

Estimer l'âge gestationnel à la naissance : étude de dérivation-validation en population générale

M. L. Urquia, Ph. D. (1, 2); T. A. Stukel, Ph. D. (2, 3); K. Fung, M. Sc. (2); R. H. Glazier, M.D. (1, 2, 3, 4, 5); J. G. Ray, M.D. (2, 3, 6)

Résumé

Introduction : Les données sur l'âge gestationnel (AG) à la naissance sont essentielles dans la recherche en santé périnatale et en santé du nourrisson, mais elles ne figurent pas toujours dans les bases de données administratives. Nous avons donc élaboré et validé un modèle de prédiction de l'AG pour les naissances simples à des fins d'études épidémiologiques.

Méthodologie : La dérivation de l'AG estimé a été calculée sur la base d'un échantillon de 130 328 nouveau-nés ayant vu le jour dans les hôpitaux de l'Ontario entre 2007 et 2009, au moyen d'une analyse de régression linéaire, avec plusieurs caractéristiques maternelles et infantiles comme variables prédictives (indépendantes). Le modèle a été validé sur un autre échantillon, de 130 329 nouveau-nés.

Résultats : La capacité discriminative du modèle linéaire selon le sexe et le poids à la naissance du nouveau-né était d'une approximation raisonnable pour les enfants nés avant la 37^e semaine de gestation ($r^2 = 0,67$; intervalle de confiance (IC) à 95 % : 0,65 à 0,68), mais pas pour les naissances à terme (37 à 42 semaines; $r^2 = 0,12$; IC à 95 % : 0,12 à 0,13). L'ajout de caractéristiques maternelles et infantiles supplémentaires n'a pas amélioré la discrimination du modèle.

Conclusion : Il est possible d'obtenir une estimation correcte de l'âge gestationnel à la naissance des enfants nés avant 37 semaines de gestation à l'aide de données localement disponibles sur le poids à la naissance et le sexe.

Mots-clés : âge gestationnel, naissance, nouveau-né, santé du nourrisson, dérivation, validation, prédiction, ensembles de données administratives, Ontario

Introduction

La gestation commence au moment de la conception et se termine à la naissance, mais on la calcule généralement à partir du premier jour des dernières menstruations. L'âge gestationnel (AG) est un prédicteur important de la morbidité et de la mortalité périnatales¹; il revêt une importance dans la datation en vue du dépistage génétique prénatal² et de la détermination du moment de l'exposition du fœtus à des agents

teratogènes^{3,4}. L'AG est également nécessaire pour déterminer correctement si un nouveau-né est petit ou gros pour son AG, tant pour la pratique clinique que pour la recherche en épidémiologie².

Dans les pays où les soins prénataux et maternels sont rares, la collecte de données statistiques de base sur les nouveau-nés peut être entravée par un manque d'information sur l'AG. Dans les pays

industrialisés, il arrive souvent que l'AG ne soit pas enregistré dans les bases de données administratives sur la santé³⁻⁵. Au Canada, comme tous les résidents permanents reçoivent leurs soins de santé dans le cadre d'un régime universel, dont les soins prénataux et périnataux et les soins aux nouveau-nés, une base de données administratives, la Base de données sur les congés des patients de l'Institut canadien d'information sur la santé (BDCP de l'ICIS), est reconnue comme une excellente source pour obtenir des estimations de population à des fins de recherche périnatale^{6,7}; toutefois, avant l'exercice 2002-2003, la BDCP de l'ICIS ne recueillait pas de données sur l'AG à la naissance en Ontario⁸, ce qui pourrait poser des problèmes dans certaines recherches sur les issues périnatales.

Cette étude vise l'élaboration et la validation d'un modèle de prédiction de l'AG pour les naissances simples qu'on puisse employer à des fins d'études épidémiologiques.

Méthodologie

Plan d'étude général

Nous avons utilisé une méthode analytique de dérivation-validation pour estimer l'AG à partir de données périnatales facilement accessibles. Nous avons réalisé une vaste étude en population générale de tous les nouveau-nés issus d'une grossesse unique et nés dans les hôpitaux de l'Ontario en 2007-2008 ou 2008-2009, période durant laquelle l'AG à la naissance a été systématiquement enregistré dans la BDCP de l'ICIS. La cohorte de dérivation était composée d'un échantillon aléatoire de la moitié des

Rattachement des auteurs

1. Centre for Research on Inner City Health, The Keenan Research Centre of the Li Ka Shing Knowledge Institute, Hôpital St. Michael, Toronto, Ontario, Canada

2. Institut de recherche en services de santé, Toronto, Ontario, Canada

3. Département de gestion et d'évaluation des politiques de santé, Université de Toronto, Toronto, Ontario, Canada

4. École de santé publique Dalla Lana, Université de Toronto, Toronto, Ontario, Canada

5. Département de médecine familiale et communautaire, Université de Toronto, Toronto, Ontario, Canada.

6. Département d'obstétrique et de gynécologie, Université de Toronto, Toronto, Ontario, Canada.

Correspondance : Marcelo Luis Urquia, Centre for Research on Inner City Health, The Keenan Research Centre of the Li Ka Shing Knowledge Institute, Hôpital St. Michael, 30, Bond Street, Toronto (Ontario) M5B 1W8; tél. : 416-864-6060, poste 77340; téléc. : 416-864-5558; courriel : marcelo.urquia@utoronto.ca

naissances vivantes survenues au cours de la même période. Cette cohorte a servi à créer un modèle prédictif fondé sur des caractéristiques infantiles. La cohorte de validation était formée de l'autre moitié des naissances et a servi à évaluer la capacité du modèle de dérivation à la naissance à prédire l'AG. Des études de simulation ont montré que la validation par échantillons fractionnés constitue une approche adaptée lorsque la taille de l'échantillon total est très grande, ce qui est le cas de notre étude ($n = 260\,657$)⁹.

Nous avons exclu de notre échantillon toutes les mortinaissances et les naissances multiples. Afin de réduire l'influence des possibles erreurs de données et observations aberrantes, nous avons également exclu les catégories suivantes : les nouveau-nés venus au monde à 23 semaines complètes de gestation ou moins et ceux nés à 43 semaines complètes de gestation ou davantage; les nouveau-nés dont la combinaison du poids à la naissance et de l'AG était peu probable sur le plan clinique¹⁰; ceux dont le séjour à l'hôpital a duré plus de 90 jours; ceux dont l'AG, le poids à la naissance ou le sexe n'étaient pas enregistrés; les enfants nés de mères âgées de moins de 16 ans ou de plus de 50 ans au moment de l'accouchement. Ont également été rejetées les observations aberrantes extrêmes de la distribution du poids à la naissance, c'est à dire les valeurs présentant un écart supérieur à deux fois ou plus l'intervalle interquartile¹¹.

Variables

En Ontario, l'AG est estimé en grande partie au moyen de la datation par échographie en début de grossesse. Depuis 2002, les services d'archives médicales des hôpitaux enregistrent l'AG selon la meilleure interprétation que fait le médecin traitant de l'ensemble des données cliniques qui figurent habituellement dans le dossier prénatal^{12,13}. De plus, le sexe du nouveau né et son poids précis à la naissance sont enregistrés dans la BDCP de l'ICIS¹⁴. Nous avons déterminé les anomalies congénitales et les maladies de la prématurité selon les codes CIM-10-CA¹⁵ inscrits dans les 25 champs de diagnostic des dossiers médicaux (tableau 1).

Analyse statistique

La dérivation de l'estimation de l'AG s'est déroulée en deux étapes¹⁶. À l'aide de la cohorte de dérivation, nous avons effectué une série d'analyses de régression linéaire et utilisé l'âge gestationnel (en nombre de semaines complétées) comme variable dépendante et plusieurs variables choisies a priori comme variables indépendantes (tableau 2).

Nous avons d'abord modélisé l'AG en utilisant une fonction spline cubique restreinte du poids à la naissance avec quatre degrés de liberté¹⁷. Nous avons ajouté au modèle de base le sexe du nouveau-né, les anomalies congénitales et chromosomiques ainsi que les maladies de la prématurité (syndrome de détresse respiratoire, leucomalacie cérébrale néonatale ou hémorragie intraventriculaire, rétinopathie de la prématurité et entérocolite nécrosante). Les détails de ces variables sont présentés dans le tableau 2.

Nous avons créé des modèles prédictifs en multipliant les coefficients de chaque variable indépendante dans les modèles de dérivation par les valeurs spécifiques qui constituent le profil de chaque individu dans la cohorte de validation. Nous avons mis à l'épreuve chaque modèle prédictif en utilisant l'AG *réel* de la cohorte de validation comme variable dépendante et l'AG *estimé* comme variable indépendante, les AG étant arrondis au nombre de semaines complètes le plus proche. Pour mesurer la discrimination du modèle, nous avons calculé le coefficient de détermination (r^2) et son intervalle de confiance (IC) à 95 %. Les modèles ont été validés pour la totalité de la cohorte de naissances et stratifiés selon le sexe du nouveau-né et selon le moment de la naissance (AG inférieur à 37 semaines et AG égal ou supérieur à 37 semaines). L'AG *réel* et l'AG *estimé* ont été représentés graphiquement selon leurs distributions de fréquence respectives (figure 1).

Nous avons généré graphiquement le taux de vrais positifs du modèle dérivé (c.-à-d. la proportion de nouveau-nés dont l'AG *réel* correspond à l'AG *dérivé*, dont l'AG *réel* varie d'une semaine par rapport à l'AG *dérivé* ou

dont l'AG *réel* varie de deux semaines par rapport à l'AG *dérivé*) sur l'axe des ordonnées et l'AG estimé sur l'axe des abscisses (figure 2).

Toutes les analyses ont été réalisées à l'aide du logiciel SAS (version 9.1, SAS Institute Inc., Cary, Caroline du Nord, États-Unis).

Résultats

Il y avait 281 406 dossiers de nouveau-nés en 2007-2008 et 2008-2009. Une fois exclues les mortinaissances, les naissances multiples et les observations aberrantes évidentes (7,4 %), notre ensemble de données contenait 260 657 dossiers de naissances simples. Les caractéristiques infantiles de la cohorte de dérivation et celles de la cohorte de validation étaient similaires (tableau 2).

Le modèle optimal comprenait une fonction spline cubique restreinte du poids à la naissance, exprimé en kilogrammes, ainsi que le sexe du nouveau-né. Le coefficient de détermination (r^2) pour ce modèle prédictif était de 0,44 (IC à 95 % : 0,43 à 0,45). L'ajout au modèle d'une anomalie congénitale ou chromosomique ou de maladies de la prématurité, ou encore la stratification selon le sexe du nouveau-né, n'ont eu aucune incidence appréciable sur le coefficient de détermination (tableau 3).

Lorsque les données ont été stratifiées selon le moment de la naissance, la capacité discriminative du modèle s'est révélée faible pour les enfants nés à terme (37 à 42 semaines; $r^2 = 0,12$; IC à 95 % : 0,12 à 0,13), mais bien meilleure pour les naissances avant terme (24 à 36 semaines; $r^2 = 0,67$; IC à 95 % : 0,65 à 0,68) (tableau 3). L'ajout de caractéristiques au modèle des enfants nés avant terme, comme l'admission dans une unité néonatale de soins intensifs, la durée d'hospitalisation du nouveau-né, la prééclampsie maternelle ou hypertension gestationnelle et le mode d'accouchement, n'a pas amélioré le coefficient de détermination (données non présentées).

Jusqu'à environ 36 semaines de gestation, la concordance était élevée dans les courbes

* Classification statistique internationale des maladies et des problèmes de santé connexes, dixième révision, améliorations canadiennes.

TABLEAU 1

Codes CIM-10-CA utilisés pour déterminer les anomalies congénitales, les maladies de la prématureté, les naissances multiples et les mortinassances dans l'étude portant sur les nouveau-nés vivants issus d'une grossesse unique, Ontario, 2007 à 2009

Variable	BDCP de l'ICIS source du dossier	CIM-10-CA
Toute anomalie congénitale ou chromosomique	Nourrisson	Q00-Q99
Maladies de la prématureté	Nourrisson	
Entérocolite nécrosante		P77
Syndrome de détresse respiratoire		P22
Leucomalacie cérébrale néonatale ou hémorragie intraventriculaire néonatale		P91.2, P52
Rétinopathie de la prématureté		H35.1
Grossesse multiple	Nourrisson	Q89.4, Z38.3-Z38.8
Grossesse multiple	Mère	O30, O31, Z37.2-Z37.7, Z38.3-Z38.8, Z37.9.0
Mort intra-utérine	Nourrisson	P95
Mort intra-utérine	Mère	O36.4, Z37.1, Z37.4, Z37.7

Abréviations : BDCP de l'ICIS, Base de données sur les congés des patients de l'Institut canadien d'information sur la santé; CIM-10-CA, *Classification statistique internationale des maladies et des problèmes de santé connexes*, dixième révision, améliorations canadiennes.

TABLEAU 2

Caractéristiques des nouveau-nés vivants issus d'une grossesse unique dans la cohorte de dérivation (n = 130 328) et dans la cohorte de validation (n = 130 329), Ontario, 2007 à 2009

Caractéristiques du nouveau-né	Cohorte de dérivation, n (%)	Cohorte de validation, n (%)
Sexe masculin	66 551 (51,06)	66 898 (51,33)
Âge gestationnel à la naissance		
À terme, 37 à 42 semaines	122 723 (94,16) ^a	122 760 (94,19) ^b
Avant terme, 24 à 36 semaines	7 605 (5,84)	7 569 (5,81)
Bien avant terme, 24 à 27 semaines	187 (0,14)	206 (0,16)
Poids moyen à la naissance ± É-T, grammes	3 392 ± 531	3 392 ± 532
Poids à la naissance ^c		
< 2 500 grammes	5 715 (4,39)	5 797 (4,45)
≥ 2 500 grammes	124 613 (95,61)	124 532 (95,55)
Anomalies congénitales ou chromosomiques ^d	5 655 (4,34)	5 677 (4,36)
Maladies de la prématureté ^{d,e}	7 587 (5,82)	7 771 (5,96)
Syndrome de détresse respiratoire	7 474 (5,73)	7 681 (5,89)
Leucomalacie cérébrale néonatale ou hémorragie intraventriculaire néonatale	206 (0,16)	207 (0,16)
Rétinopathie de la prématureté	111 (0,09)	112 (0,09)
Entérocolite nécrosante	62 (0,05)	62 (0,05)

Abréviations : CIM-10-CA, *Classification statistique internationale des maladies et des problèmes de santé connexes*, dixième révision, améliorations canadiennes; n, taille de l'échantillon; É-T, écart type.

^aL'âge gestationnel moyen (± É-T) à la naissance dans ce groupe était de 39,2 (± 1,14) semaines.

^bL'âge gestationnel moyen (± É-T) à la naissance dans ce groupe était de 39,2 (± 1,15) semaines.

^cLe poids moyen à la naissance (± É-T) était de 3 392 ± 531 grammes dans la cohorte de dérivation et de 3 392 ± 532 grammes dans la cohorte de validation.

^dAnomalies congénitales ou chromosomiques et maladies de la prématureté déterminées en fonction des codes CIM-10-CA tirés des dossiers d'hospitalisation.

^eDe nombreux nouveau-nés pouvaient être atteints de plus d'une maladie de la prématureté. C'est pourquoi le total des pourcentages n'est pas égal à 100.

de distribution de l'AG réel par rapport à l'AG dérivée, après quoi c'est une discordance marquée qui s'observait (figure 1). Dans le cas des naissances à terme, l'AG prédit ne permettait pas d'estimer correctement l'âge réel, particulièrement à 39 semaines, moment le plus fréquent de naissance (figure 1).

Le modèle de l'AG qui comprenait le sexe et le poids à la naissance du nouveau-né a eu une valeur prédictive positive de 34 % à 28 ± 1 semaine, de 67 % à 28 ± 2 semaines, de 47 % à 32 ± 1 semaine, de 74 % à 32 ± 2 semaines, de 60 % à 37 ± 1 semaine et de 85 % à 37 ± 2 semaines de gestation (figure 2).

Nous avons répété la validation en utilisant l'ensemble des données plutôt que les données de validation initiales seulement et les résultats sont demeurés les mêmes (données non présentées).

Analyse

Dans notre vaste étude de dérivation-validation portant sur une large population, le poids à la naissance et le sexe du nouveau-né pris ensemble ont permis d'obtenir une estimation correcte de l'AG chez les enfants nés avant 37 semaines de gestation, mais pas chez les enfants nés à terme.

L'ajout de caractéristiques maternelles et infantiles supplémentaires n'a pas amélioré le coefficient de détermination de notre modèle chez les enfants nés avant terme. D'autres chercheurs ont obtenu des résultats semblables lorsqu'ils ont élaboré des courbes de poids à la naissance¹⁸.

Un modèle restreint fondé sur le poids à la naissance et le sexe du nouveau-né comporte certains avantages, en ce sens que les deux variables sont saisies et enregistrées dans pratiquement tous les établissements cliniques, ceux des pays pauvres comme ceux des pays riches, en plus d'être inscrites dans de vastes ensembles de données administratives alors que l'AG n'y est pas disponible. Il convient de souligner que le poids à la naissance et le sexe du nouveau-né sont les deux variables principalement utilisées pour déterminer les valeurs de référence du poids à la naissance qui serviront à établir l'AG pour une population^{10,19,20}. Par

FIGURE 1

Comparaison entre l'âge gestationnel prédit en fonction du sexe et du poids à la naissance du nouveau-né (ligne pleine) et l'âge gestationnel réel à la naissance (ligne pointillée), données de validation (n = 130 329)

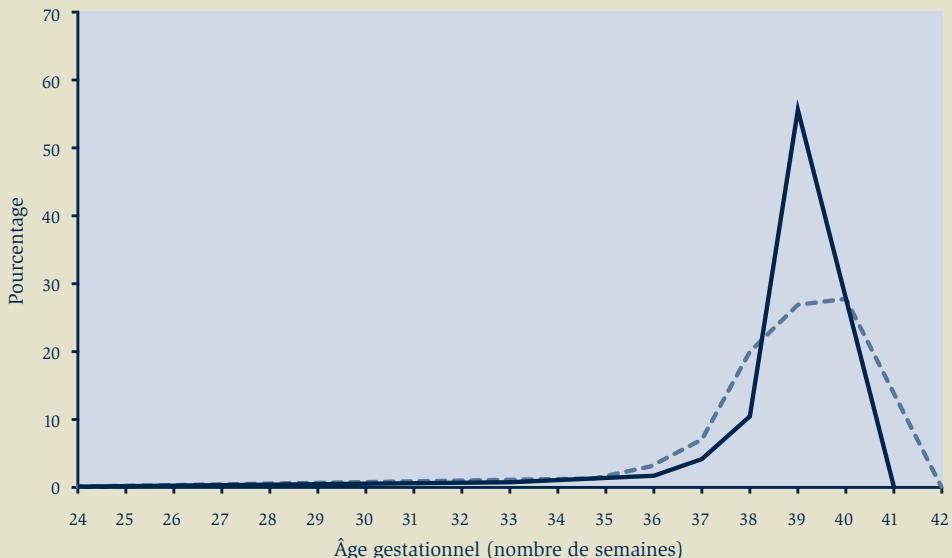


FIGURE 2

Concordance entre l'âge gestationnel dérivé et l'âge gestationnel réel des nouveau-nés vivants issus d'une grossesse unique en Ontario, de 2007-2008 à 2008-2009. Les courbes représentent le pourcentage de nouveau-nés dont l'âge gestationnel réel est le même que l'âge gestationnel dérivé (en bas), ou varie de ± 1 semaine (milieu) et de ± 2 semaines (en haut)

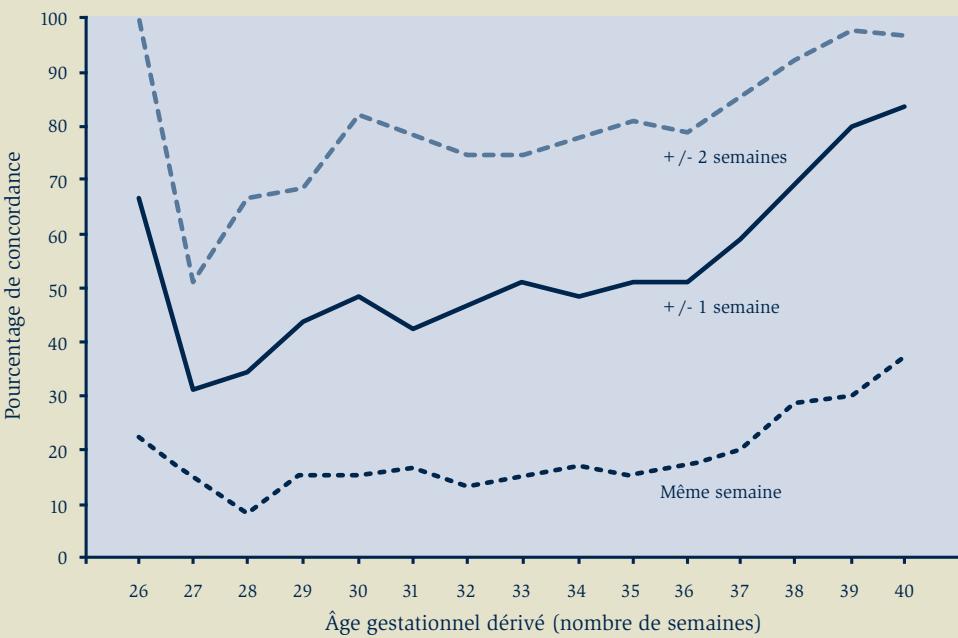


TABLEAU 3

Capacité discriminative d'un modèle créé d'après la cohorte de dérivation à prédire l'âge gestationnel à la naissance connu dans la cohorte de validation^a

Groupes	Variables du modèle	Coefficient de détermination, r^2 (IC à 95 %) ^b
Total (n = 130 329)	Poids à la naissance et sexe	0,44 (0,43 à 0,45)
	Poids à la naissance, sexe, anomalies congénitales ou chromosomiques et maladies de la prématurité ^c	0,45 (0,44 à 0,46)
Sexe		
Masculin (n = 66 898)	Poids à la naissance	0,46 (0,44 à 0,47)
	Poids à la naissance, anomalies congénitales ou chromosomiques et maladies de la prématurité ^c	0,47 (0,45 à 0,48)
Féminin (n = 63 431)	Poids à la naissance	0,43 (0,41 à 0,44)
	Poids à la naissance, anomalies congénitales ou chromosomiques et maladies de la prématurité ^c	0,44 (0,42 à 0,45)
Terme de naissance		
À terme, 37 à 42 semaines (n = 122 760)	Poids à la naissance et sexe	0,12 (0,12 à 0,13)
	Poids à la naissance, sexe, anomalies congénitales ou chromosomiques et maladies de la prématurité ^c	0,13 (0,12 à 0,13)
Avant terme, 24 à 36 semaines (n = 7 569)	Poids à la naissance et sexe	0,67 (0,65 à 0,68)
	Poids à la naissance, sexe, anomalies congénitales ou chromosomiques et maladies de la prématurité ^c	0,68 (0,67 à 0,70)

Abbreviations : CIM-10-CA, *Classification statistique internationale des maladies et des problèmes de santé connexes*, dixième révision, améliorations canadiennes; n, taille de l'échantillon.

^aLes cohortes sont composées de nouveau-nés vivants issus d'une grossesse unique en Ontario en 2007-2008 et 2008-2009.

^bLe poids à la naissance est modélisé selon une fonction spline cubique restreinte avec quatre degrés de liberté.

^cSyndrome de détresse respiratoire, leucomalacie cérébrale ou hémorragie périventriculaire néonatale, rétinopathie de la prématurité, entérocolite nécrosante, selon les codes CIM-10-CA.

conséquent, dans le cas où l'AG est absent des bases de données, nous recommandons l'utilisation des données sur le poids à la naissance et le sexe du nouveau-né pour estimer l'AG, et les chiffres sur le poids à la naissance obtenus localement pour construire les graphiques d'AG, notamment le 50^e percentile du poids à la naissance sexospécifique observé à chaque semaine d'AG. Les percentiles du poids à la naissance inférieurs (5^e, 10^e) et supérieurs (90^e, 95^e) pourraient également servir à exprimer la variabilité biologique de l'AG à un poids de naissance donné.

La faible corrélation que nous avons observée entre l'AG et le poids à la naissance après 36 semaines de gestation est digne de mention, étant donné que 94 % des enfants issus d'une

grossesse unique naissent à terme. La faible prédiction de l'AG à terme est due essentiellement à la grande variabilité du poids à la naissance à mesure que l'AG augmente. Par exemple, un graphique récent de poids à la naissance pour les nouveau-nés canadiens de sexe masculin indiquait une différence d'au moins 1 100 grammes entre le 10^e et le 90^e percentile du poids à la naissance entre 37 et 41 semaines de gestation²⁰. Cette dernière donnée témoigne de la grande variabilité du poids à la naissance dans l'intervalle « normal ». La meilleure prédiction de l'AG au cours des semaines de gestation précédentes témoigne d'une moins grande variabilité biologique. De plus, la pente du poids à la naissance est plus linéaire et plus forte aux AG inférieurs qu'à l'AG à terme²⁰.

Limites

Cette étude comporte des limites. Premièrement, nous nous sommes appuyés sur les codes CIM-10-CA contenus dans une base de données administratives dans laquelle les mesures des nouveau-nés n'ont pas été effectuées aux fins de l'étude. Deuxièmement, comme nous n'avons inclus que les enfants vivants issus d'une grossesse unique, notre approche pourrait ne pas s'appliquer aux grossesses multiples. Malheureusement, les courbes pour une population des poids à la naissance applicables aux naissances multiples demeurent peu nombreuses^{21,22}. Troisièmement, la base de données ne contenait pas d'information sur d'autres facteurs associés à la durée de gestation et au poids du nouveau-né comme l'origine ethnique des parents, l'anthropométrie maternelle ou les comportements liés à la santé durant la grossesse, facteurs qui pourraient servir à l'élaboration de graphiques spécifiques des poids à la naissance^{23,24}. L'inclusion de ces facteurs améliorerait certainement notre modèle prédictif^{25,26}. Quatrièmement, nous avons fondé nos analyses sur l'estimation clinique de l'AG (qui s'effectue généralement au moyen d'une échographie de datation en début de grossesse), qui, on le sait, diffère de l'estimation basée sur la date des dernières règles^{12,13}. Cette dernière mesure surestime les taux de naissance avant terme et après terme et présente des distributions bimodales de poids à la naissance entre 28 et 34 semaines de gestation^{20,25,27-29}. Reproduire notre méthode de validation en utilisant cette estimation comme référence pourrait appauvrir sa capacité de prédiction. Enfin, nous devons souligner que nos modèles n'ont pas été conçus pour estimer l'AG de nouveau-nés en particulier.

En conclusion, à défaut de données sur l'AG réel, il est possible d'estimer raisonnablement l'AG des nouveau-nés à l'échelle d'une population comme variable continue jusqu'à 36 semaines de gestation à l'aide du sexe et du poids à la naissance, et quoiqu'une incertitude importante soit inévitable, même après la prise en compte d'autres prédicteurs de l'AG.

Remerciements

L'étude a bénéficié du soutien de l'Institut de recherche en services de santé (IRSS), qui est financé par une subvention annuelle du ministère de la Santé et des Soins de longue

durée de l'Ontario (MSSLDO). Les positions, résultats et conclusions de l'étude n'engagent cependant que les auteurs et sont indépendants des sources de financement. Ils ne représentent pas la position officielle de l'IRSS ou MSSLDO et ne devraient par être perçus comme tels.

Références

1. Cheng YW, Nicholson JM, Nakagawa S, Bruckner TA, Washington AE, Caughey AB. Perinatal outcomes in low-risk term pregnancies: do they differ by week of gestation? *Am J Obstet Gynecol.* 2008; 199(4):370-7.
2. Behrman RE, Butler AS, editors; Committee on Understanding Premature Birth and Assuring Healthy Outcomes. Preterm birth: causes, consequences, and prevention. 1st ed. Washington (DC): National Academies Press; 2007.
3. Andrade SE, Raebel MA, Morse AN, Davis RL, Chan KA, Finkelstein JA, et al. Use of prescription medications with a potential for fetal harm among pregnant women. *Pharmacoepidemiol Drug Saf.* 2006; 15(8):546-54.
4. Hardy JR, Leaderer BP, Holford TR, Hall GC, Bracken MB. Safety of medications prescribed before and during early pregnancy in a cohort of 81,975 mothers from the UK General Practice Research Database. *Pharmacoepidemiol Drug Saf.* 2006;15(8):555-64.
5. Toh S, Mitchell AA, Werler MM, Hernandez-Diaz S. Sensitivity and specificity of computerized algorithms to classify gestational periods in the absence of information on date of conception. *Am J Epidemiol.* 2008;167(6):633-40.
6. Joseph KS, Fahey J. Système canadien de surveillance périnatale. Validation des données périnatales de la Base de données sur les congés des patients de l'Institut canadien d'information sur la santé. *Maladies chroniques au Canada.* 2009;29(3):108-113.
7. Wen SW, Liu S, Marcoux S, Fowler D. Utilité et limites des dossiers d'hospitalisation courants pour la surveillance périnatale. *Maladies chroniques au Canada.* 1997;18(3):125-32.
8. Institut canadien d'information sur la santé. Qualité de données : Base de données sur les congés des patients 2002-2003, Ottawa (Ontario): ICIS; 2004.
9. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. *J Clin Epidemiol.* 2001;54(8):774-81.
10. Alexander GR, Himes JH, Kaufman RB, Mor J, Kogan M. A United States national reference for fetal growth. *Obstet Gynecol.* 1996;87(2):163-8.
11. Tukey JW. Exploratory data analysis. Reading (MA): Addison-Wesley; 1977.
12. Wingate MS, Alexander GR, Buekens P, Vahrtian A. Comparison of gestational age classifications: date of last menstrual period vs. clinical estimate. *Ann Epidemiol.* 2007;17(6):425-30.
13. Qin C, Hsia J, Berg CJ. Variation between last-menstrual-period and clinical estimates of gestational age in vital records. *Am J Epidemiol.* 2008;167(6):646-52.
14. Institut canadien d'information sur la santé. Manuel de saisie de l'information, BDCP (pour utilisation avec la CIM-10-CA/CCI) 2006-2007, Ottawa (Ont.): ICIS; 2006.
15. Canadian Institute for Health Information. Final report: the Canadian enhancement of ICD-10 (International Statistical Classification of Diseases and Related Health Problems, Tenth Revision). Ottawa (ON): CIHI; 2001.
16. Sullivan LM, Massaro JM, D'Agostino RB Sr. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. *Stat Med.* 2004;23(10):1631-60.
17. Durrleman S, Simon R. Flexible regression models with cubic splines. *Stat Med.* 1989;8(5):551-61.
18. Hutcheon JA, Zhang X, Cnattingius S, Kramer MS, Platt RW. Customised birthweight percentiles: does adjusting for maternal characteristics matter? *BJOG.* 2008;115(11):1397-404.
19. Bonellie S, Chalmers J, Gray R, Greer I, Jarvis S, Williams C. Centile charts for birthweight for gestational age for Scottish singleton births. *BMC Pregnancy Childbirth.* 2008;8:5.
20. Kramer MS, Platt RW, Wen SW, Joseph KS, Allen A, Abrahamowicz M, et al. A new and improved population-based Canadian reference for birth weight for gestational age. *Pediatrics.* 2001;108(2):E35.
21. Ananth CV, Vintzileos AM, Shen-Schwarz S, Smulian JC, Lai YL. Standards of birth weight in twin gestations stratified by placental chorionicity. *Obstet Gynecol.* 1998;91(6):917-24.
22. Clinianaia SV, Skjaerven R, Magnus P. Birthweight percentiles by gestational age in multiple births. A population-based study of Norwegian twins and triplets. *Acta Obstet Gynecol Scand.* 2000; 79(6):450-8.
23. Gardosi J, Francis A. A customized standard to assess fetal growth in a US population. *Am J Obstet Gynecol.* 2009;201(1):25.e1-7.
24. Gardosi J, Chang A, Kalyan B, Sahota D, Symonds EM. Customised antenatal growth charts. *Lancet.* 1992;339(8788):283-7.
25. Kierans WJ, Joseph KS, Luo ZC, Platt R, Wilkins R, Kramer MS. Does one size fit all? The case for ethnic-specific standards of fetal growth. *BMC Pregnancy Childbirth.* 2008;8:1.
26. Ray JG, Jiang D, Sgro M, Shah R, Singh G, Mamdani MM. Thresholds for small for gestational age among newborns of East Asian and South Asian ancestry. *J Obstet Gynaecol Can.* 2009;31(4):322-30.
27. Joseph KS, Huang L, Liu S, Ananth CV, Allen AC, Sauve R, et al. Reconciling the high rates of preterm and postterm birth in the United States. *Obstet Gynecol.* 2007;109(4):813-22.
28. Platt RW, Abrahamowicz M, Kramer MS, Joseph KS, Mery L, Blondel B, et al. Detecting and eliminating erroneous gestational ages: a normal mixture model. *Stat Med.* 2001;20(23):3491-503.
29. Tentoni S, Astolfi P, De PA, Zonta LA. Birthweight by gestational age in preterm babies according to a Gaussian mixture model. *BJOG.* 2004;111(1):3