes qui n'avaient pas encore rempli le questionnaire²². Environ trois semaines avant la fin de la collecte téléphonique, une troisième lettre de relance a été envoyée aux parents qui n'avaient toujours pas rempli le questionnaire.

Après une période jugée suffisante, des rappels téléphoniques ont débuté ; si la personne mentionnait vouloir toujours remplir le questionnaire sur le Web, on demandait le courriel pour effectuer les relances suivantes. Sinon, les parents étaient invités à réaliser l'entrevue au téléphone s'ils étaient disponibles au moment de l'appel ; ils pouvaient également prendre un rendez-vous à un autre moment.

Puisqu'il fallait attendre que le questionnaire de l'EQDEM 2022 soit rempli par l'enseignante ou l'enseignant avant de pouvoir contacter le parent, la collecte pour l'EQPEM 2022 et l'EEM4 a été divisée en cinq lots pour tenir compte de l'avancement de la collecte de l'EQDEM. La taille exacte de chaque lot dépendait du nombre d'enfants ciblés pour l'échantillon parmi les répondants de l'EQDEM à la date de la constitution du lot. Une taille minimale de 1 000 enfants par lot était toutefois fixée pour faciliter la gestion de la collecte. Les enfants d'une même fratrie étaient tous dans le dernier lot.

Les intervieweuses et intervieweurs ont reçu une formation pour se familiariser avec le contexte, les objectifs, le contenu du questionnaire, les codes de résultat et le questionnaire informatisé associés à l'enquête. Les intervieweuses et intervieweurs étaient responsables de fournir une assistance aux personnes participantes désirant remplir le questionnaire en ligne, de procéder à des relances téléphoniques ou de mener des entrevues téléphoniques.

Le nombre de questionnaires remplis a été suivi de près. Plusieurs stratégies ont été mises en place pour que la collaboration des parents des enfants sélectionnés soit favorisée et que la gestion de la collecte soit facilitée. Par exemple, en tout temps, les parents disposaient d'un numéro de téléphone et d'une adresse courriel réservés à l'enquête.

Dans le cas des parents des enfants pour lesquels le MEQ n'avait pas de numéro de téléphone (un peu plus de 10,5 % des personnes échantillonnées dans l'ensemble et 8,6 % pour l'EEM4), des recherches ont été effectuées afin que des numéros valides soient trouvés. Les parents des enfants pour qui le numéro de téléphone était resté inconnu après la recherche ont reçu une lettre leur demandant de contacter l'ISQ.

Soulignons enfin que l'on a tout mis en œuvre pour alléger le fardeau de réponse des parents. Par exemple, les enfants d'une même fratrie avaient été ciblés pour la collecte téléphonique uniquement. Des intervieweurs et intervieweuses s'occupaient de ces questionnaires en particulier. Des instructions précises leur avaient été données, de sorte que les questions dont les réponses seraient les mêmes pour les enfants de la fratrie ne soient posées qu'une seule fois aux parents. Plusieurs tests ont été réalisés avant le début de la collecte pour s'assurer du bon fonctionnement des applications informatiques IWAQ et ITAQ.

22. Dans le cas des questionnaires transmis à partir du 6 juin 2022, une seule lettre de relance a été envoyée, puisque la période de collecte était plus courte.

Questionnaire

Le tableau 1.3 présente les thèmes de chacune des 11 sections du questionnaire de l'EQPPEM 2022 ainsi que les groupes d'enfants auxquels ils s'appliquaient.

Toutes les questions ont été révisées par une réviseuse linguistique et traduites en anglais pour qu'il soit possible de mener des entrevues dans les deux langues.

Pour les deux modes, la durée moyenne pour remplir un questionnaire de l'EEM4 était d'environ 38 minutes.

Tableau 1.3

Contenu du questionnaire de l'EQPPEM 2022 adressé aux parents des enfants sélectionnés selon le groupe de l'EEM4

Section du questionnaire		Groupe d'enfants		
		EEM4 Enfants- maternelle 4 ans	EEM4 Enfants- témoins	EQPPEM
FAM	Description de la famille	X	X	X
SDG	Parcours préscolaire en service de garde	X	X	X
EDU	Activités éducatives	X	X	X
MAT4	Évaluation de la maternelle 4 ans B - Questions posées uniquement aux parents dont les enfants n'avaient pas fréquenté la maternelle 4 ans	-	X	-
	Évaluation de la maternelle 4 ans C - Questions posées uniquement aux parents dont les enfants avaient fréquenté la maternelle 4 ans	X	-	-
MAT5	Passage à la maternelle 5 ans	X	X	X
SAN	Santé de l'enfant	X	X	X
ACT	Activités physiques et utilisation des écrans	X	X	X
PAR	Contexte familial	X	X	X
RES	Lieu de résidence et quartier	X	X	X
OCC	Occupation	X	X	X
COV	Répercussions de la pandémie de COVID-19	X	X	X
DEM	Caractéristiques démographiques et socio-économiques	X	X	X

1.2.4 Résultats de collecte

Le taux de réponse obtenu à l'EQDEM 2022 est supérieur à celui anticipé. En effet, combiné au taux d'admissibilité, le taux de réponse à l'EQDEM a atteint 92,3 % plutôt que 90,0 %, comme prévu dans les hypothèses.

Pour déterminer le statut de réponse à l'EEM4 des enfants, on devait tenir compte de leur statut de réponse à l'EQDEM. Les enfants qui étaient inadmissibles à l'EQDEM étaient obligatoirement considérés comme inadmissibles à l'EEM4. De même, les enfants qui étaient non-répondants à l'EQDEM étaient automatiquement considérés comme non-répondants à l'EEM4. De plus, lorsque le statut d'inadmissibilité ou de non-répondant à l'EQDEM d'un enfant était établi plus tôt, les parents n'étaient pas invités à répondre au questionnaire EQPPM/EEM4, et ce, pour alléger la gestion de la collecte.

Au total, les parents de 3 933 enfants de maternelle 5 ans²³ ont répondu à l'EEM4, ce qui correspond à un taux de réponse pondéré global de 60,0 %. Ce taux prend en considération la non-réponse à l'EQDEM 2022 et la non-réponse à l'EEM4.

Le tableau 1.4 présente le nombre d'enfants échantillonnés ainsi que les taux d'admissibilité et de réponse à l'EEM4. Ces taux tiennent compte de l'admissibilité et de la réponse à l'édition 2022 de l'EQDEM et de l'EQPPM.

Tableau 1.4
Taux d'admissibilité et de réponse pondérés selon le type de participation à l'EEM4

Échantillon	Enfants admissibles	Taux d'admissibilité ¹	Taux de réponse ²	Répondants
	n	%		n
Enfants-maternelle 4 ans	3 461	98,4 %	55,0 %	1 911
Enfants-témoins	3 445	96,3 %	61,6 %	2 022
Total	6 906	96,8 %	60,0 %	3 933

1. Correspond au rapport du nombre pondéré d'enfants admissibles au nombre d'enfants de la base de sondage.

2. Correspond au rapport du nombre pondéré d'enfants pour qui un questionnaire à l'EEM4 a été rempli au nombre pondéré d'enfants admissibles.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

1.3 Traitement des données

La présente section aborde les éléments de validation des données recueillies lors de la collecte, la méthode de pondération retenue ainsi que l'analyse et le traitement de la non-réponse partielle.

1.3.1 Validation et saisie

Le recours à la collecte assistée par ordinateur, que ce soit la collecte Web ou téléphonique, permet une saisie automatique des réponses fournies par les parents. La très grande majorité des questions sont autocodées, c'est-à-dire que l'intervieweuse ou l'intervieweur ou encore le parent n'a qu'à cocher le ou les choix de réponses appropriés ou à les sélectionner dans le menu déroulant.

Entièrement informatisées, les collectes téléphoniques et Web permettent d'effectuer plusieurs validations de base au moment de remplir le questionnaire, notamment une vérification du respect des choix de réponses pour les questions fermées ou de l'adéquation des sauts de section pour les blocs de questions ne concernant qu'un sous-groupe de répondants. Bien que les questionnaires informatisés aient été amplement testés avant le début de la collecte, tous les codes de réponse et les sauts de section de même que la cohérence des réponses d'un même questionnaire ont été validés *a posteriori*.

23. Ce nombre comprend tous les enfants visés par l'étude, dont les EHDA qui sont exclus des résultats présentés dans ce rapport et dans le recueil statistique.

De plus, plusieurs questions comportent un choix de réponses ouvert « Autre » qui offre la possibilité de préciser une réponse différente de celles des catégories proposées. Dans ces cas, les réponses aux questions ont été examinées. Si elles correspondaient à l'un des choix de réponses offerts, la réponse était reclassée. Dans le cas contraire, la réponse était laissée dans le choix « Autre ».

En ce qui a trait à la validation des réponses fournies à chaque question, les valeurs maximales et minimales aux questions numériques ouvertes ont été examinées ; toutes les valeurs jugées aberrantes ont été remplacées par des valeurs manquantes.

1.3.2 Pondération de description

Stratégie de pondération de description

La pondération consiste à attribuer à chaque enfant pour qui un questionnaire a été rempli par le parent un poids statistique, c'est-à-dire une valeur qui correspond au nombre d'enfants qu'il « représente » dans la population visée. Ce poids tient compte de la probabilité de sélection des enfants échantillonnés, prédéterminée par le plan d'échantillonnage de l'EEM4, ainsi que de la non-réponse à l'EQDEM 2022 et à l'EEM4. La pondération est essentielle pour qu'une inférence adéquate soit faite à cette population, puisque cette dernière n'a pas été observée dans sa totalité. En effet, l'EEM4 est menée à l'aide d'un échantillon et, comme toutes les enquêtes, elle n'est pas exempte de non-réponse.

De plus, les enfants ayant fréquenté la maternelle 4 ans à temps plein ont été suréchantillonnés en vue de garantir un nombre de répondants suffisants dans ce groupe pour les besoins de comparaison de l'EEM4. Cela affecte les probabilités de sélection et fait en sorte que la répartition de l'échantillon selon la fréquentation de la maternelle 4 ans à temps plein ne représente pas bien celle de la population visée. D'où la nécessité de tenir compte de la pondération pour la production des estimations.

La non-réponse a touché inégalement certains groupes d'enfants de l'échantillon. Cette situation a conduit à un ajustement de la pondération en vue de réduire les biais potentiels de non-réponse totale dans les estimations. Cette non-réponse peut survenir à différents moments du processus de collecte puisque les centres de services scolaires, les commissions scolaires et les écoles privées

peuvent refuser le jumelage des données recueillies auprès du personnel enseignant avec celles fournies par les parents. De leur côté, les parents peuvent également refuser que l'enseignante ou l'enseignant remplisse le questionnaire de leur enfant ou encore refuser de participer aux enquêtes auprès des parents. Il s'agit dans tous ces cas d'une non-réponse à l'EEM4.

La pondération des répondants doit aussi offrir des ajustements pour la couverture de la base de sondage pour améliorer la qualité des estimations produites concernant certains totaux de population connus ou estimés.

La stratégie de pondération établie par l'ISQ pour l'EEM4 tient donc compte de la probabilité de sélection des enfants échantillonnés, de l'ajustement de la non-réponse totale à l'EQDEM et à l'EEM4 ainsi que d'un ajustement aux totaux de la population pour la vulnérabilité, l'âge et le sexe des enfants.

Comme l'estimation de la variance et l'élaboration des tests sont faites à partir de poids d'autoamorçage (*bootstrap*), ceux-ci sont calculés en même temps que les poids dans toutes les étapes de pondération.

Description des étapes de pondération

A – Poids de base

La première étape de la pondération consiste à calculer le poids de départ, aussi appelé « poids de base », attribué à chaque enfant en vertu de sa probabilité de sélection. Chacun des 7 070 enfants échantillonnés pour l'EEM4 reçoit donc un poids équivalent à l'inverse de la probabilité de sélection dans la strate à laquelle il appartient. Les poids de base des enfants de l'échantillon varient ainsi d'une strate à l'autre. Pour chacune des strates, le poids de base correspond au quotient du nombre d'enfants dans la base de sondage par le nombre d'enfants sélectionnés dans l'échantillon.

Probabilité de sélection :

$$P_{strate} = \frac{\text{Nombre d'enfants de la strate dans l'échantillon}}{\text{Nombre d'enfants de la strate dans la population}}$$

Poids de base :

$$\begin{aligned} W_{strate} &= \frac{1}{P_{strate}} \\ &= \frac{\text{Nombre d'enfants de la strate dans la population}}{\text{Nombre d'enfants de la strate dans l'échantillon}} \end{aligned}$$

B – Ajustement pour la non-réponse à l'EQDEM

Pour être répondant à l'EEM4, il faut avoir un questionnaire rempli par l'enseignante ou l'enseignant à l'EQDEM 2022. Les poids de base de l'échantillon sont alors ajustés pour que les poids des enfants de l'échantillon pour qui un questionnaire à l'EQDEM 2022 a été rempli correspondent à la somme des poids de base des enfants admissibles à l'EEM4.

Puisque le taux de réponse à l'échelle des écoles et du personnel enseignant est élevé (plus de 90 %) ²⁴ et que la probabilité d'avoir un questionnaire rempli à l'EQDEM 2022 est indépendante du fait d'être sélectionné pour l'EEM4, l'ajustement est effectué à partir des caractéristiques des enfants associées à la probabilité de répondre à l'EQDEM. Les groupes d'enfants présentant des caractéristiques semblables et des probabilités de répondre à l'EQDEM semblables (classes de non-réponse) utilisés pour cette étape d'ajustement proviennent alors de la pondération de l'EQDEM. Les variables considérées pour former les classes de non-réponse à l'EQDEM sont la région sociosanitaire, le réseau d'enseignement (public/privé), le fait d'être codé EHDAA ou DAA selon les fichiers du MEQ, la langue maternelle de l'enfant et le lieu de naissance des parents.

Pour chaque sous-population, l'ajustement s'effectue ensuite à l'intérieur de chacun de ces groupes : le poids de base des enfants admissibles à l'EEM4 pour qui un questionnaire à l'EQDEM 2022 est disponible est multiplié par l'inverse du taux de réponse (pondéré) dans le groupe homogène de pondération. Le premier ajustement pour la non-réponse est ainsi terminé.

C – Ajustement pour la non-réponse à l'EEM4

La non-réponse observée à l'échelle des enfants correspond essentiellement au refus des parents de participer à l'EEM4. Dans quelques rares cas, elle correspond au refus du jumelage des renseignements fournis par l'enseignante ou l'enseignant avec ceux de l'enquête auprès des parents (EQPPM et donc EEM4). Puisque seuls les enfants « répondants » à l'EEM4 servent à produire les estimations, on leur accorde un poids ajusté pour qu'ils « représentent » ceux pour qui nous n'avons pas de

questionnaire rempli. Pour ce faire, on analyse d'abord les caractéristiques des enfants qui sont associées à la probabilité que le parent réponde au questionnaire et aux paramètres à estimer. On utilise la méthode du score ²⁵ pour déterminer les groupes homogènes de pondération, c'est-à-dire les sous-groupes de l'échantillon à l'intérieur desquels la propension à répondre est assez uniforme. Pour l'ajustement des poids à la non-réponse au questionnaire des parents, on a formé deux classes de pondération à partir de la région administrative, du réseau d'enseignement et de l'indice de défavorisation sociale. Les probabilités de réponse au questionnaire sont estimées au moyen d'un modèle de régression logistique. Cette approche permet de réduire au minimum les biais potentiels dus à la non-réponse.

Pour chaque sous-population, les poids ajustés pour la non-réponse à l'EQDEM 2022 des enfants admissibles répondants à l'EEM4 sont donc multipliés par l'inverse du taux de réponse (pondéré) dans le groupe de pondération formé pour l'ajustement à la non-réponse de l'EEM4.

D – Poststratification

La dernière étape de la pondération est la poststratification. Cet ajustement permet de s'assurer que la somme des poids de tous les répondants de l'enquête correspond à la population des enfants de maternelle 5 ans visée par l'enquête, notamment pour certaines caractéristiques associées aux mesures d'intérêt.

Comme mentionné en introduction du rapport, l'analyse de la relation entre la fréquentation de la maternelle 4 ans à temps plein et le développement de l'enfant à la maternelle 5 ans comme mesuré dans l'EQDEM fait partie des objectifs de l'EEM4. L'EQDEM 2022, qui est basée sur un recensement des enfants inscrits à la maternelle 5 ans en 2021-2022, offre des estimations très précises du niveau de développement relatif de ces enfants à leur entrée à l'école. En effet, des totaux estimés peuvent être produits à partir de l'information obtenue du personnel enseignant au sujet des enfants répondants de l'EQDEM visés par l'EEM4. Des estimations de l'EQDEM résultant du croisement entre un indicateur de vulnérabilité

24. Pour plus de détails sur les résultats de collecte de l'EQDEM, consulter le rapport [Méthodologie de l'Enquête québécoise sur le développement des enfants à la maternelle \(EQDEM\) 2022](#).

25. Pour plus de renseignements sur la méthode du score, consulter Haziza et Beaumont (2007) et Eltinge et Yansaneh (1997).

(vulnérable dans au moins un domaine de développement ou non²⁶), l'âge²⁷ et le sexe des enfants ont donc servi à la poststratification des répondants de l'EEM4.

Cette étape d'ajustement des poids des répondants aux totaux de la population selon les caractéristiques d'âge, de sexe et d'indice de vulnérabilité conclut la pondération de description.

1.3.3 Non-réponse partielle

Taux de non-réponse partielle

La pondération permet de réduire les biais associés à la non-réponse totale, mais ne comporte pas d'ajustement pour la non-réponse partielle. Celle-ci se produit lorsque pour une question, les réponses de certains répondants à l'enquête sont manquantes ou invalides. Il est connu qu'une non-réponse partielle importante, au même titre que la non-réponse totale, peut entraîner certains biais dans les estimations s'il s'avère que les non-répondants (partiels) présentent des caractéristiques différentes de celles des répondants et, de surcroît, si ces caractéristiques sont liées aux mesures d'intérêt.

Pour chaque question, le taux de non-réponse partielle pondéré se définit comme le rapport du nombre pondéré d'enfants pour lesquels la réponse à la question est manquante ou invalide au nombre pondéré d'enfants admissibles à la question. Pour chaque indicateur (construit à partir de plusieurs questions), le taux de non-réponse partielle correspond au rapport du nombre pondéré d'enfants pour lesquels l'indicateur ne peut être construit au nombre d'enfants admissibles à l'indicateur. Plus ce taux est élevé, plus le risque de biais induit par la non-réponse partielle est grand.

Étant donné que les données de l'EEM4 ont été collectées en même temps que celles de l'EQPPEM, une seule analyse de la non-réponse partielle a été conduite pour les deux études lors du traitement de la non-réponse

partielle de l'EQPPEM. Les taux de non-réponse partielle présentés ici sont ceux calculés pour l'EQPPEM 2022. Compte tenu de la précision attendue pour les estimations, on a émis l'hypothèse qu'une non-réponse partielle inférieure à 5 % a une incidence négligeable sur les estimations, sauf lorsque cette non-réponse partielle est caractérisée.

D'une part, le taux de non-réponse partielle pour la majorité des questions et des indicateurs est en dessous du seuil critique de 5 %, ce qui signifie qu'il ne devrait pas y avoir d'incidence sur les conclusions des analyses effectuées une fois la précision des estimations prise en compte. D'autre part, les questions relatives à l'âge, en années et en mois, au début ou à la fin d'un événement présentent un taux de non-réponse partielle plus élevé que le seuil critique. Cela pourrait s'expliquer par le fait que les parents peuvent avoir plus de difficulté à se souvenir des détails recherchés (en années et en mois). Cette non-réponse partielle fait augmenter le risque d'avoir des taux de non-réponse partielle relativement élevés pour les indicateurs construits à partir de ces questions.

Âge au moment de la séparation des parents

La question portant sur l'âge en années et en mois de l'enfant à la séparation des parents présente des taux de non-réponse partielle supérieurs au seuil critique, ce qui fait augmenter le risque d'avoir un taux de non-réponse partielle relativement élevé pour l'indicateur composite « Enfants vivant actuellement avec leurs deux parents ou dont les parents sont séparés (selon l'âge qu'ils avaient au moment de la séparation) »²⁸ (SEPARATION_4CAT), comparativement aux autres indicateurs. En effet, le taux de non-réponse partielle est d'environ 6 % pour le nombre d'années de l'enfant au moment de l'événement (FAM_Q5_AN) et de 17 % pour le nombre de mois (FAM_Q5_MO), ce qui induit une non-réponse partielle caractérisée selon le type de famille pour l'indicateur « Enfants vivant actuellement avec leurs deux parents ou dont les parents sont séparés (selon l'âge qu'ils avaient

26. Pour plus de détails sur le concept de vulnérabilité, les domaines de développement et les indicateurs qui y sont associés, consulter le [rapport de l'EQDEM 2022](#).

27. Six groupes d'âge ont été construits à partir de la date de naissance des enfants : Après septembre 2016 (plus jeune) / Juillet, août et septembre 2016 / Avril, mai et juin 2016 / Janvier, février et mars 2016 / Octobre, novembre et décembre 2015 et Avant octobre 2015 (plus vieux).

28. Cet indicateur est construit par la combinaison de l'indicateur du type de famille (intacte, recomposée – parents ensemble, recomposée – parents séparés et monoparentale) et des variables liées à l'âge de l'enfant au moment de la séparation des parents (FAM_Q5_AN et FAM_Q5_MO).

au moment de la séparation) ». La non-réponse partielle de l'indicateur « Enfants vivant actuellement avec leurs deux parents ou dont les parents sont séparés (selon l'âge qu'ils avaient au moment de la séparation) » est donc d'environ 3 %, mais les non-répondants partiels sont tous des enfants dont les parents sont séparés. Le risque de biais ne peut donc pas être écarté. Cet indicateur a été traité à l'étape de l'imputation, abordée dans la prochaine section du rapport.

Âge au début et à la fin de la fréquentation d'un service de garde

Par ailleurs, d'autres questions portant sur l'âge au début et à la fin de la fréquentation de chaque service de garde présentent une non-réponse partielle inférieure au seuil critique, mais tout de même supérieure à la plupart des questions. Le taux de non-réponse partielle le plus élevé est d'environ 1,5 % pour les questions sur l'âge au début et à la fin de la fréquentation de chaque service de garde. Bien que cette non-réponse partielle ne soit pas associée à un risque de biais préoccupant, un traitement d'imputation a été effectué pour assurer une cohérence entre tous les indicateurs de fréquentation des services de garde.

Revenu du ménage

Les indicateurs utilisant le revenu du ménage (avant impôts et autres retenues), aussi traités dans la section sur l'imputation, présentent un taux de non-réponse partielle au-dessus du seuil critique soit 5,85 %.

Raisons de non-fréquentation d'un service de garde

La question portant sur les principales raisons de la non-fréquentation d'un service de garde avant l'entrée à la maternelle 5 ans (SDG_Q1a) présente aussi un taux de non-réponse partielle supérieur au seuil critique, soit 14,7 %. Cette non-réponse est caractérisée selon le lieu de naissance des parents ou du parent seul. En effet, le taux de non-réponse partielle est plus élevé chez les enfants dont les parents, ou le parent seul, sont nés à l'extérieur du Canada que chez les enfants dont au moins un des parents est né au Canada. Comme les proportions d'enfants qui n'ont pas fréquenté un service de garde pour diverses raisons sont significativement différentes entre ces deux groupes d'enfants, l'EQPPEM 2022 et l'EEM4 pourraient légèrement surestimer la proportion d'enfants qui n'ont pas fréquenté un service de garde car le parent a fait le choix de rester à la maison avec

son enfant, et sous-estimer le fait qu'un enfant n'a pas fréquenté de service de garde car aucune place n'était disponible.

Signalons enfin que la variable relative aux principales raisons de la non-fréquentation d'un service de garde avant l'entrée à la maternelle 5 ans n'a pas été imputée. Les risques de biais ne représentaient pas un réel problème pour les fins exploratoires auxquelles ces données devaient servir.

Imputation

Lorsque le risque de biais associé à la non-réponse partielle est jugé préoccupant ou touche une variable de croisement, l'imputation des données manquantes est une solution souvent mise de l'avant. Pour une variable présentant une non-réponse partielle, l'imputation consiste à attribuer une valeur à un individu en remplacement d'une donnée manquante ou invalide au moyen des renseignements disponibles à son sujet dans la base de sondage ou dans les données du questionnaire d'enquête lui-même.

Pour l'EQPPEM 2022 et l'EEM4, on a procédé à l'imputation des données sur le revenu du ménage avant impôts et autres retenues, l'âge au début et à la fin de la fréquentation d'un service de garde, l'âge auquel l'enfant a été initié à la lecture par un adulte, l'âge de l'enfant à la séparation des parents ainsi que le genre du parent et de la conjointe ou du conjoint pour les non-répondants partiels. L'imputation du revenu a été effectuée au moyen de la méthode déductive et aléatoire par donneur (*hot deck*) et par modélisation. Pour les autres questions, on a utilisé la méthode aléatoire par donneur (*hot deck*) (Lohr 2010 ; Statistique Canada 2003) pour l'imputation.

Revenu du ménage

Les questions sur le revenu du ménage servent à la construction de l'indicateur de niveau de revenu du ménage (en quatre catégories) utilisé comme variable de croisement. L'indicateur est calculé à partir de la question sur le revenu total du ménage avant impôts et autres retenues. La personne répondante avait la possibilité de fournir l'information sur le revenu soit sous forme de valeur exacte (DEM_Q8), bien qu'approximative, soit sous forme de catégorie en indiquant la tranche de revenu (DEM_Q9 et DEM_Q9A à DEM_Q9F). Parfois, la déclaration sous forme de catégorie de revenus n'est pas

assez précise pour permettre l'attribution d'une valeur à l'indicateur de niveau de revenu. Lorsque ces cas sont ajoutés à ceux pour lesquels aucune réponse n'a été fournie au sujet du revenu du ménage, le taux pondéré de non-réponse partielle atteint près de 6 %. L'imputation déductive et aléatoire par donneur (*hot deck*) basée sur l'information disponible sur la catégorie du revenu a permis de réduire cette non-réponse partielle à 3,9 %.

Les cas pour lesquels aucune information n'était disponible ont été imputés par modélisation (régression), c'est-à-dire qu'on a attribué aux non-répondants partiels une valeur de revenu probable à partir des caractéristiques qui lui sont potentiellement corrélées telles que le nombre de personnes du ménage, l'âge et le sexe des parents, le niveau de scolarité du ou des parents, la région administrative, l'occupation du ou des parents, le revenu en tranche et l'indice de défavorisation matérielle et sociale. Il a ainsi été possible d'obtenir un indicateur de niveau de revenu pour chaque enfant de maternelle 5 ans.

Âge au moment de la séparation des parents

L'indicateur composite « Enfants vivant actuellement avec leurs deux parents ou dont les parents sont séparés (selon l'âge qu'ils avaient au moment de la séparation) » en quatre catégories était calculé pour les enfants qui vivaient avec leurs deux parents biologiques ou adoptifs, dans une famille monoparentale ou dans une famille recomposée – parents séparés²⁹. Les enfants ne vivant avec aucun des parents biologiques ou adoptifs étaient exclus de cet indicateur puisque la qualité de la réponse aux questions utilisées dans la construction de l'indicateur ne pouvait être garantie. En effet, il est possible que pour un enfant vivant dans une famille d'accueil par exemple, l'âge à la séparation déclaré corresponde à l'âge que l'enfant avait au moment de la séparation des parents de la famille d'accueil alors qu'il y est arrivé après la séparation ou encore qu'il corresponde à l'âge de la séparation de ses parents biologiques alors que l'enfant ne vivait plus avec eux. L'âge à la séparation (en années et en mois) a donc été imputé par donneur pour les enfants vivant dans une famille monoparentale ou dans une famille avec parents biologiques ou adoptifs séparés, ce qui a permis d'obtenir un indicateur composite « Enfants qui

vivent avec leurs deux parents ou l'âge qu'ils avaient lors de la séparation de leurs parents » pour tous les enfants concernés par l'indicateur.

Âge au début et à la fin de la fréquentation d'un service de garde

L'âge au début et à la fin de la fréquentation de chacun des services de garde a servi à la construction de plusieurs indicateurs tels que l'âge au début de la fréquentation du premier service de garde en mois (AGE_1SDG_MOIS), la durée cumulative en mois en services de garde en cinq catégories (DUREE_GARDE_TOT_5CAT) et l'intensité de fréquentation du service de garde (INTENSITE_GD). Comme ces questions ont été utilisées dans le calcul de plusieurs indicateurs, elles ont aussi été imputées, de sorte qu'il y avait pour chaque milieu de garde fréquenté par un enfant un âge au début et à la fin de la fréquentation du milieu de garde pour les composantes année et mois.

Par ailleurs, il est important de mentionner que l'imputation a pour effet d'augmenter artificiellement le nombre de répondants et répondantes, ce qui entraîne une sous-estimation de la variance des estimations produites. Il est alors recommandé d'interpréter avec prudence les tests de différences statistiques dont le seuil observé est très près du seuil de signification fixé, en particulier pour les indicateurs utilisant le revenu du ménage et l'indicateur composite « Enfants qui vivent avec leurs deux parents ou l'âge qu'ils avaient lors de la séparation de leurs parents », qui présentent des taux d'imputation plus importants.

29. La famille recomposée – parents séparés se définit comme toute famille constituée d'un des parents de l'enfant de maternelle 5 ans et d'une nouvelle conjointe ou d'un nouveau conjoint.

1.4 Précisions des estimations et tests statistiques

1.4.1 Précisions des estimations

Les données tirées d'enquêtes statistiques comportent des erreurs dites d'échantillonnage, attribuables au fait que seule une partie des unités de la population est sélectionnée pour participer à l'enquête et qu'il y a de la non-réponse globale. Ces erreurs se répercutent sur la précision des estimations produites. La complexité du plan d'échantillonnage, qui n'est pas un plan aléatoire simple, et la non-réponse globale à l'EQDEM 2022 et à l'EEM4 influencent la précision des estimations. Il est nécessaire de mesurer la précision de chaque estimation et d'en tenir compte dans l'interprétation des résultats inférés à la population visée.

L'estimation de la variance et les tests statistiques sont effectués à l'aide de poids d'autoamorçage, aussi appelés « poids *bootstrap* », qui facilitent la mesure de la variabilité des statistiques non linéaires. Une série de 500 poids d'autoamorçage a été créée afin que l'estimation de la variance et les tests statistiques tiennent adéquatement compte du plan d'échantillonnage complexe et de tous les ajustements de non-réponse et de poststratification apportés à la pondération. Pour ce faire, on a d'abord sélectionné 500 échantillons d'autoamorçage selon un plan d'échantillonnage avec remise à partir de l'échantillon initial. Ensuite, toutes les étapes de la pondération ont été appliquées à chacun des échantillons, ce qui a généré 500 poids d'autoamorçage (Rust et Rao 1996). Ces poids sont utilisés dans l'estimation de la variance et dans les tests statistiques à l'aide de logiciels tels que SAS et SUDAAN.

Les mesures de précision présentées dans le cadre de cette étude sont l'intervalle de confiance et le coefficient de variation (CV). Les estimations dont le CV est supérieur à 15 % sont annotées dans les tableaux s'il y a lieu³⁰.

Le CV se calcule comme suit :

$$CV = \frac{\sqrt{\text{variance de l'estimation}}}{\text{estimation}}$$

1.4.2 Tests statistiques

Pour l'analyse des données de l'EEM4 présentée dans le recueil statistique, l'ISQ a d'abord effectué un test statistique d'indépendance du khi-deux³¹ pour déceler un lien possible entre l'indicateur analysé et une variable de croisement correspondant généralement à une caractéristique socio-économique ou autre. Les tests dont le seuil de signification observé est inférieur à 5 % montrent une association significative.

Des tests d'égalité de proportions ont ensuite été réalisés pour que les associations révélées par les tests du khi-deux significatifs soient mieux comprises. Notons qu'un écart important entre deux proportions n'est pas nécessairement significatif d'un point de vue statistique. De même, deux intervalles disjoints n'indiquent pas nécessairement une différence significative des statistiques estimées. Tout dépend de la variabilité associée à chaque estimation. La comparaison est basée sur la construction d'une statistique de Wald à partir de la différence des transformations *logit*³² des statistiques estimées. La variance et la covariance de la différence des *logits* et pour la construction des intervalles de confiance sont calculées à l'aide des poids d'autoamorçage. Le seuil de signification pour qu'un écart soit considéré comme significatif est de 5 % pour l'EEM4.

Pour apprécier la qualité des tests de comparaison des estimations, on peut leur associer une puissance statistique. Celle-ci correspond à la probabilité de conclure à un écart significatif sur la base des données lorsque les estimations comparées sont réellement différentes. Elle dépend de la taille d'échantillon : lorsque les tailles d'échantillon diminuent, les écarts permettant de conclure à une différence statistiquement significative augmentent.

30. Pour plus de renseignements sur la présentation des CV, consulter la section 1.5.

31. 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 (Rao et Scott, 1981). Plus précisément, c'est la statistique F correspondant à cette correction du test du khi-deux qui est utilisée dans les analyses.

32. La transformation $\text{logit} = \log\left(\frac{p}{1-p}\right)$ 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).

1.5 Présentation des résultats du recueil statistique

La présente section fournit des indications générales sur l'ensemble des proportions présentées dans le recueil statistique.

En général, les estimations de proportions sont arrondies à une décimale près dans les tableaux du recueil. En raison de ces arrondissements, la somme des proportions de certains tableaux peut être légèrement différente de 100 %.

Les estimations dont le CV est inférieur ou égal à 15 % sont suffisamment précises et sont présentées sans indication à cet égard (Institut de la statistique du Québec). La présentation de celles dont le CV est supérieur à 15 % est basée sur le système de classification du tableau 1.5. De plus, puisque les résultats présentés sont basés sur un échantillon, ils comprennent un certain degré d'imprécision. De ce fait, lors de leur interprétation, des expressions appropriées devraient être utilisées pour rendre compte du fait que les statistiques fournies sont des estimations et non des valeurs exactes. On dira par exemple « Environ 89 % des enfants qui fréquentaient la maternelle 5 ans... » plutôt que « 88,8 % des enfants qui fréquentaient la maternelle 5 ans... ».

Pour les analyses bivariées du recueil statistique, en présence d'une association significative selon le test du khi-deux, la lettre « a » en exposant ajoutée aux statistiques présentées indique les proportions d'une catégorie de la variable d'analyse qui diffèrent significativement entre les enfants-maternelle 4 ans et les enfants-témoins. Une même lettre révèle un écart significatif entre deux catégories.

Il peut arriver que deux proportions qui semblent différentes ne le soient pas d'un point de vue statistique, en raison notamment du petit nombre d'individus sur lequel est basée l'estimation. 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.

La valeur p indique le seuil observé qui correspond à la probabilité que la statistique du test prenne une valeur au moins aussi grande que la valeur observée de cette statistique lorsque l'hypothèse nulle est vraie (Hines et autres, 2003). Puisque les tests sont effectués au seuil de 5 %, une valeur p inférieure à 0,05 signale que le test du khi-deux est significatif, ce qui veut dire que pour au moins une catégorie de la variable d'analyse, la proportion chez les enfants-maternelle 4 ans est statistiquement différente de la même proportion chez les enfants-témoins.

Tableau 1.5

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

Précision de l'estimation	Signe distinctif	Réserves à mentionner lors de la diffusion
$CV \leq 15 \%$	(Aucun signe)	Estimation diffusée sans mention
$15 \% < CV \leq 25 \%$	*	Estimation à interpréter avec prudence
$CV > 25 \%$	**	Estimation à utiliser avec circonspection ; fournie à titre indicatif seulement

2 Aspects méthodologiques propres à l'étude comparative

L'EEM4 est une enquête cas-témoins qui compare la sous-population des enfants-maternelle 4 ans à la sous-population des enfants-témoins. Or, ces deux sous-populations présentent des caractéristiques différentes pour ce qui est de leur taille et de la répartition des enfants à l'intérieur des strates principales formées par les écoles ou les regroupements d'écoles. Il est alors nécessaire de les standardiser en vue d'optimiser la comparaison du développement et de la transition à la maternelle 5 ans des enfants-maternelle 4 ans et des enfants-témoins. Des contrôles ont donc été effectués lors du plan d'échantillonnage ainsi que lors du traitement postcollecte des données pour rendre la sous-population des enfants-témoins le plus semblable possible à celle des enfants-maternelle 4 ans.

La deuxième partie du rapport présente les méthodes de standardisation des deux sous-populations ainsi que les éléments méthodologiques de l'analyse comparative. Elle décrit d'abord la population visée pour la comparaison, l'échantillon et la pondération de comparaison. Pour ce qui est des éléments d'analyse, elle introduit les indicateurs à l'étude, soit les différents indicateurs de développement ainsi que l'indicateur de transition à la maternelle 5 ans³³. Pour finir, elle détaille la démarche de modélisation et le mode adopté pour la présentation des résultats d'analyse de comparaison.

2.1 Plan d'échantillonnage pour la comparaison

La section 1.1 a fourni les renseignements pertinents sur la population visée et les échantillons sélectionnés pour l'EEM4 en général. La présente section s'intéresse à la description des éléments particuliers qui ont permis de contrôler la population visée et l'échantillon en vue d'une comparaison optimale³⁴.

2.1.1 Population visée pour la comparaison

Comme mentionné précédemment, la population visée a été divisée en deux sous-populations que l'on veut comparer pour l'analyse mettant en lien le développement des enfants et leur transition à la maternelle 5 ans avec la fréquentation de la maternelle 4 ans à temps plein :

- les enfants-maternelle 4 ans qui constituent la sous-population d'intérêt pour la comparaison ;
- les enfants-témoins qui constituent la sous-population avec laquelle la comparaison est effectuée.

En vue de réduire les biais possibles dans la comparaison, la sous-population des enfants-témoins devrait être le plus semblable possible à la sous-population des enfants-maternelle 4 ans aux fins de comparaison. Pour ce faire, des exclusions supplémentaires ont été effectuées dans la population visée telle que définie lors du plan d'échantillonnage (section 1.1.1). Les enfants se trouvant en situation d'admissibilité exceptionnelle ont

33. L'indicateur de transition à la maternelle 5 ans est décrit à la section 2.3.2.

34. Des représentations schématiques qui combinent le plan d'échantillonnage de description et celui pour la comparaison sont fournies dans l'annexe II (global) et dans l'annexe III (par sous-population).

été exclus de la population visée pour la comparaison puisqu'ils n'auraient pas pu fréquenter la maternelle 4 ans en 2020-2021. Il s'agit précisément des enfants en situation de reprise de la maternelle 5 ans, de passage précoce à la maternelle 5 ans lié à la démonstration d'aptitudes particulières ou encore d'une entrée à un âge inhabituel associée à une faible densité de population ou à une situation familiale particulière. Ces exclusions correspondent à un peu moins de 2 % des enfants de maternelle 5 ans fréquentant l'une des écoles admissibles à l'étude en 2021-2022. Ces enfants en situation d'admission exceptionnelle se trouvent quasi exclusivement dans la sous-population des enfants-témoins.

De plus, la sous-population d'intérêt a été retenue comme sous-population de référence pour la standardisation des distributions des enfants des deux sous-populations.

La taille et la distribution de la sous-population des enfants-maternelle 4 ans sert donc de référence pour le traitement des données des deux groupes d'enfants de l'échantillon de comparaison, c'est-à-dire que l'on considère que la taille et la distribution de la sous-population des enfants-témoins correspondent à celles de la sous-population des enfants-maternelle 4 ans. Il s'agit d'effectuer un contrôle de la distribution des enfants-témoins entre écoles ou regroupements d'écoles (milieu scolaire) et un contrôle des caractéristiques pertinentes pour les analyses comparatives de l'EEM4 (milieu familial et environnemental).

La population de référence pour la comparaison compte 10 869 enfants de maternelle 5 ans³⁵. À titre indicatif, le tableau 2.1 présente la répartition en fonction du sexe et de la date de naissance des enfants-maternelle 4 ans.

Tableau 2.1

Population de référence pour la comparaison, enfants-maternelle 4 ans selon le sexe et la date de naissance, Québec, 2021-2022

Sexe	Date de naissance	Nombre d'enfants	Proportion (%)
Masculin	Après juin 2016	1 234	11,35
	Entre avril et juin 2016	1 297	11,93
	Entre janvier et mars 2016	1 331	12,25
	Avant décembre 2015	1 691	15,56
	Sous-total	5 553	51,09
Féminin	Après juin 2016	1 168	10,75
	Entre avril et juin 2016	1 231	11,33
	Entre janvier et mars 2016	1 267	11,66
	Avant décembre 2015	1 650	15,18
	Sous-total	4 316	48,91
Total		10 869	100

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

35. Il y avait 10 870 enfants-maternelle 4 ans dans la population visée (voir section 1.1.1); on en a retiré un pour la comparaison en raison de sa situation d'admissibilité exceptionnelle.

2.1.2 Taille et répartition de l'échantillon et des répondants pour la comparaison

La sélection de l'échantillon a été effectuée avant la détermination de la population visée pour la comparaison. Les enfants sélectionnés pour l'EEM4 qui ne se retrouvaient

pas dans la population visée pour la comparaison ont été retirés de l'échantillon pour la comparaison. Le tableau 2.2 présente le nombre d'enfants sélectionnés et le nombre de répondants de l'échantillon EEM4 admissibles à la comparaison ainsi que l'effet plan et le nombre de répondants efficaces. Le tableau 2.3, quant à lui, présente leur répartition selon le sexe et la date de naissance.

Tableau 2.2

Nombre d'enfants échantillonnés et nombre d'enfants répondants admissibles à la comparaison, Québec, 2022

Échantillon	Admissibles à la comparaison		Effet de plan pour la comparaison	Nombre de répondants efficaces ¹
	Nombre d'enfants échantillonnés	Nombre d'enfants répondants		
Enfants-maternelle 4 ans	3 460	1 910	1,00	1 910
Enfants-témoins	3 379	1 992	1,09	1 835
Total	6 839	3 902		

1. C'est-à-dire l'équivalent en termes de répondants qui auraient été obtenus au moyen d'un échantillon aléatoire simple.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

Tableau 2.3

Répartition des enfants répondants par types de participation selon le sexe et le groupe d'âge

Répondants	Sexe	Date de naissance	Nombre d'enfants	
Maternelle 4 ans	Masculin	Après juin 2016	210	
		Entre avril et juin 2016	232	
		Entre janvier et mars 2016	231	
		Avant décembre 2015	291	
		Sous-total	964	
		Féminin	Après juin 2016	193
	Entre avril et juin 2016		226	
	Entre janvier et mars 2016		229	
	Avant décembre 2015		298	
	Sous-total		946	
	Total			1 910
	Témoins	Masculin	Après juin 2016	288
Entre avril et juin 2016			267	
Entre janvier et mars 2016			268	
Avant décembre 2015			218	
Sous-total			1 041	
Féminin			Après juin 2016	268
		Entre avril et juin 2016	246	
		Entre janvier et mars 2016	236	
		Avant décembre 2015	201	
		Sous-total	951	
		Total		

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

2.2 Pondération de comparaison

Comme mentionné au début de la présente partie, la sous-population des enfants-maternelle 4 ans diffère sensiblement de celle des enfants-témoins en ce qui a trait à leur répartition dans les strates principales. Afin d'éliminer l'effet de ces différences, on a défini une population de référence qui sert de base commune d'inférence pour la comparaison. En vue de tenir compte de cette standardisation à la population de référence dans le traitement de l'échantillon, une pondération différente a été effectuée. Celle-ci permet de produire des estimations optimales pour la comparaison des sous-populations.

Remarque sur la pondération de comparaison

Il est important de mentionner que la pondération de comparaison ne permet pas d'estimer des caractéristiques de la population des enfants visée par l'EEM4. Elle est construite précisément aux fins des analyses de comparaison présentées dans la troisième partie. Autrement dit, les estimations de proportions produites avec les poids de comparaison pourraient être différentes de celles obtenues avec les poids de description utilisés dans le recueil statistique.

2.2.1 Stratégie de pondération pour la comparaison

Les poids des enfants de l'échantillon de l'EEM4 admissibles à la comparaison sont calculés par rapport à la sous-population de référence pour la comparaison. À l'instar de la pondération pour la description de l'EEM4, la pondération et la production de poids d'autoamorçage (*bootstrap*) pour les analyses comparatives sont effectuées en suivant les quatre étapes de la pondération de

description. L'estimation de la variance et l'élaboration des tests de comparaison sont faites à partir des poids d'autoamorçage.

2.2.2 Particularités de la pondération de comparaison

Les étapes de pondération pour la comparaison sont les mêmes que celles de la pondération pour l'analyse descriptive de la maternelle 4 ans (section 1.3.3). La différence entre les deux pondérations repose essentiellement sur la standardisation des distributions des deux sous-populations à comparer. Elle se traduit par un calcul des poids de comparaison suivant les quatre mêmes étapes de la pondération de description tout en se basant sur les caractéristiques de la population de référence. Les particularités de la stratégie de pondération élaborée pour la comparaison sont présentées pour chaque étape.

A – Poids de comparaison de base

Avec la définition retenue de la population visée pour la comparaison, on suppose, autant pour les enfants-maternelle 4 ans que pour les enfants-témoins, un « pseudo-tirage » d'enfants à l'intérieur de chaque strate secondaire dont la taille correspond au nombre d'enfants-maternelle 4 ans dans la strate principale³⁶.

Pour chacun des groupes d'enfants (maternelle 4 ans et témoins), les poids de base des enfants de chaque strate principale sont établis en vertu de leur pseudo-probabilité de sélection. Ils correspondent à l'inverse de la « probabilité de sélection » dans la population de référence. Évidemment, il ne s'agit pas d'une réelle probabilité de sélection, mais plutôt d'une probabilité correspondant au fait d'être dans l'échantillon admis pour la comparaison. On obtient ainsi un poids de base pour chacun des enfants de l'échantillon selon la strate à laquelle il appartient.

Poids de base des enfants-maternelle 4 ans :

$$W_{\text{Stratecomp.}} = \frac{\text{Nombre d'enfants-maternelle 4 ans dans la strate principale de la population}}{\text{Nombre d'enfants-maternelle 4 ans dans la strate principale de l'échantillon}}$$

36. Rappelons que la strate principale correspond aux écoles ou regroupements d'écoles, alors que la strate secondaire correspond aux écoles ou regroupements d'écoles par sous-population (enfants-maternelle 4 ans et enfants-témoins).

Poids de base des enfants-témoins :

$$W_{Stratecomp.} = \frac{\text{Nombre d'enfants – maternelle 4 ans dans la strate principale de la population}}{\text{Nombre d'enfants – témoins dans la strate principale de l'échantillon}}$$

B – Ajustement des poids de comparaison pour la non-réponse à l'EQDEM

Les poids de comparaison de base sont ajustés pour que les poids des enfants de l'échantillon pour qui un questionnaire à l'EQDEM a été rempli correspondent à la somme des poids de base des enfants admissibles à la comparaison.

Pour chaque groupe, le poids de comparaison de base des enfants admissibles à la comparaison et pour qui un questionnaire à l'EQDEM est disponible est multiplié par l'inverse du taux de réponse (pondéré) dans le groupe homogène de pondération. Le premier ajustement pour la non-réponse est ainsi terminé. Les groupes de pondération pour l'ajustement à la non-réponse EQDEM sont les mêmes que ceux de la pondération de description et proviennent de la pondération de l'EQDEM 2022.

C – Ajustement des poids de comparaison pour la non-réponse à l'EEM4

Les groupes homogènes de pondération pour l'ajustement à la non-réponse de l'EEM4 sont les mêmes que ceux construits pour la pondération de description de la maternelle 4 ans.

Pour les enfants admissibles à la comparaison et qui sont répondants à l'EEM4, les poids de comparaison de l'étape précédente sont multipliés par l'inverse du taux de réponse (pondéré) dans chaque groupe de pondération.

D – Poststratification de comparaison

La dernière étape de la pondération de comparaison est la poststratification. Cet ajustement permet de s'assurer que la somme des poids de comparaison de tous les enfants répondants de l'échantillon des enfants admis à la comparaison correspond bien à la population de comparaison, notamment pour certaines caractéristiques d'intérêt pour la comparaison.

Rappelons que l'analyse comparative a entre autres pour but de mettre en lien le niveau de développement des enfants inscrits à la maternelle 5 ans en 2021-2022 avec la fréquentation ou non de la maternelle 4 ans à temps plein l'année précédente. Dans ce sens, la caractéristique analysée ne peut être utilisée comme facteur de contrôle dans la standardisation des sous-populations à comparer.

Pour la comparaison, aucun indicateur de vulnérabilité n'est donc utilisé dans la poststratification. Seuls l'âge et le sexe des enfants ont servi à la poststratification des répondants de l'EEM4 admis à la comparaison.

Pour chaque groupe d'enfants, la somme des poids des répondants de l'EEM4 admis à la comparaison correspond aux totaux de la population de référence selon les caractéristiques d'âge et de sexe des enfants. Cette étape d'ajustement des poids de chaque groupe d'enfants au nombre d'enfants de la population de référence pour la comparaison conclut la pondération de comparaison.

2.3 Mesures du développement des enfants et de la transition à la maternelle 5 ans

2.3.1 Vulnérabilité par domaine de développement et vulnérabilité globale

Dans le cadre de l'EQDEM, le développement des enfants à la maternelle 5 ans est mesuré à l'aide de l'IMDPE³⁷. Cet instrument, créé en 1999 par Dan R. Offord et Magdalena Janus du Offord Center for Child Studies (OCCS) à l'Université McMaster, a été élaboré en collaboration avec des spécialistes du développement de l'enfant, du personnel enseignant et du personnel éducatif en services de garde.

L'IMDPE est conçu pour mesurer les aptitudes des enfants à la maternelle 5 ans sur la base des observations du personnel enseignant. Il s'appuie sur les habiletés et comportements qui sont pour la plupart attendus chez un enfant en âge de fréquenter la maternelle 5 ans et non sur des points de référence inspirés d'un programme scolaire. Il ne s'agit donc pas d'une mesure basée sur les exigences du Programme-cycle de l'éducation préscolaire québécois. Les auteurs de l'IMDPE préconisent plutôt un critère empirique en considérant les enfants vulnérables

comme ceux dont le score équivaut au seuil correspondant au 10^e centile d'une population de référence (provinciale ou nationale) ou encore qui se trouve sous ce seuil. La réalisation de l'EQDEM en 2012, qui était la première édition québécoise de cette enquête, a permis de définir la population de référence et d'établir le seuil québécois de la vulnérabilité pour chaque domaine de développement.

Depuis, ces mesures repères ont été utilisées dans les éditions subséquentes de l'EQDEM (2017 et 2022) en raison des objectifs de comparaison dans le temps. Pour simplifier le traitement des données, cette option est également retenue pour les analyses comparatives de l'EEM4. En d'autres termes, ce sont les seuils établis à partir de la distribution des scores de l'EQDEM 2012 qui servent de point de référence pour établir la proportion d'enfants considérés comme vulnérables parmi les enfants-maternelle 4 ans ou parmi les enfants-témoins. Un enfant est considéré comme vulnérable lorsque son score dans un domaine de développement est égal ou inférieur au seuil empirique établi pour le domaine à partir de la population de référence. Les mesures du développement des enfants sont donc relatives.

Le tableau 2.4 décrit les cinq domaines de l'IMDPE ainsi que les seuils de 2012 (seuils de référence) ayant servi à effectuer les mesures de la vulnérabilité des enfants des sous-populations à comparer.

Tableau 2.4

Domaines de l'IMDPE¹, sujets abordés et seuil de référence

Domaine de développement	Sujets abordés	Seuil de référence
Santé physique et bien-être	Développement physique général, motricité, alimentation et habillement, propreté, ponctualité, état d'éveil	7,6923
Compétences sociales	Habiletés sociales, confiance en soi, sens des responsabilités, respect des pairs, des adultes, des règles et des routines, habitudes de travail et autonomie, curiosité	5,9615
Maturité affective	Comportement prosocial et entraide, crainte et anxiété, comportement agressif, hyperactivité et inattention, expression des émotions	5,8333
Développement cognitif et langagier	Intérêt et habiletés en lecture, en écriture et en mathématiques, utilisation adéquate du langage	5,8333
Habiletés de communication et connaissances générales	Capacité à communiquer de façon à être compris, capacité à comprendre les autres, articulation claire, connaissances générales	5,0000

1. L'IMDPE a été élaboré par Magdalena Janus et Dan R. Offord du OCCS (© McMaster University, Ontario).

Source : Institut de la statistique du Québec, *Enquête québécoise sur le développement des enfants à la maternelle 2012*.

37. L'IMDPE est largement utilisé dans les autres provinces canadiennes et à l'étranger sous l'appellation *Early Development Instrument* (EDI). Pour plus d'informations à propos des expériences canadiennes, se référer au site Web de l'OCCS : www.offordcentre.com.

Dans le questionnaire de l'EQDEM, les cinq domaines de développement sont couverts par un peu plus d'une centaine de questions factuelles³⁸. Celles-ci portent sur des comportements simples et observables par le personnel enseignant. Les enseignantes et enseignants devaient remplir un questionnaire pour chacun de leurs élèves inscrits à la maternelle 5 ans. Les réponses fournies par les membres du personnel enseignant à chaque question de l'EQDEM 2022 sont ramenées sur une échelle de 0 à 10. La moyenne des valeurs associées aux questions liées à un domaine constitue le score de l'enfant dans le domaine de développement en question. Chaque enfant obtient ainsi cinq scores moyens variant de 0 à 10, soit un pour chaque domaine de développement mesuré par l'IMDPE. Plus le score est faible, plus l'enfant serait susceptible de présenter des difficultés dans le domaine.

2.3.2 Transition à la maternelle 5 ans

L'EEM4 s'intéresse aussi à la transition à la maternelle 5 ans comme vécue par l'enfant, du point de vue des parents. Ces derniers ont répondu à cinq questions relatives au passage de leur enfant à la maternelle 5 ans à l'aide d'une échelle de réponse de Likert à quatre points (encadré 2.1). Les réponses obtenues à ces questions sont additionnées. Leur somme est ramenée sur une échelle de 0 à 10. La valeur obtenue constitue le score de l'enfant pour la transition à la maternelle 5 ans. Une mesure relative est ensuite obtenue au moyen d'un indicateur dichotomique de la transition à la maternelle 5 ans basé sur un seuil déterminé selon la distribution des scores des enfants. Ce seuil équivaut au score correspondant au 10^e centile de la distribution des scores. Il est établi à six pour cette édition. Les enfants dont le score de transition est inférieur à six appartiennent donc à la catégorie des enfants qui ont eu un passage à la maternelle plutôt difficile. À l'inverse, les enfants dont le score est d'au moins six sont classés dans la catégorie des enfants qui ont eu un passage à la maternelle moins difficile.

Encadré 2.1

Transition à la maternelle 5 ans

L'indicateur de la transition à la maternelle 5 ans est construit à partir de cinq énoncés pour lesquels le parent indiquait son degré d'accord.

Les cinq énoncés sont les suivants :

- Avant le début de la maternelle 5 ans, l'enfant était content à l'idée de commencer l'école.
- Les premières semaines d'école ont été une période stressante pour l'enfant. (item inversé)
- Pendant les premières semaines d'école, l'enfant a eu de la facilité à se faire des amis.
- Pendant les premières semaines d'école, l'enfant aimait son enseignant ou son enseignante.
- Pendant les premières semaines d'école, l'enfant était content d'aller à l'école.

Les quatre points de l'échelle de Likert indiquant les niveaux du degré d'accord sont les suivants :

- 1 = Tout à fait en désaccord
- 2 = Plutôt en désaccord
- 3 = Plutôt d'accord
- 4 = Tout à fait d'accord

38. Pour plus de détails, consulter le [Questionnaire-imdpe-2022.pdf](#).

2.4 Démarche de modélisation et présentation des résultats

Bien que certains facteurs aient été contrôlés dans le plan d'échantillonnage ou par la pondération, d'autres caractéristiques possiblement associées au développement des enfants ou à la transition à la maternelle 5 ans pourraient différencier les enfants-maternelle 4 ans des enfants-témoins. Ces facteurs ont alors un potentiel de confusion par rapport à l'association étudiée. Pour contrôler les effets de confusion potentiels sur la relation entre le développement de l'enfant (ou la transition à la maternelle 5 ans) et la fréquentation ou non de la maternelle 4 ans à temps plein, on a effectué des analyses multidimensionnelles (aussi appelées « multivariées ») à l'aide de régressions logistiques. La présente section décrit la méthodologie utilisée dans ces analyses de régression, notamment les hypothèses, les variables et les étapes de modélisation. Les résultats de ces analyses sont présentés dans la troisième partie du rapport.

2.4.1 Hypothèse de l'analyse comparative

Les tests statistiques présentés dans cette section visent à voir si l'on peut rejeter l'hypothèse qu'il n'y a pas de différence entre les deux groupes d'enfants, lorsque l'on tient compte des caractéristiques socio-économiques des familles ou du parcours préscolaire des enfants. En ce sens, pour des enfants ayant des caractéristiques semblables, les tests de comparaison effectués postulent un lien d'indépendance entre la fréquentation de la maternelle 4 ans à temps plein et le niveau de développement relatif des enfants à leur entrée à la maternelle ou encore la transition vécue à ce moment.

2.4.2 Variables de modélisation

Variables dépendantes

Afin de modéliser le lien entre le développement relatif des enfants comme perçu par l'enseignante ou l'enseignant, ou la transition à la maternelle 5 ans comme perçue par les parents, et la fréquentation de la maternelle 4 ans à temps plein, on a construit des indicateurs dichotomiques à partir des scores obtenus pour chacun des phénomènes. Ces variables binaires indiquent si l'enfant est vulnérable ou non dans le domaine de

développement considéré (vulnérabilité relative dans chacun des cinq domaines de développement mesurés par l'IMDPE). Deux indicateurs composites de vulnérabilité, dans au moins un domaine et dans au moins deux domaines de développement, ont aussi été considérés. Enfin, une autre variable binaire indique si l'enfant a eu un passage à la maternelle plutôt difficile selon ses parents. Chacun de ces huit indicateurs dichotomiques représente une variable dépendante analysée pour lequel un modèle spécifique est construit par la suite.

Variable indépendante d'intérêt

Puisque l'étude consiste à comparer la sous-population des enfants ayant fréquenté la maternelle 4 ans à temps plein à celle des enfants ne l'ayant pas fréquentée, l'indicateur de fréquentation ou non de la maternelle 4 ans à temps plein est la seule variable indépendante d'intérêt pour tous les modèles. Cette variable est donc fondamentale et doit être incluse dans tous les modèles multivariés.

L'information ayant permis de définir cette variable indépendante provient des fichiers du MEQ.

Variables indépendantes caractéristiques

On inclut aussi dans la modélisation toutes les variables qui peuvent modifier le lien entre les variables dépendantes et la fréquentation de la maternelle 4 ans à temps plein ou non. Ces variables qui peuvent servir de contrôle entre les enfants des deux groupes sont des caractéristiques potentiellement confondantes ou susceptibles d'être associées au développement ou à la transition à la maternelle 5 ans.

Étant donné l'objectif poursuivi par l'analyse multivariée, soit le contrôle des effets potentiellement confondants, seules les caractéristiques des enfants antérieures à l'entrée des enfants à la maternelle 5 ans sont prises en considération.

Les variables caractéristiques peuvent provenir de différentes sources. Certaines variables liées aux caractéristiques de l'enfant et de son parcours préscolaire telles que l'IMSE de l'élève, la fréquentation de la maternelle 4 ans à mi-temps ou la participation au programme Passe-Partout l'année précédente, le nombre de parents inscrits dans le dossier de l'élève et la langue d'usage à la maison ont été construites à partir des informations de

la base de sondage. Certaines des caractéristiques des parents et de la famille ainsi que des pratiques parentales en matière de littératie ont été construites à partir des données recueillies auprès des parents. La présence d'un plan d'intervention en maternelle 5 ans n'est pas utilisée dans la modélisation puisqu'elle est vue comme très proche des variables dépendantes (vulnérabilité).

Le tableau 2.5 présente les différentes caractéristiques prises en compte dans les analyses comparatives de l'EEM4.

2.4.3 Étapes de modélisation

Le plan d'échantillonnage et la pondération ont permis de réduire les différences selon certaines caractéristiques connues entre les deux sous-populations à comparer. Par exemple, en construisant les strates principales à partir de l'école ou du regroupement d'écoles ayant un niveau de défavorisation (mesuré par l'IMSE) semblable, on tente d'éliminer certains écarts entre les enfants-témoins et les enfants-maternelle 4 ans quant à la défavorisation. De même, par la pondération basée sur la population de référence pour la comparaison décrite dans la section 2.2, l'écart entre les distributions des enfants des deux sous-populations selon l'âge et le sexe est contrôlé.

À l'étape d'analyse des liens avec la fréquentation de la maternelle 4 ans à temps plein, des différences demeurent entre les enfants-maternelle 4 ans et les enfants-témoins en ce qui a trait à d'autres caractéristiques possiblement associées au développement des enfants ou à la transition à la maternelle 5 ans. Ces caractéristiques concernent par exemple l'environnement familial des enfants qui n'était pas connu avant la collecte. Ces différences devraient être prises en compte dans l'analyse des liens. Des modèles multivariés sont donc construits en deux étapes à l'aide de régressions logistiques. La première étape consiste à analyser individuellement l'association entre chaque variable dépendante et la variable indépendante d'intérêt (fréquentation de la maternelle 4 ans à temps plein), l'association entre chaque variable dépendante et chaque caractéristique ainsi que l'association entre la variable indépendante d'intérêt et chaque caractéristique. La deuxième étape examine l'association entre chaque variable dépendante et la variable indépendante d'intérêt en tenant compte simultanément de plusieurs variables caractéristiques.

Tableau 2.5

Variables caractéristiques potentiellement confondantes prises en compte dans les analyses multivariées

Caractéristiques individuelles de l'enfant	Caractéristiques des parents et de la famille	Caractéristiques du parcours préscolaire de l'enfant
• Sexe¹ • Âge¹ • Lieu de naissance • Langue parlée le plus souvent à la maison • Indicateurs liés à la santé	• Âge de la mère et du père • Zone de résidence • Indicateurs liés au type de famille • Indicateurs liés au lieu de naissance des parents ou à leur immigration • Indicateurs liés à la scolarité des parents • Indicateurs liés au revenu du ménage • Indicateurs liés à la santé des parents • Indicateurs liés à l'occupation des parents • Indicateurs liés au soutien disponible à la famille • Indicateurs liés aux pratiques parentales en matière de littératie	• Indicateurs liés à fréquentation d'un service de garde • Indicateurs liés au type de service de garde fréquenté • Indicateurs liés au profil de garde de l'enfant

1. Avec la pondération à la population de comparaison, le sexe et l'âge de l'enfant ne sont pas considérés comme des variables confondantes. Cependant, ces données sont conservées pour assurer qu'elles n'interagissent pas avec la fréquentation de la maternelle 4 ans à temps plein dans la relation avec le développement ou la transition à la maternelle 5 ans.

Analyses bivariées

L'analyse bivariée permet la comparaison du développement global ou de la transition à la maternelle 5 ans entre la sous-population des enfants ayant fréquenté la maternelle 4 ans à temps plein et celle des enfants-témoins. Cette analyse est menée à l'aide d'un modèle de régression logistique pour chaque domaine de développement ou pour la transition à la maternelle 5 ans avec comme variable indépendante la fréquentation ou non de la maternelle 4 ans à temps plein. On peut alors établir l'existence ou non d'une association entre chaque variable dépendante et la fréquentation de la maternelle 4 ans à temps plein.

Toutefois, il est possible que certaines caractéristiques des deux sous-populations comparées présentent des différences qui expliquent ou modulent chacune des relations. Ces caractéristiques sont d'abord suggérées par la littérature, puis vérifiées à l'aide de modèles de régression logistique simple ayant comme variable dépendante la fréquentation de la maternelle 4 ans à temps plein et comme variable indépendante, chacune des variables indépendantes caractéristiques.

Plus de 170 modèles de régression logistique simple³⁹ ont été construits. À la fin de cette étape, les caractéristiques pouvant avoir un effet confondant ou pouvant moduler la relation entre les variables dépendantes et la fréquentation de la maternelle 4 ans à temps plein sont retenues : ce sont les caractéristiques significativement associées à la fois à la fréquentation de la maternelle 4 ans à temps plein et au domaine de développement considéré ou à la transition à la maternelle 5 ans, au seuil de 5 %.

En plus de l'indicateur de fréquentation de la maternelle 4 ans à temps plein, qui est la variable indépendante d'intérêt, les caractéristiques retenues sont considérées dans les modèles des analyses multivariées de l'étape suivante.

Analyses multivariées

L'analyse multivariée a pour objectif de contrôler les effets potentiellement confondants sur la relation entre la fréquentation de la maternelle 4 ans à temps plein et le développement de l'enfant ou la transition à la maternelle 5 ans. Pour ce faire, des modèles de régression logistique multiple sont construits pour chaque variable dépendante. D'abord, toutes les caractéristiques significatives identifiées par les analyses bivariées sont introduites dans le modèle comme facteurs, en plus de la variable indépendante d'intérêt. Ensuite, les facteurs qui deviennent non significatifs en présence des autres sont progressivement retirés, un à la fois : le facteur avec la valeur *p* supérieure à 5 % la plus élevée est chaque fois retiré. L'opération est répétée jusqu'à ce que tous les facteurs qui restent dans le modèle soient significatifs au seuil de 5 %. Il s'agit de la procédure pas-à-pas descendant (*backward elimination*), basée sur la statistique *F* (Shah 1976) correspondant à un ajustement de Satterthwaite du test du khi-deux. Un modèle distinct est produit pour chacun des sept indicateurs de vulnérabilité et pour la transition à la maternelle 5 ans.

Certaines caractéristiques introduites dans les modèles peuvent être fortement corrélées. Par exemple, l'indicateur « Plus haut diplôme obtenu par l'un ou l'autre des parents » est corrélé à l'indicateur « Diplôme de la mère ». Un même indicateur peut aussi être décliné de plusieurs façons : en deux catégories (Diplôme de niveau secondaire ou moins et Diplôme de niveau collégial ou plus) ou en quatre catégories (Aucun diplôme, Diplôme de niveau secondaire, Diplôme de niveau collégial et Diplôme universitaire). Maintes tentatives successives de modélisations ont alors été réalisées par l'introduction, tour à tour, des différentes variables mesurant un même phénomène. Chaque fois qu'un choix s'impose, la variable indépendante qui entraîne le plus grand changement est conservée⁴⁰.

39. Les 170 modèles de régression logistique simple proviennent de l'équation suivante : (8 variables dépendantes x 19 variables indépendantes) + (1 variable indépendante d'intérêt x 18 variables indépendantes caractéristiques).

40. Des modèles plus simples, ne contenant que la variable spécifiant si l'enfant avait fréquenté la maternelle 4 ans à temps plein ou non et une autre variable d'intérêt, ont été faits pour voir quelle variable, parmi celles qui étaient fortement corrélées, avait le plus d'effet confondant sur la relation de la variable dépendante avec la fréquentation de la maternelle 4 ans à temps plein, c'est-à-dire laquelle entraînait le plus grand changement dans l'estimation du paramètre de la variable de fréquentation de la maternelle 4 ans à temps plein lorsqu'elle était introduite dans le modèle.

Une fois les caractéristiques de contrôle retenues, les interactions entre les facteurs significatifs et la fréquentation ou non de la maternelle 4 ans à temps plein ont été testées pour vérifier si ce programme préscolaire a le même effet sur le domaine de développement ou la transition à la maternelle 5 ans selon les modalités des variables de contrôle.

Mentionnons enfin que les variables à conserver pour chaque modèle sont d'abord sélectionnées sans que les enfants pour lesquels elles présentent des valeurs manquantes soient pris en compte. En revanche, une fois le modèle final choisi, ces enfants sont ajoutés par la création de modalités correspondant aux valeurs inconnues. Comme la non-réponse partielle est assez faible, cette approche ne cause pas de problèmes particuliers.

Présentation des résultats des analyses multivariées

Les rapports de cotes sont présentés dans ce rapport comme résultats des régressions logistiques ajustées. Pour un modèle donné, en présence d'un lien significatif entre la variable dépendante et une variable indépendante, les différences entre deux catégories peuvent être interprétées à l'aide du rapport de cotes, une mesure relative de la force de la relation observée. Une cote se définit par $p/(1-p)$. Dans le cas présent, il s'agit du rapport de la probabilité d'être considéré comme vulnérable (ou d'avoir eu un passage à la maternelle plutôt difficile) à la probabilité de ne pas l'être (ou d'avoir eu un passage à la maternelle moins difficile). Le rapport de cotes comparant les catégories 1 (enfants-maternelle 4 ans) et 2 (enfants-témoins) d'une variable indépendante se définit comme suit :

$$RC = \frac{p_1/(1 - p_1)}{p_2 / (1 - p_2)}$$

Dans le cas de la modélisation d'un indicateur de développement, un rapport de cotes de 1 signifie qu'il n'y a pas de différence entre les catégories comparées. Un rapport de cotes significativement supérieur à 1 indique que la probabilité d'être considéré comme vulnérable est plus élevée chez les enfants de la catégorie 1 que chez ceux de la catégorie 2. Enfin, un rapport de cotes significativement inférieur à 1 indique que les enfants de la catégorie 1 sont moins susceptibles d'être considérés comme vulnérables que ceux de la catégorie 2.

Pour la transition à la maternelle 5 ans, un rapport de cotes significativement supérieur à 1 indique que la probabilité d'avoir eu un passage à la maternelle 5 ans plutôt difficile est plus élevée chez les enfants de la catégorie 1 que chez ceux de la catégorie 2. Inversement, un rapport de cotes significativement inférieur à 1 indique que les enfants de la catégorie 1 sont moins susceptibles d'avoir eu un passage à la maternelle 5 ans plutôt difficile que ceux de la catégorie 2.

Dans cette étude cas-témoins, les rapports de cotes d'intérêt sont ceux qui mettent en parallèle la cote des enfants-maternelle 4 ans et celle des enfants-témoins. Pour le calcul du rapport de cotes non ajusté associé à la comparaison des enfants-maternelle 4 ans avec les enfants-témoins, la probabilité d'être considéré comme vulnérable ou d'avoir eu un passage à la maternelle 5 ans plutôt difficile est estimée directement avec la proportion d'enfants considérés comme vulnérables ou qui ont eu un passage plutôt difficile parmi les enfants-maternelle 4 ans et de la même probabilité parmi les enfants-témoins.

En revanche, un rapport de cotes ajusté tient compte des effets de confusion contrôlés dans le modèle. Il nous informe donc sur l'association entre la fréquentation de la maternelle 4 ans à temps plein et le développement de l'enfant ou la transition à la maternelle 5 ans.

Mentionnons qu'en présence d'une interaction entre la fréquentation ou non de la maternelle 4 ans à temps plein et une variable indépendante caractéristique, l'analyse reposant sur le rapport de cotes marginal comparant les enfants maternelle 4 ans aux enfants-témoins n'a pas de sens. La comparaison de ces deux sous-populations doit alors être faite à l'intérieur de chacun des niveaux de la variable indépendante avec laquelle il y a interaction. Par exemple, en présence d'une interaction entre l'indicateur de fréquentation et le sexe de l'enfant, les rapports de cotes sont calculés séparément pour les garçons et les filles. Les rapports de cotes dits non ajustés seraient alors calculés à partir du modèle de régression logistique contenant uniquement l'indicateur de fréquentation de la maternelle 4 ans à temps plein, le sexe et le terme d'interaction entre ces deux variables, tandis que les rapports de cotes dits ajustés seraient obtenus à partir du modèle complet contenant aussi les autres variables indépendantes prises en considération.

Il est important de noter qu'un rapport de cotes ne peut être interprété quantitativement comme un risque relatif que lorsque la prévalence du phénomène étudié (variable dépendante) par un modèle de régression logistique est faible pour toutes les combinaisons des variables facteurs. Ce n'est en général pas le cas pour les mesures du développement de l'enfant ou de la transition à la maternelle 5 ans étudiées dans ce rapport. Par conséquent, on s'en tiendra à la seule interprétation du sens des relations observées.

Pour chaque rapport de cotes estimé, un intervalle de confiance à 95 % est présenté. Lorsqu'un intervalle de confiance pour le rapport de cotes ne contient pas la valeur 1, on conclut qu'il y a une différence significative entre cette modalité et la modalité de référence à laquelle elle est comparée au seuil de 5 %.

3 Résultats des analyses comparatives de la fréquentation de la maternelle 4 ans

La présente partie du rapport porte essentiellement sur les résultats de comparaison entre la sous-population des enfants-maternelle 4 ans et la sous-population des enfants-témoins. Ces résultats permettent d'établir si des liens ont été décelés entre les variables dépendantes et la variable indépendante d'intérêt (section 2.4.2) lorsque les caractéristiques significatives sont contrôlées.

En vue d'une interprétation et d'une utilisation adéquates de l'ensemble des résultats de l'EEM4, on discute d'abord des résultats présentés dans le recueil statistique et dans le présent rapport, avant de s'attarder aux résultats de la comparaison du niveau de développement des enfants et de la transition à la maternelle 5 ans.

3.1 Analyses descriptives (recueil statistique) et analyses comparatives (section 3.2)

Le recueil statistique de l'EEM4 offre des analyses descriptives, alors que, dans ce rapport, à la section 3.2, on présente des résultats d'analyses comparatives. Ces analyses ont des objectifs différents.

Les tableaux 3.1 et 3.2 présentent, pour les indicateurs à l'étude, les résultats de l'analyse descriptive et de l'analyse de comparaison globale⁴¹.

Pour obtenir des estimations de proportions d'enfants relativement vulnérables ou d'enfants ayant eu un passage à la maternelle 5 ans plutôt difficile chez les enfants ayant fréquenté la maternelle 4 ans, on recommande l'utilisation des renseignements du recueil statistique. On y trouve aussi l'information concernant la sous-population des enfants n'ayant pas fréquenté la maternelle 4 ans.

Pour ce qui est de la comparaison des deux sous-populations, on recommande d'utiliser les résultats des sections 3.2 et suivantes du présent rapport. La comparaison y est faite de manière optimale pour « isoler » le lien entre la fréquentation de la maternelle 4 ans et le développement des enfants ou la transition à la maternelle 5 ans. Rappelons aussi que l'étude permet de vérifier s'il y a ou non une association entre la maternelle 4 ans et le développement (ou la transition à la maternelle); elle ne permet toutefois pas d'établir s'il s'agit ou non d'une relation de cause à effet.

41. Les analyses de comparaison globale correspondent aux analyses bivariées effectuées comme première étape des analyses multivariées (section 2.4.3).

Tableau 3.1

Résultats des analyses bivariées des différents domaines de développement selon la population et les poids de description et de comparaison, enfants visés par l'EEM4, Québec, 2022

Domaine de développement	Sous-population	Avec population et poids de description (recueil statistique)		Avec population et poids de comparaison (section 3.2.1)	
		Proportions d'enfants vulnérables (%)	Intervalle de confiance (IC à 95 %)	Rapport de cotes ¹ (RC)	Intervalle de confiance (IC à 95 %)
Au moins un domaine	Maternelle 4 ans	30,8 ^a	30,1 - 31,4	1,09	0,95 - 1,26
	Témoins	29,2 ^a	28,3 - 30,0	1,00	...
Au moins deux domaines	Maternelle 4 ans	13,8	12,5 - 15,3	1,02	0,83 - 1,24
	Témoins	14,2	12,6 - 15,9	1,00	...
Santé physique et bien-être	Maternelle 4 ans	10,7	9,5 - 12,1	1,04	0,83 - 1,31
	Témoins	9,8	8,4 - 11,4	1,00	...
Compétences sociales	Maternelle 4 ans	11,2	9,9 - 12,6	1,12	0,89 - 1,41
	Témoins	10,5	9,1 - 12,0	1,00	...
Développement cognitif et langagier	Maternelle 4 ans	10,6	9,4 - 12,0	0,86	0,69 - 1,07
	Témoins	12,4	11,0 - 14,0	1,00	...
Habilités de communication et connaissances générales	Maternelle 4 ans	9,8	8,6 - 11,2	0,84	0,66 - 1,06
	Témoins	11,1	9,6 - 12,7	1,00	...
Maturité affective	Maternelle 4 ans	12,5	11,1 - 14,0	1,13	0,92 - 1,39
	Témoins	11,9	10,4 - 13,5	1,00	...

a : Pour un domaine de développement donné, exprime une différence significative entre les proportions d'enfants considérés comme vulnérables parmi les enfants-maternelle 4 ans et parmi les enfants-témoins au seuil de 0,05.

1. Un rapport de cotes significativement supérieur à 1, c'est-à-dire dont la borne inférieure de l'intervalle de confiance est supérieure à 1, indique que les enfants-maternelle 4 ans sont plus susceptibles que les enfants-témoins d'être vulnérables dans le domaine considéré, alors qu'un rapport de cotes significativement inférieur à 1, c'est-à-dire dont la borne supérieure de l'intervalle de confiance est inférieure à 1, indique qu'ils le sont moins. Un rapport de cotes dont l'intervalle de confiance comprend la valeur 1 indique une différence non significative entre les enfants-maternelle 4 ans et les enfants-témoins pour ce qui est de la probabilité d'être vulnérables. Les enfants-témoins représentent la catégorie de référence. Le rapport de cotes vaut 1 pour la catégorie de référence.

... Indique qu'un intervalle de confiance n'est pas construit pour le rapport de cotes de la catégorie de référence.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

Comme le montre ce tableau, environ 31% des enfants qui fréquentaient la maternelle 5 ans en 2021-2022 dans une école ayant au moins trois enfants qui ont fréquenté la maternelle 4 ans à temps plein et qui ont fréquenté eux-mêmes la maternelle 4 ans (enfants-maternelle 4 ans) sont considérés comme vulnérables dans au moins un domaine de développement. Toujours selon les résultats de l'analyse de description, cette proportion serait statistiquement supérieure à celle estimée pour les enfants-témoins.

Cependant, lorsqu'on utilise la population et les poids de comparaison, c'est-à-dire lorsqu'on contrôle la taille et la distribution des sous-populations et que l'on ne tient plus compte de la vulnérabilité dans la pondération, on ne décèle plus de différence significative entre les deux sous-populations pour ce qui est de la vulnérabilité dans au moins un domaine de développement. En effet, l'intervalle de confiance du rapport de cotes de vulnérabilité entre les enfants-maternelle 4 ans et les enfants-témoins contient la valeur 1⁴².

42. Pour plus d'informations sur le rapport de cotes, consulter la sous-section [Présentation des résultats des analyses multivariées](#) de la section 2.4.3.

Autrement dit, lorsque les sous-populations sont rendues comparables de façon optimale par la pondération de comparaison, c'est-à-dire lorsque la sous-population des enfants-témoins ressemble davantage à la sous-population des enfants-maternelle 4 ans, on ne détecte plus de différence significative entre les enfants-maternelle 4 ans et les enfants-témoins pour ce qui est de la vulnérabilité dans au moins un domaine de développement. Cela suggère que la différence trouvée par l'analyse descriptive s'explique par les caractéristiques des deux sous-populations et ne serait pas associée à la fréquentation de la maternelle 4 ans.

L'analyse descriptive montre aussi qu'environ 7 % des enfants qui fréquentaient la maternelle 5 ans en 2021-2022 dans une école ayant au moins trois enfants qui ont fréquenté la maternelle 4 ans à temps plein et qui ont eux-mêmes fréquenté la maternelle 4 ans (enfants-maternelle 4 ans) ont vécu un passage à la maternelle considérée comme plutôt difficile selon la perception des parents (tableau 3.2). Cette proportion serait statistiquement moins élevée que celle estimée pour les enfants-témoins. L'analyse comparative confirme cette différence qui ne serait pas seulement le fait des caractéristiques distinctes des enfants-maternelle 4 ans et des enfants-témoins.

3.2 Résultats des analyses comparatives

La présente section examine d'abord le lien entre la fréquentation de la maternelle 4 ans à temps plein et la vulnérabilité relative des enfants à la maternelle 5 ans dans différents domaines de développement, puis celui entre la fréquentation de la maternelle 4 ans à temps plein et la transition à la maternelle 5 ans, tout en tenant compte de facteurs ayant un potentiel de confusion. Le potentiel de confusion sur la relation observée provient des variables indépendantes associées à la fois au domaine de développement et à la fréquentation de la maternelle 4 ans à temps plein.

Des modèles de régression logistique multiple sont ajustés pour chacune des variables de vulnérabilité au moyen des échantillons et des poids de comparaison. Cette démarche est ensuite reprise pour analyser le lien entre la fréquentation de la maternelle 4 ans à temps plein et la transition à la maternelle 5 ans.

Rappelons que les résultats des analyses comparatives sont présentés exclusivement sous forme de rapports de cotes et que toutes les analyses de comparaison qui suivent ont été réalisées sur la population de comparaison et au moyen de la pondération de comparaison.

Tableau 3.2

Résultats des analyses bivariées de la transition à la maternelle 5 ans selon la population et les poids utilisés, enfants visés par l'EEM4, Québec, 2022

Sous-population	Avec population et poids de description (recueil statistique)		Avec population et poids de comparaison (section 3.2.2)	
	Proportions d'enfants dont le passage à la maternelle 5 ans est considéré comme plutôt difficile (%)	Intervalle de confiance (IC à 95 %)	Rapport de cotes	Intervalle de confiance (IC à 95 %)
Maternelle 4 ans	7,3 ^a	6,1 - 8,6	0,70†	0,56 - 0,88
Témoins	9,3 ^a	7,9 - 10,9	1,00	...

a : Exprime une différence significative entre les proportions d'enfants pour lesquels le passage à la maternelle 5 ans est considéré comme plutôt difficile parmi les enfants-maternelle 4 ans et parmi les enfants-témoins au seuil de 0,05.

† Rapport de cotes significativement différent de 1,00 ($p < 0,05$). Le rapport de cotes inférieur à 1 indique que les enfants-maternelle 4 ans sont significativement moins susceptibles d'avoir eu un passage à la maternelle plutôt difficile que les enfants-témoins. Les enfants-témoins représentent la catégorie de référence. Le rapport de cotes vaut 1 pour la catégorie de référence.

... Indique qu'un intervalle de confiance n'est pas construit pour le rapport de cotes de la catégorie de référence.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

3.2.1 Vulnérabilité et fréquentation de la maternelle 4 ans selon les analyses comparatives

Comme mentionné à la section 2.4.3, l'analyse du lien entre le développement des enfants et la fréquentation de la maternelle 4 ans à temps plein commence par une analyse bivariée, avant de considérer les facteurs potentiellement confondants (analyse multivariée). Le tableau 3.3 présente les rapports de cotes non ajustés découlant des analyses bivariées⁴³ et les rapports de cotes ajustés obtenus à partir des modèles multivariés.

Considérant la population de comparaison, pour chacune des mesures de développement de l'étude, l'analyse bivariée de la vulnérabilité relative en fonction de la fréquentation de la maternelle 4 ans à temps plein n'a pas permis de détecter une différence significative entre la sous-population des enfants-maternelle 4 ans et celle des enfants-témoins. Les rapports de cotes non ajustés ne sont pas significativement différents de 1, au seuil de 5 % (tableaux 3.1 et 3.3). Avec les échantillons de répondants disponibles pour la comparaison, l'étude devait permettre de détecter un écart de l'ordre de 2,8 % entre les proportions d'enfants vulnérables chez les enfants ayant fréquenté la maternelle 4 ans et chez les enfants-témoins avec une puissance de 80 %. Il y aurait donc tout lieu de croire que les écarts réels entre les enfants-témoins et les enfants-maternelle 4 ans en ce qui a trait aux proportions d'enfants relativement vulnérables seraient inférieurs à 2,8 %. D'ailleurs, le plus grand écart observé entre les proportions d'enfants vulnérables chez les enfants-maternelle 4 ans et chez les enfants-témoins lors de l'analyse de comparaison était de 1,8 %⁴⁴.

Malgré les résultats non significatifs des analyses bivariées, on a mené des analyses en contrôlant les facteurs ayant un potentiel de confusion sur le lien entre la fréquentation de la maternelle 4 ans à temps plein et la vulnérabilité dans les différents domaines de développement. Celles-ci permettent de s'assurer que des facteurs non contrôlés lors de l'échantillonnage ou de la pondération ne masquent pas les différences qui auraient pu être observées entre les deux sous-populations. Il s'agit de prendre en considération, dans les analyses, les facteurs qui correspondent aux caractéristiques associées à la fois à la vulnérabilité dans le domaine de développement analysé et à la fréquentation de la maternelle 4 ans à temps plein (analyses multivariées) et qui permettent d'obtenir des rapports de cotes ajustés. Pour chaque domaine de développement, les facteurs⁴⁵ considérés dans l'analyse multivariée sont indiqués en note au bas du tableau 3.3.

L'analyse multivariée montre que pour les domaines « Maturité affective », « Développement cognitif et langagier » et « Habiletés de communication et connaissances générales », la fréquentation de la maternelle 4 ans à temps plein est associée à la vulnérabilité lorsque certains facteurs sont contrôlés (tableau 3.3).

43. En présence d'une interaction entre une variable caractéristique et l'indicateur de fréquentation de la maternelle 4 ans à temps plein, les analyses bivariées sont menées à l'intérieur de chacun des niveaux de la variable caractéristique avec laquelle il y a interaction.

44. Les proportions estimées avec les poids de comparaison ne sont pas présentées dans ce rapport puisqu'elles n'ont aucun intérêt.

45. Les indicateurs des facteurs retenus dans les modèles des analyses multivariées ainsi que leurs niveaux sont présentés à l'annexe IV.

Tableau 3.3

Rapports de cotes liant la fréquentation de la maternelle 4 ans à la vulnérabilité relative, enfants de la population de comparaison, Québec, 2022

		Sans correction des effets de confusion ¹		Avec correction des effets de confusion	
		Rapport de cotes	Intervalle de confiance (IC à 95 %)	Rapport de cotes ajusté	Intervalle de confiance (IC à 95 %)
			de à		de à
Vulnérabilité dans au moins un domaine de développement²					
	Maternelle 4 ans	1,09	0,95 1,26	0,98	0,85 1,14
	Témoins	1,00		1,00	
Vulnérabilité dans au moins deux domaines de développement³					
	Maternelle 4 ans	1,02	0,83 1,24	0,88	0,72 1,08
	Témoins	1,00		1,00	
Vulnérabilité dans le domaine « Santé physique et bien-être »⁴					
	Maternelle 4 ans	1,04	0,83 1,31	0,91	0,72 1,15
	Témoins	1,00		1,00	
Vulnérabilité dans le domaine « Compétences sociales »⁵					
	Maternelle 4 ans	1,12	0,89 1,41	0,84	0,62 1,13
	Témoins	1,00		1,00	
Vulnérabilité dans le domaine « Développement cognitif et langagier »⁶					
	Maternelle 4 ans	0,86	0,69 1,07	0,73 [†]	0,57 0,92
	Témoins	1,00		1,00	
Vulnérabilité dans le domaine « Habiletés de communication et connaissances générales »⁷					
Faible revenu	Maternelle 4 ans	0,49 [†]	0,33 0,73	0,41 [†]	0,27 0,62
	Témoins	1,00		1,00	
Revenu élevé	Maternelle 4 ans	0,98	0,72 1,33	0,72	0,50 1,02
	Témoins	1,00		1,00	
Vulnérabilité dans le domaine « Maturité affective »⁸					
Diplôme d'études secondaires ou moins	Maternelle 4 ans	0,80	0,56 1,15	0,76	0,52 1,09
	Témoins	1,00		1,00	
Diplôme d'études collégiales	Maternelle 4 ans	0,82	0,53 1,29	0,82	0,52 1,30
	Témoins	1,00		1,00	
Diplôme universitaire	Maternelle 4 ans	1,73 [†]	1,20 2,49	1,75 [†]	1,20 2,54
	Témoins	1,00		1,00	

† Rapport de cotes significativement différent de 1,00 ($p < 0,05$). Un rapport de cotes inférieur à 1 indique que les enfants-maternelle 4 ans sont significativement moins susceptibles d'être considérés comme vulnérables dans un domaine de développement donné que les enfants-témoins. Un rapport de cotes supérieur à 1 indique que les enfants-maternelle 4 ans sont significativement plus susceptibles d'être considérés comme vulnérables dans un domaine de développement donné que les enfants-témoins.

... Indique qu'un intervalle de confiance n'est pas construit pour le rapport de cotes de la catégorie de référence.

1. En présence d'une interaction entre une variable caractéristique et l'indicateur de fréquentation de la maternelle 4 ans à temps plein, les analyses bivariées sont menées et présentées à l'intérieur de chacun des niveaux de la variable caractéristique avec laquelle il y a interaction.
2. L'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : plus haut diplôme obtenu par l'un ou l'autre des parents, niveau de revenu du ménage et niveau d'engagement relatif dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans.
3. L'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : plus haut diplôme obtenu par l'un ou l'autre des parents, niveau de revenu du ménage, nombre de déménagements au cours des cinq dernières années et niveau d'engagement relatif dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans.
4. L'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : plus haut diplôme obtenu par l'un ou l'autre des parents, niveau de revenu du ménage, nombre de déménagements au cours des cinq dernières années, niveau d'engagement relatif dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans et problème de santé des parents limitant les activités quotidiennes à la maison.
5. L'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : plus haut diplôme obtenu par l'un ou l'autre des parents, niveau de revenu du ménage, nombre de déménagements au cours des cinq dernières années, niveau d'engagement relatif dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans et durée totale de garde.
6. L'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : plus haut diplôme obtenu par l'un ou l'autre des parents et niveau de revenu du ménage.
7. En plus de l'interaction entre la fréquentation de la maternelle 4 ans et le niveau de revenu du ménage, l'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : durée totale de garde et langue parlée le plus souvent à la maison.
8. En plus de l'interaction entre la fréquentation de la maternelle 4 ans et le plus haut diplôme obtenu par l'un ou l'autre des parents, l'analyse multivariée permet la correction des effets de confusion des caractéristiques suivantes : nombre de déménagements au cours des cinq dernières années et niveau d'engagement relatif dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

Vulnérabilité dans le domaine « Développement cognitif et langagier »

Dans le domaine « Développement cognitif et langagier », la probabilité d'être considéré comme vulnérable est moins élevée parmi les enfants-maternelle 4 ans ($RC_{ajusté} = 0,73$) que parmi les enfants-témoins une fois la standardisation à la population de référence effectuée et lorsqu'on contrôle les effets du plus haut diplôme obtenu par l'un ou l'autre des parents et du niveau de revenu du ménage.

Vulnérabilité dans le domaine « Habiletés de communication et connaissances générales »

Dans le domaine « Habiletés de communication et connaissances générales », on observe que la probabilité d'être vulnérable est moins élevée chez les enfants-maternelle 4 ans que chez les enfants-témoins. De plus, dans ce domaine de développement, l'écart entre les enfants-maternelle 4 ans et les enfants-témoins est modulé par le niveau de revenu du ménage. Chez les enfants vivant dans un ménage à faible revenu, la probabilité d'être considéré comme vulnérable est significativement moins élevée parmi les enfants-maternelle 4 ans ($RC_{ajusté} = 0,41$) que parmi les enfants-témoins.

Vulnérabilité dans le domaine « Maturité affective »

L'analyse multivariée a permis de contrôler les effets de confusion dus à la durée cumulative en mois en services de garde de l'enfant, à la langue parlée le plus souvent à la maison et au revenu du ménage dans lequel l'enfant vit. La probabilité d'être considéré comme vulnérable est alors modulée selon la scolarité des parents.

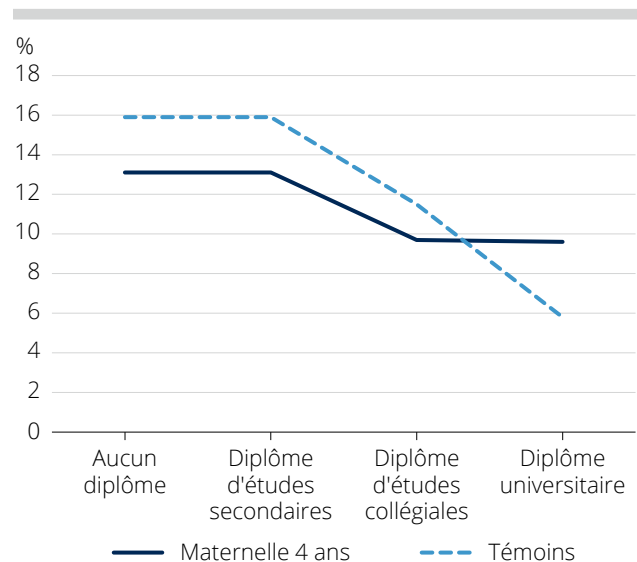
Chez les enfants dont le plus haut diplôme obtenu par l'un ou l'autre des parents est un diplôme universitaire, la probabilité d'être considéré comme vulnérable est significativement plus élevée parmi les enfants-maternelle 4 ans ($RC_{ajusté} = 1,75$) que parmi les enfants-témoins.

Chez les enfants dont les parents n'ont aucun diplôme ou seulement un diplôme d'études secondaires, bien que l'écart ne soit pas statistiquement significatif en raison d'un plus faible effectif, il semble y avoir une tendance selon laquelle les enfants-maternelle 4 ans seraient moins vulnérables que les enfants-témoins. L'intervalle de confiance à 95 % suggère en effet une valeur inférieure à 1 pour le rapport de cotes comparant les enfants-maternelle 4 ans aux enfants-témoins.

La figure 3.1 illustre la proportion espérée⁴⁶ d'enfants vulnérables selon le plus haut diplôme obtenu par l'un ou l'autre des parents pour les enfants-maternelle 4 ans et les enfants-témoins.

Figure 3.1

Proportion espérée¹ d'enfants considérés comme vulnérables dans le domaine de la maturité affective selon le plus haut diplôme obtenu par la mère ou le père, enfants de la population de comparaison, Québec, 2022



1. Correspond à la moyenne des valeurs espérées du modèle de régression logistique multiple ajusté.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

46. Pour une catégorie de variable dépendante, la proportion espérée correspond à la moyenne des valeurs espérées du modèle de régression logistique multiple ajusté.

3.2.2 Transition à la maternelle 5 ans selon le parent et fréquentation de la maternelle 4 ans selon les analyses comparatives

Une association significative entre la transition à la maternelle 5 ans et la fréquentation de la maternelle 4 ans à temps plein a été détectée dans l'analyse bivariée (tableau 3.4), tout comme c'était le cas dans l'analyse

descriptive. Cette association demeure significative lorsqu'on tient compte du facteur *Problème de santé des parents limitant les activités quotidiennes à la maison*, qui est le seul facteur confondant significatif. En effet, l'analyse multivariée montre que les enfants-maternelle 4 ans ont une plus faible probabilité que les enfants-témoins d'avoir vécu un passage à la maternelle 5 ans plutôt difficile du point de vue des parents ($RC_{\text{ajusté}} = 0,71$)

Tableau 3.4

Rapports de cotes liant la fréquentation de la maternelle 4 ans et la transition à la maternelle 5 ans, enfants de la population de comparaison, Québec, 2022

	Sans correction des effets de confusion			Avec correction des effets de confusion ¹		
	Rapport de cotes	Intervalle de confiance (IC à 95 %)		Rapport de cotes ajusté	Intervalle de confiance (IC à 95 %)	
		de	à		de	à
Maternelle 4 ans	0,70 [†]	0,56	0,88	0,71 [†]	0,56	0,90
Témoins	1,00	...	...	1,00	...	...

† Rapport de cotes significativement différent de 1,00 ($p < 0,05$). Le rapport de cotes inférieur à 1 indique que les enfants-maternelle 4 ans sont significativement moins susceptibles d'avoir eu un passage à la maternelle plutôt difficile que les enfants-témoins.

... Indique qu'un intervalle de confiance n'est pas construit pour le rapport de cotes de la catégorie de référence.

1. L'analyse multivariée permet la correction des effets de confusion de la caractéristique suivante : problème de santé des parents limitant les activités quotidiennes à la maison.

Source : Institut de la statistique du Québec, *Étude pour l'évaluation de la maternelle 4 ans à temps plein 2022*.

Conclusion

Synthèse des résultats

Le principal objectif de l'*Étude pour l'évaluation de la maternelle 4 ans* consiste à vérifier si la fréquentation de la maternelle 4 ans à temps plein est associée au développement de l'enfant et à une transition réussie à la maternelle 5 ans chez les enfants qui ont fréquenté la maternelle 5 ans en 2021-2022 dans les écoles qui avaient offert la maternelle 4 ans l'année précédente. Dans cette étude, le niveau de développement est mesuré par la vulnérabilité relative dans cinq domaines de développement ainsi que par deux mesures composites de développement (la vulnérabilité dans au moins un domaine de développement et la vulnérabilité dans au moins deux domaines de développement). La transition à la maternelle 5 ans est quant à elle mesurée par un indicateur relatif qui indique si le passage de l'enfant à la maternelle 5 ans est considéré comme plutôt difficile selon la perception des parents.

Des informations ont été recueillies pour l'EEM4 dans le cadre de l'EQPPEM menée auprès des parents. L'échantillon de l'EEM4 a été constitué, d'une part, des enfants ayant fréquenté la maternelle 4 ans à temps plein en 2020-2021 (enfants-maternelle 4 ans) et, d'autre part, des enfants qui ne l'ont pas fréquentée (enfants-témoins), mais qui possèdent certaines caractéristiques semblables. Les informations récoltées ont été jumelées à celles recueillies dans le cadre d'une enquête menée auprès du personnel enseignant sur le développement des enfants de maternelle 5 ans (EQDEM 2022). L'ensemble de ces informations permet de comparer les indicateurs de développement et de transition à la maternelle 5 ans des deux groupes d'enfants.

Tout a été mis en œuvre lors de l'échantillonnage pour contrôler les caractéristiques des deux sous-populations d'enfants, assurer le meilleur « pairage » possible des enfants-maternelle 4 ans avec les enfants-témoins et tenir compte de la non-réponse éventuelle aux enquêtes afin que des échantillons de répondants de taille suffisante soient obtenus. En vue de mieux cerner l'association entre la fréquentation de la maternelle 4 ans et le développement des enfants ou leur transition, une population de comparaison a été définie. Une pondération particulière pour la comparaison, notamment par la standardisation des enfants-témoins à la sous-population de référence des enfants-maternelle 4 ans, contribue également à la qualité des analyses comparatives présentées dans le présent rapport.

Pour répondre à l'objectif principal de l'EEMA, on a d'abord mené des analyses bivariées entre chaque indicateur à l'étude et la fréquentation de la maternelle 4 ans à temps plein. Ces analyses n'ont pas révélé de différences significatives entre les enfants-maternelle 4 ans et les enfants-témoins en ce qui concerne la vulnérabilité dans chacun des cinq domaines de développement ainsi que la vulnérabilité dans au moins un domaine ou dans au moins deux domaines de développement.

Les analyses bivariées montrent cependant une différence significative entre les enfants-maternelle 4 ans et les enfants-témoins en ce qui a trait à la transition à la maternelle 5 ans.

Pour s'assurer que l'association éventuelle entre les indicateurs de développement et la fréquentation de la maternelle 4 ans à temps plein ne soit pas masquée par certains facteurs confondants, on a ensuite tenu compte des caractéristiques associées à la fois au développement et à la fréquentation de la maternelle 4 ans dans les analyses par l'utilisation de régressions logistiques multivariées (analyses multivariées). De même, ces caractéristiques ont aussi été considérées pour vérifier si l'association entre la transition à la maternelle 5 ans et la fréquentation de la maternelle 4 ans à temps plein demeure lorsque ces facteurs potentiellement confondants sont contrôlés.

Rappelons que dans le cas précis de l'EEM4, les analyses multivariées ne visent pas à trouver les facteurs qui expliquent le mieux la vulnérabilité relative des enfants (modèle explicatif) ou ceux qui permettent de la prédire (modèle prédictif). Les analyses cherchent plutôt à vérifier si l'association détectée ou non entre la variable indépendante d'intérêt (fréquentation de la maternelle 4 ans à temps plein ou non) et les indicateurs de développement des enfants ou de la transition à la maternelle 5 ans, peut être modifiée lorsque différents facteurs potentiellement confondants sont inclus dans le modèle.

Dans le cas de la transition à la maternelle 5 ans, l'association avec la fréquentation de la maternelle 4 ans à temps plein demeure significative. Pour certains domaines de développement, le lien entre les indicateurs et la fréquentation de la maternelle 4 ans à temps plein devient significatif lorsque des facteurs confondants sont contrôlés.

Transition à la maternelle 5 ans

Le contrôle du facteur *Problème de santé des parents limitant les activités quotidiennes à la maison* dans l'analyse a permis de confirmer que la différence entre les enfants-maternelle 4 ans et les enfants-témoins demeure significative pour ce qui est de la transition à la maternelle 5 ans. Les enfants ayant fréquenté la maternelle 4 ans en 2020-2021 seraient moins susceptibles que les enfants-témoins d'avoir eu un passage à la maternelle 5 ans plutôt difficile selon la perception des parents.

Développement cognitif et langagier

Le contrôle des facteurs *Plus haut diplôme obtenu par l'un ou l'autre des parents* et *Niveau de revenu du ménage* a permis de relever une différence entre les enfants-maternelle 4 ans et les enfants-témoins dans le domaine « Développement cognitif et langagier ». En effet, la probabilité d'être considéré comme vulnérable dans ce domaine est plus faible chez les enfants qui ont fréquenté la maternelle 4 ans l'année précédente que chez les enfants qui ne l'ont pas fréquentée.

Habiletés de communication et connaissances générales

Le contrôle des facteurs *Durée cumulative en mois en services de garde*, *Langues parlées le plus souvent à la maison* et *Niveau de revenu du ménage* a montré que la probabilité d'être vulnérable dans le domaine « Habiletés de communication et connaissances générales » n'est pas uniforme selon le niveau de revenu du ménage. Chez les enfants vivant dans des familles à faible revenu, les enfants-maternelle 4 ans sont moins susceptibles d'être vulnérables dans ce domaine, alors que chez les enfants vivant dans des familles à revenu élevé, on ne détecte toujours pas de différence significative entre les enfants-maternelle 4 ans et les enfants-témoins.

Maturité affective

Le contrôle des facteurs *Nombre de déménagements vécus au cours des cinq dernières années*, *Niveau d'engagement relatif dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans* et *Plus haut diplôme obtenu par l'un ou l'autre des parents* a permis de faire ressortir que la probabilité d'être vulnérable dans le domaine « Maturité affective » est modulée selon le *Plus haut diplôme obtenu par l'un ou l'autre des parents*. Chez les enfants dont l'un ou l'autre parent a au plus un diplôme d'études secondaires, l'Intervalle de confiance à 95 % pour le rapport de cotes ajusté suggère une tendance selon laquelle les enfants-maternelle 4 ans seraient moins susceptibles d'être considérés comme vulnérables comparativement aux enfants-témoins. Chez les enfants dont l'un ou l'autre parent a au moins un diplôme universitaire, les enfants-maternelle 4 ans ont une plus grande probabilité d'être considérés comme vulnérables que les enfants-témoins.

Dans certains cas, la maternelle 4 ans semble contribuer positivement au développement des enfants-maternelle 4 ans en général ; par exemple en ce qui a trait au développement dans le domaine « Développement cognitif et langagier ». De plus, cette contribution est à l'avantage des groupes « défavorisés » notamment par la scolarité des parents ou par le niveau de revenu du ménage. C'est le cas pour le développement dans les domaines « Maturité affective » et « Habiletés de communication et connaissances générales ».

Portée, limites et discussion

Tout a été mis en œuvre pour assurer la qualité et la représentativité de l'EEM4. La disponibilité d'une base de sondage construite à partir de fichiers administratifs récents du MEQ a par exemple permis une bonne identification des enfants-maternelle 4 ans et des enfants-témoins de la population visée. En plus d'offrir une excellente couverture, cette base comporte plusieurs renseignements facilitant le contact avec les parents des enfants sélectionnés. D'ailleurs, la possibilité de réaliser une stratification fine des enfants de la population visée assure l'élaboration d'un plan d'échantillonnage efficace qui répond aux objectifs de comparaison poursuivis par l'EEM4 ; les échantillons des deux sous-populations d'enfants visés ont ainsi pu être constitués de manière optimale.

Les procédures inférentielles utilisées ont aussi fait l'objet d'une attention particulière. Les pondérations utilisées ont été calculées de façon à assurer la fiabilité de l'inférence à la population visée des enfants de maternelle 5 ans qui, en 2021-2022, ont fréquenté les écoles qui comptaient au moins trois enfants ayant fréquenté la maternelle 4 ans à temps plein l'année précédente. Les mesures de redressement effectuées réduisent au minimum les biais potentiels dus à la non-réponse à l'EQDEM 2022 et à l'EEM4, qui est du reste assez faible.

Comme c'est le cas pour toutes les enquêtes populationnelles, bien que toutes les précautions possibles aient été prises pour que la qualité des données soit assurée et que les biais soient réduits au minimum, certaines limites doivent être considérées. Par exemple, il n'est pas possible de garantir l'exactitude des réponses fournies par

les parents qui ont pu avoir de la difficulté à se souvenir des événements survenus il y a plusieurs mois, voire quelques années, comme c'est le cas pour les questions relatives à l'âge, en années et en mois, au début ou à la fin d'un événement tel que la fréquentation des services de garde. De plus, certains parents ne connaissaient peut-être pas certaines informations ; cela peut avoir été le cas de parents adoptifs au sujet des premières années de vie de leur enfant. Il faut aussi mentionner que les parents apprécient de façon subjective certains phénomènes. C'est le cas par exemple de l'appréciation des parents relativement à l'intégration de leur enfant à la maternelle 5 ans. À cela s'ajoute le fait que l'EEM4 est aussi tenue par les limites de l'EQDEM 2022 puisqu'elle utilise les données de cette enquête.

L'EEM4 est une enquête transversale⁴⁷ qui permet de faire ressortir les liens entre deux variables et de déceler des différences entre des groupes de population. Ce genre d'étude ne permet toutefois pas d'établir de lien de causalité entre les caractéristiques étudiées. Par exemple, la différence constatée à l'analyse comparative entre le développement cognitif et langagier des enfants-maternelle 4 ans et celui des enfants-témoins ne signifie pas que la maternelle 4 ans est la cause de cet écart et que c'est grâce à ce programme que ces enfants ont un meilleur développement dans ce domaine. L'EEM4 ne permet pas plus d'affirmer que la fréquentation de la maternelle 4 ans est la cause du déficit en matière de maturité affective des enfants-maternelle 4 ans comparativement aux enfants-témoins dont l'un ou l'autre des parents a au moins un diplôme universitaire.

En lien avec certains objectifs de l'étude, les analyses descriptives présentées dans le recueil statistique⁴⁸ s'appuient sur des méthodes d'analyse bivariée ; elles ne tiennent pas compte de facteurs confondants. Les proportions produites dans le recueil statistique fournissent ainsi une bonne description de deux sous-populations d'enfants visés par l'EEM4. Le présent rapport offre pour sa part une comparaison optimale des enfants-maternelle 4 ans et des enfants-témoins ; cette comparaison comporte tous les éléments accessibles pour mieux « cerner » les différences associées à la fréquentation de la maternelle 4 ans.

47. Une enquête transversale mesure un phénomène au cours d'une période donnée.

48. Le recueil statistique de l'EEM4 peut être consulté en ligne : (lien à venir).

Les données de l'EEM4 ont aussi fait l'objet d'analyses multivariées afin de mieux répondre aux objectifs de comparaison des enfants ayant fréquenté la maternelle 4 ans à temps plein en 2020-2021 avec les enfants-témoins. Par ce moyen, les effets potentiellement confondants de certains facteurs exogènes ont pu être contrôlés de façon plus appropriée, même si certaines limites demeurent. Par exemple, la sous-population des enfants-témoins est composée d'enfants pouvant avoir des profils très variés de parcours préscolaires. Certains enfants peuvent avoir fréquenté d'autres programmes préscolaires publics tels que le programme Passe-Partout ou la maternelle 4 ans à temps partiel, d'autres encore peuvent avoir fréquenté des services de garde ou être restés à la maison jusqu'à leur entrée à l'école. Cette diversité pourrait faire en sorte que la « forte » vulnérabilité des uns soit compensée par la « faible » vulnérabilité des autres dans cette sous-population (paradoxe de Simpson). Bien que certaines variables concernant le parcours préscolaire des enfants avant l'entrée à la maternelle 5 ans aient pu être contrôlées dans l'analyse multivariée, l'EEM4 ne permet pas de préciser si le fait que des différences significatives soient moins présentes pourrait être lié à ce « rééquilibrage » de la vulnérabilité relative parmi les enfants-témoins.

En outre, la variabilité associée aux estimations peut être déterminante pour qu'un écart entre deux proportions soit statistiquement significatif. Cette variabilité dépend du nombre de répondants qui permet de produire les estimations : plus petit est le nombre de répondants, plus grande sera la variabilité et plus grand sera l'écart susceptible d'être détecté. On peut ainsi observer qu'un écart d'une même amplitude soit significatif chez un groupe d'enfants et non significatif chez un autre groupe. C'est par exemple ce qui est observé dans la comparaison du développement des enfants-maternelle 4 ans et des enfants-témoins dans le domaine « Maturité affective » : des écarts de même ampleur sont estimés chez ceux dont les parents ont au plus un diplôme d'études secondaires et chez ceux dont au moins un des parents a un diplôme universitaire. Cet écart est significatif chez ces derniers, mais non significatif chez les premiers.

Par ailleurs, il faut dire que le niveau de développement des enfants avant la fréquentation de la maternelle 4 ans n'a pas été mesuré. Il n'est alors pas possible d'affirmer hors de tout doute que le fait qu'on ne détecte pas de différences significatives peut être considéré comme un apport positif de la maternelle 4 ans à temps plein. Les enfants-maternelle 4 ans auraient peut-être été plus vulnérables s'ils n'avaient pas fréquenté le programme. Pour valider l'effet de mise à niveau, il faudrait étudier le niveau de développement de ces enfants avant et après leur entrée à la maternelle 4 ans.

L'EEM4 pourrait comporter des biais de « sélection ». Par exemple, la plupart des parents qui décident que leur enfant fréquente la maternelle 4 ans à temps plein pourraient avoir un préjugé favorable envers l'école ou du moins un désir de faciliter la transition scolaire de leur enfant. Ils seraient ainsi ouverts à mettre en place des actions visant à préparer leur enfant à son entrée à l'école. Bien que certaines variables concernant les pratiques parentales en littératie, qui sont probablement liées à ce préjugé favorable envers l'école, aient pu être contrôlées avec l'analyse multivariée, l'EEM4 ne permet pas la séparation complète de ce biais de « sélection » des bienfaits du programme lui-même.

Finalement, rappelons que les indicateurs de vulnérabilité mesurent cinq domaines du développement des enfants en contexte scolaire. Il est possible que la fréquentation à temps plein de la maternelle 4 ans ait des effets positifs sur d'autres domaines du développement non mesurés par l'IMDPE. Il faut d'ailleurs rappeler que pour les indicateurs de l'IMDPE, il est possible qu'il existe une relation entre la vulnérabilité et la fréquentation de la maternelle 4 ans, mais que celle-ci soit trop ténue pour qu'on puisse la détecter. Avec les échantillons disponibles, on s'est assuré d'avoir une puissance statistique de 80 % pour détecter des écarts de plus 2,8 % entre les enfants-maternelle 4 ans et les enfants-témoins sur le plan de la vulnérabilité par domaine de développement et des écarts de plus de 4 % pour la mesure composite de vulnérabilité dans au moins un ou deux domaines de développement. La puissance statistique est plus faible pour des écarts moins importants.

En résumé, l'analyse de l'ISQ comportant plusieurs moyens de contrôle des biais de confusion potentiels suggère que la fréquentation de la maternelle 4 ans à temps plein « aiderait » davantage les enfants « défavorisés » par la scolarité des parents ou par le revenu du ménage dans lequel ils vivent. Elle suggère aussi que, globalement, les enfants des deux sous-populations arrivent à la maternelle 5 ans avec des écarts de développement tenus si on considère la puissance statistique de l'EEM4, ce qui en soi pourrait constituer une atteinte des objectifs de ce programme préscolaire : la maternelle 4 ans à temps plein pourrait avoir permis de réduire les écarts de développement des enfants de maternelle 5 ans en élevant le niveau des enfants qui l'ont fréquentée au même niveau que les autres qui auraient pu être favorisés par d'autres facteurs.

Annexe 1

Différences méthodologiques entre l'EQPPEM 2022 et l'EEM4 2022

L'EEM4 est une étude cas-témoins conçue pour comparer deux sous-populations sur la base d'une variable indépendante d'intérêt, à savoir la fréquentation de la maternelle 4 ans à temps plein. L'EQPPEM 2022, quant à elle, vise à fournir des informations complémentaires sur les facteurs associés à la vulnérabilité relative des enfants de maternelle 5 ans à l'échelle provinciale. La définition de la population visée, les méthodes utilisées pour le traitement et l'analyse des données de l'EEM4 diffèrent donc de celles de l'EQPPEM 2022. Ces différences peuvent expliquer des différences dans les résultats de chacune des enquêtes.

Différences par rapport à la population visée

L'EQPPEM 2022 vise l'ensemble des enfants ayant fréquenté la maternelle 5 ans en 2021-2022 (environ 87 000 enfants), alors que L'EEM4 vise seulement ceux de ces enfants qui ont fréquenté la maternelle 5 ans dans une école comptant au moins trois enfants ayant fréquenté la maternelle 4 ans à temps plein l'année précédente (environ 47 000 enfants).

À des fins de comparaison, l'EEM4 exclut aussi de la population visée pour la comparaison les enfants se trouvant en situation d'admissibilité exceptionnelle (enfants en situation de reprise de la maternelle 5 ans, de passage précoce à la maternelle 5 ans ou d'entrée à un âge inhabituel).

Ces différences de définition peuvent avoir les incidences suivantes :

- L'EQPPEM 2022 inclut dans le groupe des enfants-maternelle 4 ans les enfants qui ont fréquenté la maternelle 4 ans à temps plein l'année précédente et qui étaient inscrits dans des écoles comptant moins de trois enfants-maternelle 4 ans ;
- Le nombre d'enfants n'ayant pas fréquenté la maternelle 4 ans à temps plein l'année précédente est plus important dans l'EQPPEM 2022 ;
- Presque tous les enfants en situation d'admissibilité exceptionnelle sont dans le groupe d'enfants qui n'ont pas fréquenté la maternelle 4 ans à temps plein l'année précédente. Cela aurait une incidence sur les proportions d'enfants vulnérables parmi les enfants-témoins.

Différences par rapport aux échantillons de répondants

Pour l'EQPPEM 2022, tous les enfants ayant fréquenté la maternelle 4 ans à temps partiel ou encore à temps plein dans une école comptant moins de trois enfants-maternelle 4 ans⁴⁹ sont dans le groupe des enfants ayant fréquenté la maternelle 4 ans.

Pour l'EEM4, seuls les enfants ayant fréquenté la maternelle 4 ans à temps plein dans une école comptant au moins trois enfants-maternelle 4 ans sont considérés comme enfants-maternelle 4 ans.

49. Selon la définition d'enfants-maternelle 4 ans de l'EEM4.

Les estimations de l'EQPPEM 2022 sont produites à partir d'un échantillon de 14 895 enfants, alors que celles de l'EEM4 sont obtenues à partir d'un échantillon d'un peu plus de 3 900 enfants. L'EQPPEM présente donc une meilleure puissance statistique que l'EEM4 pour détecter des écarts assez petits.

Différences par rapport au traitement des données

La pondération des données pour chacune des enquêtes est faite sur la base de son échantillon de répondants et de sa population visée. Les différences de taille de population et d'échantillon de répondants expliquent les différences de poids, même pour les enfants présents dans tous les échantillons.

De plus, pour l'EQPPEM et le recueil statistique de l'EEM4, les poids ont été redressés sur les totaux de l'âge et du sexe des enfants ainsi que sur celui des indicateurs de vulnérabilité de leurs populations respectives. Contrairement aux poids utilisés pour les estimations de l'EQPPEM 2022 et du recueil statistique de l'EEM4, les poids de comparaison n'ont pas été redressés sur les totaux des indicateurs de vulnérabilité, qui sont les mesures du phénomène à l'étude. Ici encore, les différences des totaux utilisés pour le redressement dans les différentes pondérations se répercutent sur les poids obtenus pour les estimations.

La pondération de comparaison utilisée dans l'EEM4 a permis de réduire les disparités liées aux caractéristiques des sous-populations des enfants-maternelle 4 ans et des enfants-témoins par la standardisation des poids à une population de référence. Cette stratégie a permis la comparaison des indicateurs de développement entre les enfants-maternelle 4 ans et les enfants-témoins dont les caractéristiques de population étaient similaires.

Différences par rapport aux analyses

La fréquentation ou non de la maternelle 4 ans à temps plein est la variable indépendante d'intérêt dans l'EEM4. Elle est donc omniprésente dans les analyses. Dans l'EQPPEM 2022, la fréquentation ou non de la maternelle 4 ans à temps plein est l'un des facteurs qui décrivent le parcours préscolaire. À l'instar des autres facteurs, elle est retenue dans la modélisation sur la base de ses performances.

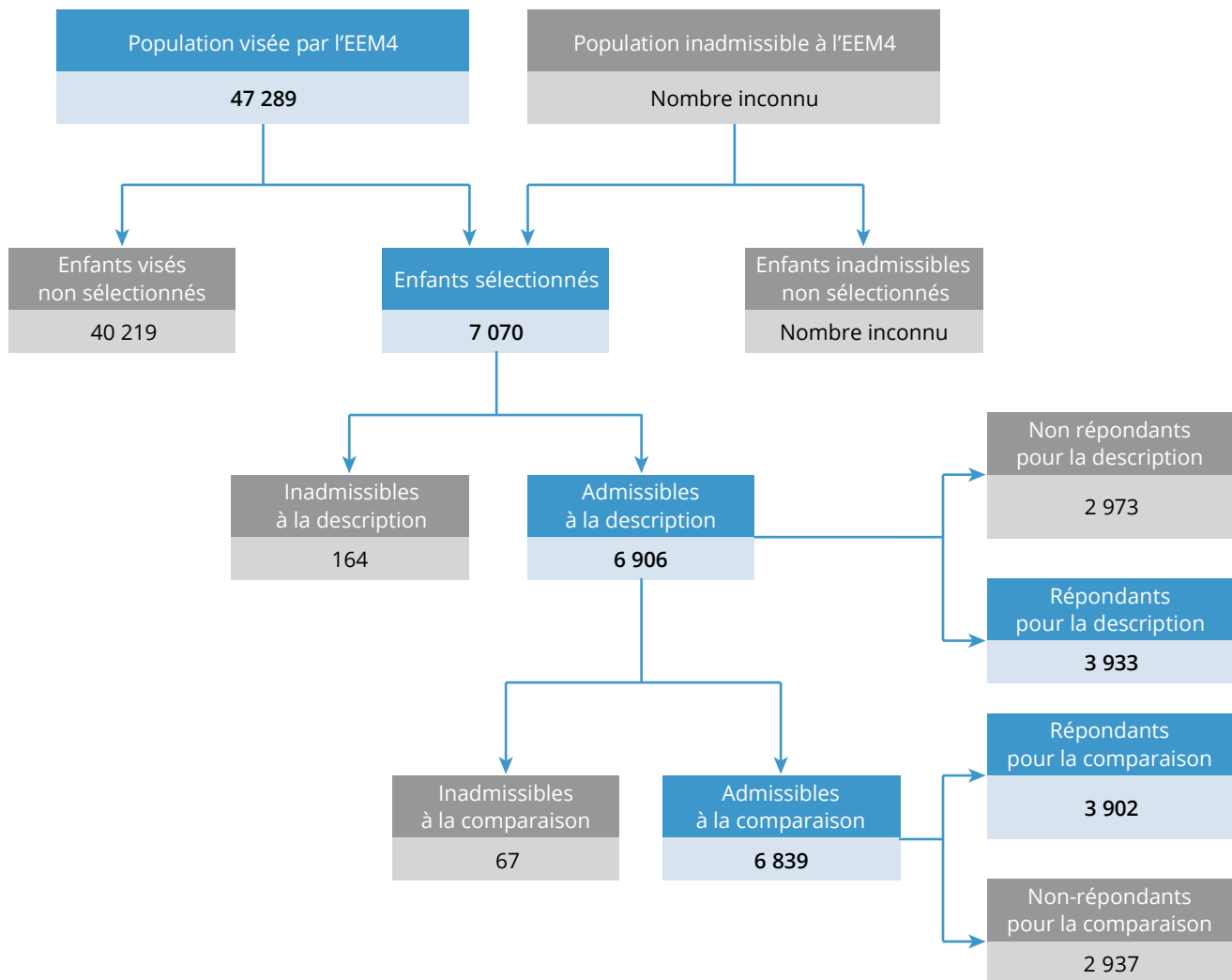
Au regard de ce qui précède, on peut retenir qu'il est important de préciser les objectifs souhaités pour déterminer l'inférence qui convient puisque les estimations et les conclusions peuvent sembler différentes :

- Lorsqu'on souhaite faire une inférence descriptive concernant l'ensemble des enfants ayant fréquenté la maternelle 5 ans en 2021-2022, sans tenir compte des caractéristiques populationnelles selon la fréquentation de la maternelle 4 ans à temps plein ou à temps partiel, les estimations et conclusions de l'EQPPEM 2022 sont recommandées ;
- Les proportions du recueil statistique de l'EEM4 sont recommandées lorsqu'on souhaite :
 - produire des estimations de proportions pour l'ensemble des enfants qui fréquentaient la maternelle 5 ans en 2021-2022 dans une école où il y avait au moins trois enfants ayant fréquenté la maternelle 4 ans en 2020-2021 ;
 - présenter, sans les comparer, les proportions de la population visée par l'EEM4 ;
 - présenter des estimations de la population et les comparer sans contrôler les disparités des caractéristiques entre la sous-population des enfants-maternelle 4 ans et la sous-population des enfants-témoins.
- Les rapports de cotes et les conclusions des analyses comparatives de l'EEM4 sont recommandés lorsque l'objectif est de comparer les sous-populations des enfants-maternelle 4 ans et des enfants-témoins de la façon la plus optimale, c'est-à-dire à caractéristiques similaires.

L'EQPPEM 2022 donne de l'information sur le lien entre la fréquentation de la maternelle 4 ans et la vulnérabilité pour la population des enfants de maternelle 5 ans en 2021-2022 en considérant les caractéristiques propres des « sous-populations » par rapport à la maternelle 4 ans et en contrôlant les totaux de l'indicateur de développement pour qu'ils correspondent à ceux de la population visée. Les analyses comparatives de l'EEM4, quant à elles, permettent de tirer des conclusions sur la comparaison à caractéristiques similaires des sous-populations (standardisation des sous-populations par la redéfinition de la population visée et de l'échantillon), sans contrôler les indicateurs de vulnérabilité qui sont alors à l'étude. Elle permet de réduire l'influence des différences qui seraient associées aux caractéristiques des sous-populations.

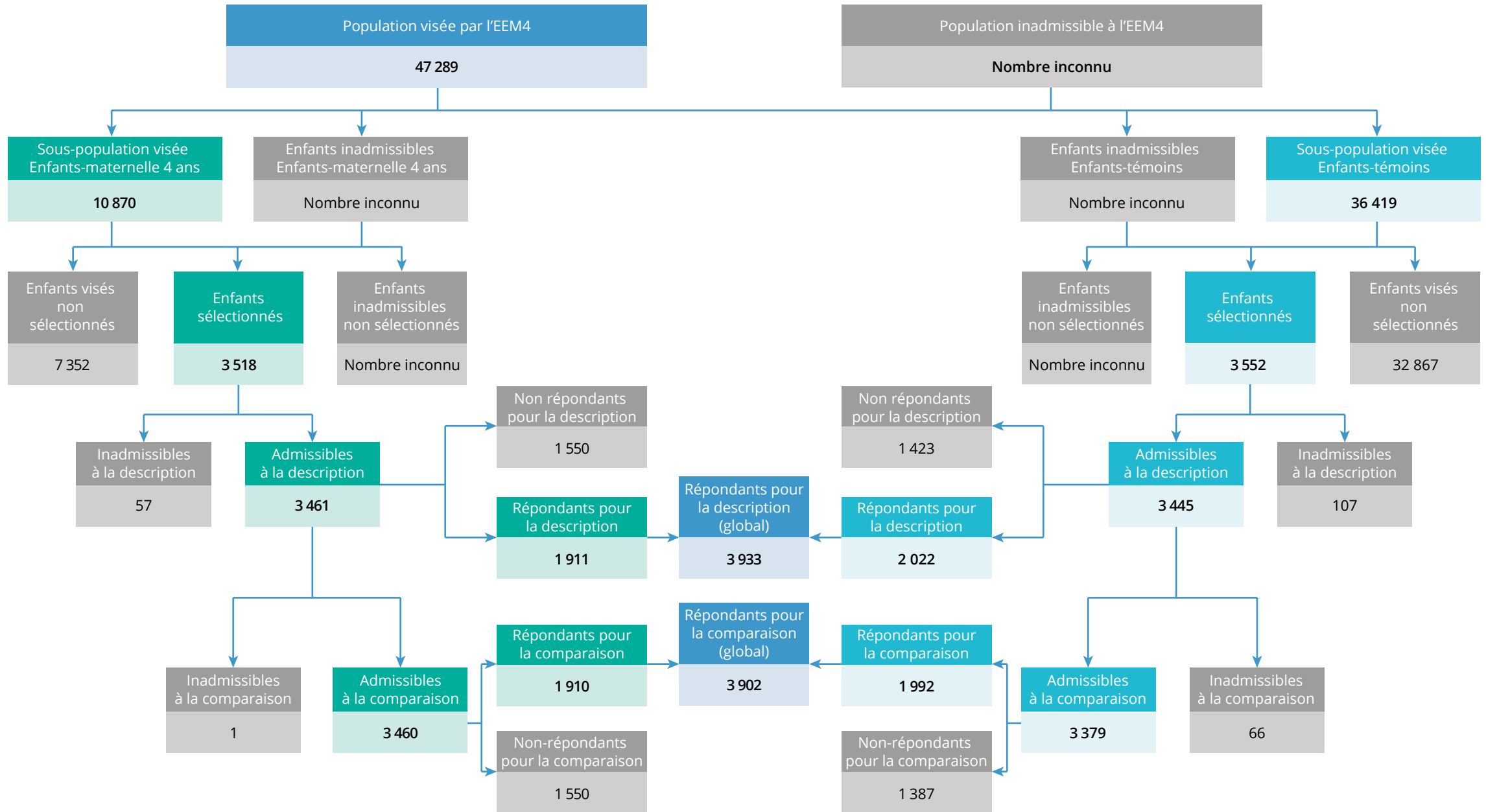
Annexe 2

Plan d'échantillonnage de l'EEM4 2022 – global



Annexe 3

Plan d'échantillonnage de l'EEM4 2022 – par sous-population



Annexe 4

Modalités des indicateurs retenus dans la modélisation

Indicateur	Modalités de l'indicateur
Durée cumulative en mois en services de garde	- De 0 à 2 mois - De 3 à 12 mois - De 13 à 24 mois - De 25 à 36 mois - De 37 à 48 mois 48 mois ou plus
Engagement dans des activités de littératie l'année avant l'entrée à la maternelle 5 ans	- Enfants souvent engagés dans des activités de littératie - Autres (enfants parfois, presque jamais ou jamais engagés dans des activités de littératie)
Langue parlée le plus souvent à la maison	- Français avec ou sans autre langue (sauf l'anglais) - Anglais avec ou sans autre langue (sauf le français) - Français et anglais avec ou sans autre langue - Autre(s) langue(s) seulement
Niveau de revenu du ménage	- Faible revenu - Revenu moyen-faible - Revenu moyen-élevé - Revenu élevé
Nombre de déménagements vécus au cours des cinq dernières années	- Aucun - Un - Deux - Trois ou plus
Plus haut diplôme obtenu par l'un ou l'autre des parents	- Aucun diplôme - Diplôme d'études secondaires - Diplôme d'études collégiales - Diplôme universitaire
Problème de santé des parents limitant les activités quotidiennes à la maison	- Les deux parents (ou le parent seul) sont limités par un problème ou un trouble chronique - Un des deux parents est limité par un problème ou un trouble chronique - Aucun parent n'est limité par un problème ou un trouble chronique

Bibliographie

- AUGER, Alexis, et Amélie GROLEAU (2023). *Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2022. Rapport statistique*, [En ligne], Québec, Institut de la statistique du Québec, Tome 1: *Portrait des caractéristiques, de l'environnement et du parcours préscolaire des enfants de maternelle 5 ans pour le Québec et ses régions*, 158 p. [statistique.quebec.ca/fr/fichier/parcours-prescolaire-enfants-maternelle-2022-rapport-statistique-tome-1.pdf] (Consulté le 5 février 2024).
- COURTEAU, Marilène, et Amélie DUCHARME (2023). *Enquête québécoise sur le développement des enfants à la maternelle 2022. Méthodologie de l'enquête*, [En ligne], Québec, Institut de la statistique du Québec, 78 p. [statistique.quebec.ca/fr/fichier/enquete-quebecoise-developpement-enfants-maternelle-2022-methodologie.pdf] (Consulté le 23 février 2024).
- DUCHARME, Amélie, Jesseca PAQUETTE et Sonia DALY (2023). *Enquête québécoise sur le développement des enfants à la maternelle 2022. Portrait statistique pour le Québec et ses régions administratives*, [En ligne], Québec, Institut de la statistique du Québec, 153 p. [statistique.quebec.ca/fr/fichier/enquete-quebecoise-developpement-enfants-maternelle-2022.pdf] (Consulté le 5 mars 2024).
- ELTINGE, John L., et Ibrahim 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], vol. 23, n° 1, p. 37-45. [www150.statcan.gc.ca/n1/fr/pub/12-001-x/1997001/article/3103-fra.pdf?st=1aQIUklh] (Consulté le 12 avril 2024).
- GAMACHE, Philippe, Denis HAMEL et Christine Blaser (2019). *L'indice de défavorisation matérielle et sociale : en bref. Guide méthodologique*, [En ligne], Québec, Institut national de santé publique du Québec, 10 p. [www.inspq.qc.ca/sites/default/files/santescope/indice-defavorisation/guidemethodologiquefr.pdf] (Consulté le 11 mai 2024).
- GOUVERNEMENT DU QUÉBEC (2022, mis à jour le 21 janvier). *Maternelle 4 ans*, [En ligne]. [www.quebec.ca/education/prescolaire-primaire-et-secondaire/maternelle/] (Consulté le 15 mars 2024).
- GROLEAU, Amélie, et Alexis AUGER (2023). *Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2022. Rapport statistique*, [En ligne], Québec, Institut de la statistique du Québec, Tome 2 : *Mieux comprendre la vulnérabilité des enfants de maternelle 5 ans : les facteurs associés*, 110 p. [statistique.quebec.ca/fr/fichier/parcours-prescolaire-enfants-maternelle-2022-rapport-statistique-tome-2.pdf] (Consulté le 5 janvier 2024).
- HAZIZA, David, et Jean-François 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)].
- HINES, William H., et autres (2003). *Probabilités et statistiques pour ingénieurs*, Québec, Chenelière Éducation, 587 p.

- INSTITUT DE LA STATISTIQUE DU QUÉBEC (2024). *Étude pour l'évaluation de la maternelle 4 ans à temps plein. Cahier technique*, Québec, L'Institut, 440 p. [Document interne non publié].
- KELELEKELA, Mbuyi, et Amélie GROLEAU (2023). *Enquête québécoise sur le parcours préscolaire des enfants de maternelle 2022. Méthodologie de l'enquête*, [En ligne], Québec, Institut de la statistique du Québec, 28 p. [statistique.quebec.ca/fr/fichier/eqppem-2022-methodologie.pdf] (Consulté le 16 janvier 2024).
- KORN, E. L., et B. I. GRAUBARD (1999). *Analysis of Health Surveys*, New York, John Wiley & Sons, 382 p.
- LAPOINTE, France, et Micha SIMARD (2019). *Étude pour l'évaluation du programme Passe-Partout 2017. Méthodologie de l'étude*, Québec, Institut de la statistique du Québec, 84 p. [Document interne non publié].
- LAPOINTE, France, et Micha SIMARD (2019). *Étude pour l'évaluation du programme Passe-Partout 2017. Rapport d'analyse – Volets A et B combinés*, Québec, Institut de la statistique du Québec, 34 p. [Document interne non publié].
- LEGAULT, Guy (1993). *Services éducatifs donnés aux enfants de quatre ans de milieu économiquement faible. Effet sur la diplomation*, Québec, Ministère de l'Éducation, 12 p.
- MINISTÈRE DE L'ÉDUCATION, DU LOISIR ET DU SPORT (MELS) [Québec] (2007). *L'organisation des services éducatifs aux élèves à risque et aux élèves handicapés ou en difficulté d'adaptation ou d'apprentissage (EHDA)*, [En ligne], Québec, Gouvernement du Québec, 25 p. [www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/adaptation_serv_compl/19-7065.pdf] (Consulté le 11 mai 2024).
- MINISTÈRE DE L'ÉDUCATION [Québec] (2023). *Évaluation de la maternelle 4 ans à temps plein. Points de vue d'enseignantes et d'enseignants, de membres de directions d'école et de conseillères pédagogiques. Rapport d'étape 2023*, [En ligne], Québec, Gouvernement du Québec, 95 p. [www.education.gouv.qc.ca/fileadmin/site_web/documents/PSG/recherche_evaluation/Rapport-maternelle-4-ans.pdf].
- MINISTÈRE DE L'ÉDUCATION ET DE L'ENSEIGNEMENT SUPÉRIEUR [Québec] (2017). *Rapport d'évaluation de la maternelle 4 ans à temps plein en milieu défavorisé*, [En ligne], Québec, Gouvernement du Québec, 63 p. [cdn-contenu.quebec.ca/cdn-contenu/adm/min/education/publications-adm/education/evaluations-programmes/Rapport-evaluation-maternelle-4ans-milieu-defavorise.pdf].
- SIMARD, Micha, Amélie LAVOIE et Nathalie AUDET (2018). *Enquête québécoise sur le développement des enfants à la maternelle 2017*, [En ligne], Québec, Institut de la statistique du Québec, 126p. [statistique.quebec.ca/fr/fichier/enquete-quebecoise-sur-le-developpement-des-enfants-a-la-maternelle-2017-portrait-statistique-pour-le-quebec-et-ses-regions-administratives.pdf] (Consulté le 13 février 2024).
- 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. (Consulté le 22 novembre 2021).
- SHAH, Babubhai. V. (1976). *The F-statistic with Satterthwaite correction for numerator degrees of Freedom*, [s. l.], Research Triangle Institute. [Document interne non publié].

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

statistique.quebec.ca