

# Mise à jour de l'indice de défavorisation de 2011: défis méthodologiques entourant l'utilisation des données de l'ENM

Denis Hamel et Philippe Gamache

Conférence Politique sociale et Inégalités de la santé

7 mai 2014

[www.inspq.qc.ca](http://www.inspq.qc.ca)

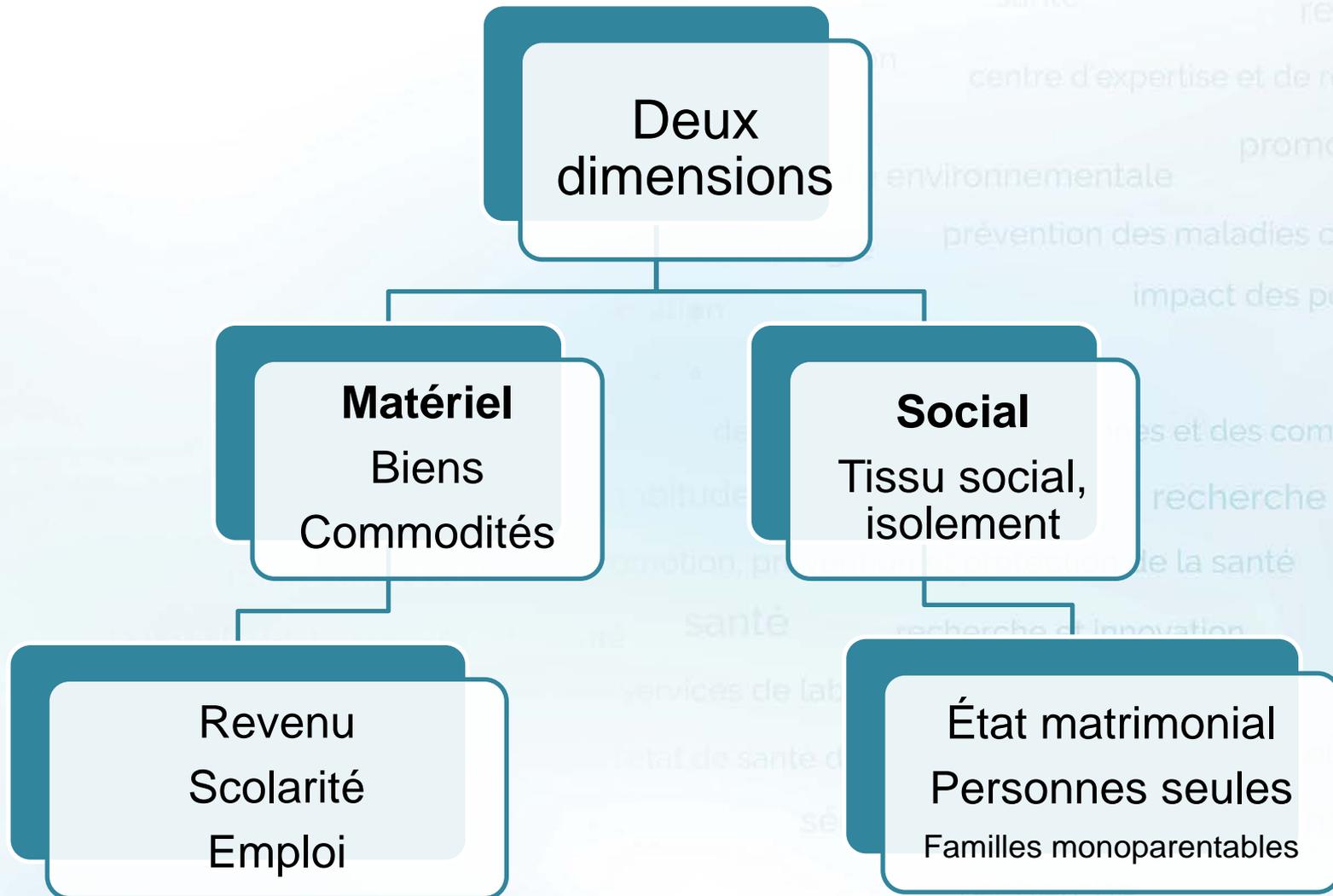
# Contexte



## Trois utilités/avantages de l'indice de défavorisation:

- Ajoute une information socio-économique dans les bases de données administratives;
- Permet de mesurer les inégalités sociales de santé et de suivre ces inégalités dans le temps et dans l'espace;
- Regroupe des déterminants de la santé plutôt que d'utiliser un seul indicateur.

# Contexte: concept défavorisation



# Unité géographique de base

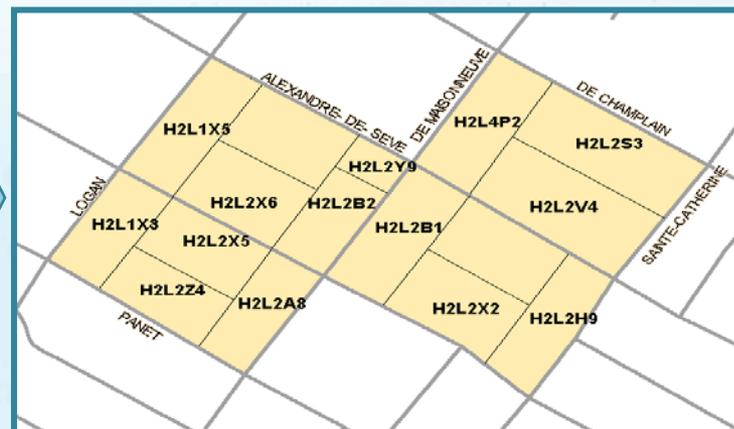
## Aire de diffusion



400 à 800 personnes

Recensement  
Renseignements socioéconomiques

## Codes postaux



- de 100 personnes

Bases de données administratives  
Code postal



# Indicateurs retenus

## Indicateurs ajustés:

- Revenu moyen
  - Taux d'emploi
  - Proportion de personnes sans DES
  - Proportion de personnes seules
  - Proportion de personnes séparées, divorcées ou veuves
- 
- 15 ans et plus

## Indicateur non ajusté:

- Proportion de familles monoparentales

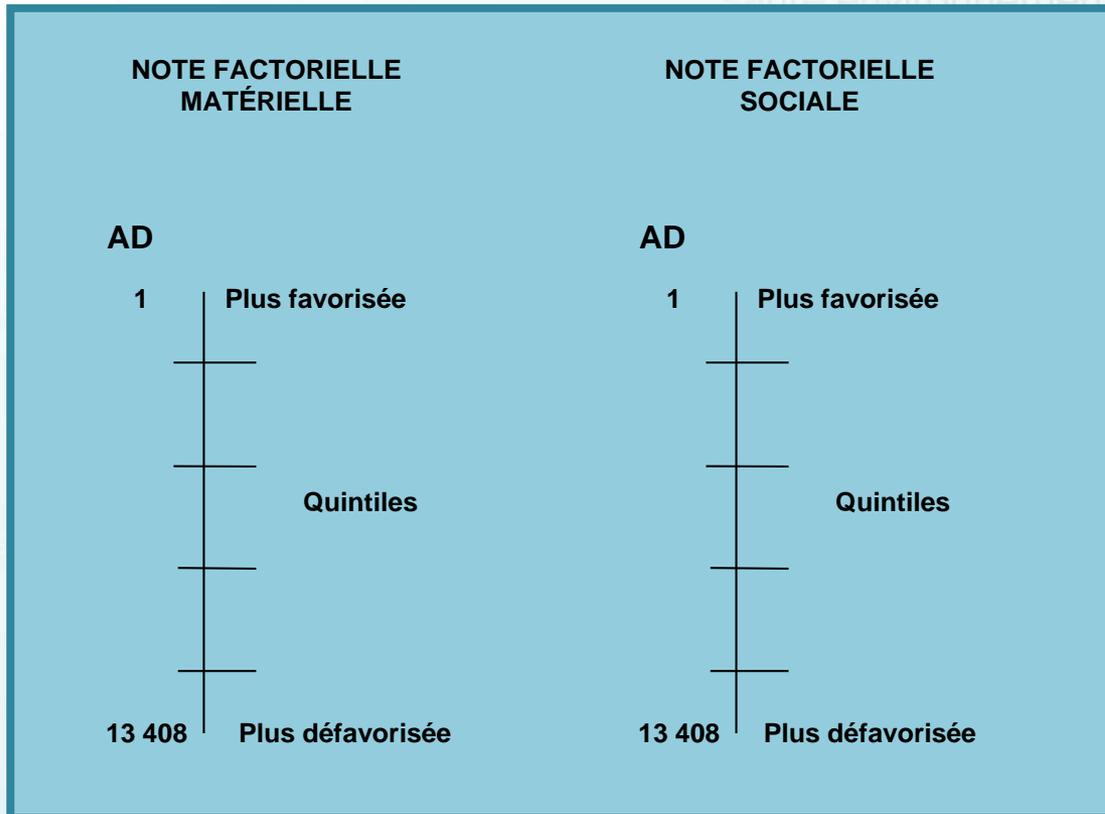
# Intégration des indicateurs

## Analyse en composantes principales:

Indicateur	Composante	
	matérielle	sociale
Sans diplôme d'études secondaires	-0,85	0,04
Rapport emploi/population	0,75	-0,18
Revenu personnel moyen	0,83	-0,28
Personnes vivant seules	-0,12	0,82
Personnes séparées, divorcées ou veuves	-0,12	0,85
Familles monoparentales	-0,21	0,68
<b>Variance expliquée</b>	34%	33%
<b>Variance cumulée</b>	34%	67%

# Classement des aires de diffusion

## Création des quintiles de défavorisation:



# Versions actuellement disponibles

## Recensement:

- 1991, 1996, 2001 et 2006

## Géographiques:

- Québec, Canada, régions canadiennes, rural/urbain, RMR
  - Nouvelles ACP
- Régions, RLS, CLSC
  - Redistribution des notes factorielles

## Demandes spéciales (pondération):

- Municipalités, CSSS, CLSC, quartiers

# De nombreux utilisateurs

- **Québec:** Ministère de la Santé et des services sociaux, Institut de la statistique du Québec (ISQ), Agences régionales de santé, Autres ministères: éducation, solidarité sociale.
- **Fédéral:** Agence de santé publique du Canada, Institut canadien de l'information sur la santé (ICIS)
- Organismes dans les autres provinces canadiennes
  - Alberta Health Services, Cancer Care Nova Scotia, Population Data BC, Saskatoon Health Region, etc.
- Chercheurs universitaires

# Défis posés par l'indice de 2011

## 1. Qualité des données: Taux global de non-réponse

[www.inspq.qc.ca](http://www.inspq.qc.ca)

# Conséquences

## Remplacement du questionnaire long du recensement par l'ENM

- Participation volontaire: taux de réponse non pondéré  
En 2011, Qc: 72%: Canada: 69%  
En 2006, Qc et Canada environ 94%

## Impact possible sur trois indicateurs:

- Revenu moyen
- Taux d'emploi
- Scolarité

# Taux global de non-réponse (TGN)

## TGN:

- Indicateur « imparfait » de la qualité des données
- Combine le taux de non-réponse total (ménage) et de la non-réponse partielle (question)
- TGN élevé → risque élevé de biais (biais potentiel)
- Seuil de suppression de Statistique Canada:  $TNG > 50\%$  alors qu'il était de 25% pour les recensements antérieurs

# TGN > 50%, pour le Québec



## Environ 25% des municipalités:

- Seulement 3% de la population (petites municipalités)

## Environ 7% des aires de diffusion:

- 6,4% de la population
- Problème moins important en zones urbaines:
  - Montréal: 2,9% des AD; Québec: 3,9% des AD

# Comment composer avec les TGN ?

## Le seuil de 50% est arbitraire:

- Deux AD à 50% n'ont probablement pas le même biais

Exemple: Une AD «multiculturelle» de Montréal avec un TGN de 40% est-elle moins biaisée qu'une AD «homogène» de Chandler avec un TGN de 60% ?

# Comment composer avec les TGN ?

## Identifier les biais est très ardu:

- Le profil des répondants vs. les non-répondants est inconnu.
- Un changement de 2006 à 2011 est-il réel ou est-il le résultat d'un biais causé par la non-réponse ?
- L'absence d'un changement de 2006 à 2011 est-elle réelle ou est-ce qu'un véritable changement a été atténué par le biais relié à la non-réponse ?
- Aucune mesure de correction simple applicable à toutes les aires de diffusion

# Du positif dans tout ça?

## Trois points positifs:

- Le nombre d'AD ne respectant pas le seuil de Statistique Canada est relativement faible.
  - Porter attention à certaines AD.
- Environ 95% des territoires d'AD de 2011 sont identiques à ceux de 2006 (lien géographique unique)
  - Ouvre la porte aux comparaisons avec 2006.
- Le caractère relatif de l'indice
  - Création de quintiles par l'intégration de plusieurs indicateurs.

# Défis posés par l'indice de 2011

## 2. Standardisation des indicateurs

[www.inspq.qc.ca](http://www.inspq.qc.ca)

# Traditionnellement...

- Ajustement des proportions et du revenu moyen selon quatre groupes d'âge et le sexe:
  - Méthode: standardisation directe
  - Population de référence: province de Québec ou Canada
  - Nécessaire, car la structure d'âge et de sexe peut varier passablement à l'échelle de l'AD
  - Assure la comparabilité entre les unités géographiques

# Valeurs supprimées



« Dans le cas de toutes les variables quantitatives, une statistique est supprimée si le nombre d'enregistrements réel (non arrondi ni pondéré) à partir desquels ont été calculées est inférieur à 4. »

- Dans les faits, les statistiques problématiques ne sont pas supprimées; une valeur «0» est plutôt affectée.

# Les zéros - Exemple

Total - Sexe				
Total - Groupe ...	15 - 24 ans	25 - 44 ans	45 - 64 ans	65 ans et plus
70	0	0	20	30



Au moins un «faux zéro»

Sexe masculin				
Total - Groupe ...	15 - 24 ans	25 - 44 ans	45 - 64 ans	65 ans et plus
15	0	0	0	0



Au moins un «faux zéro»

Sexe féminin				
Total - Groupe ...	15 - 24 ans	25 - 44 ans	45 - 64 ans	65 ans et plus
55	0	0	0	20



Au moins un «faux zéro»



# Valeurs supprimées

% aire de diffusion sans aucune suppression:

Indicateur	Âge et sexe	Âge seulement
Revenu	72%	81%
Emploi	58%	75%
Scolarité	19%	54%

# Solutions envisagées pour la mise à jour de l'indice de défavorisation 2011

[www.inspq.qc.ca](http://www.inspq.qc.ca)

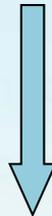
# Qualités recherchées

- Répondre aux défis méthodologiques de l'ENM
  - Ajustement âge-sexe
  - Minimiser les biais « possibles » dus aux TGN élevés
- Comparable aux versions précédentes
  - Appliquer les solutions aux données de 2006 et comparer le nouveau regroupement en quintiles avec la version antérieure de l'indice
- Facilement reproductible à l'échelle des AD au Québec et au Canada

# Classification des AD

## Groupe 1: AD éliminées

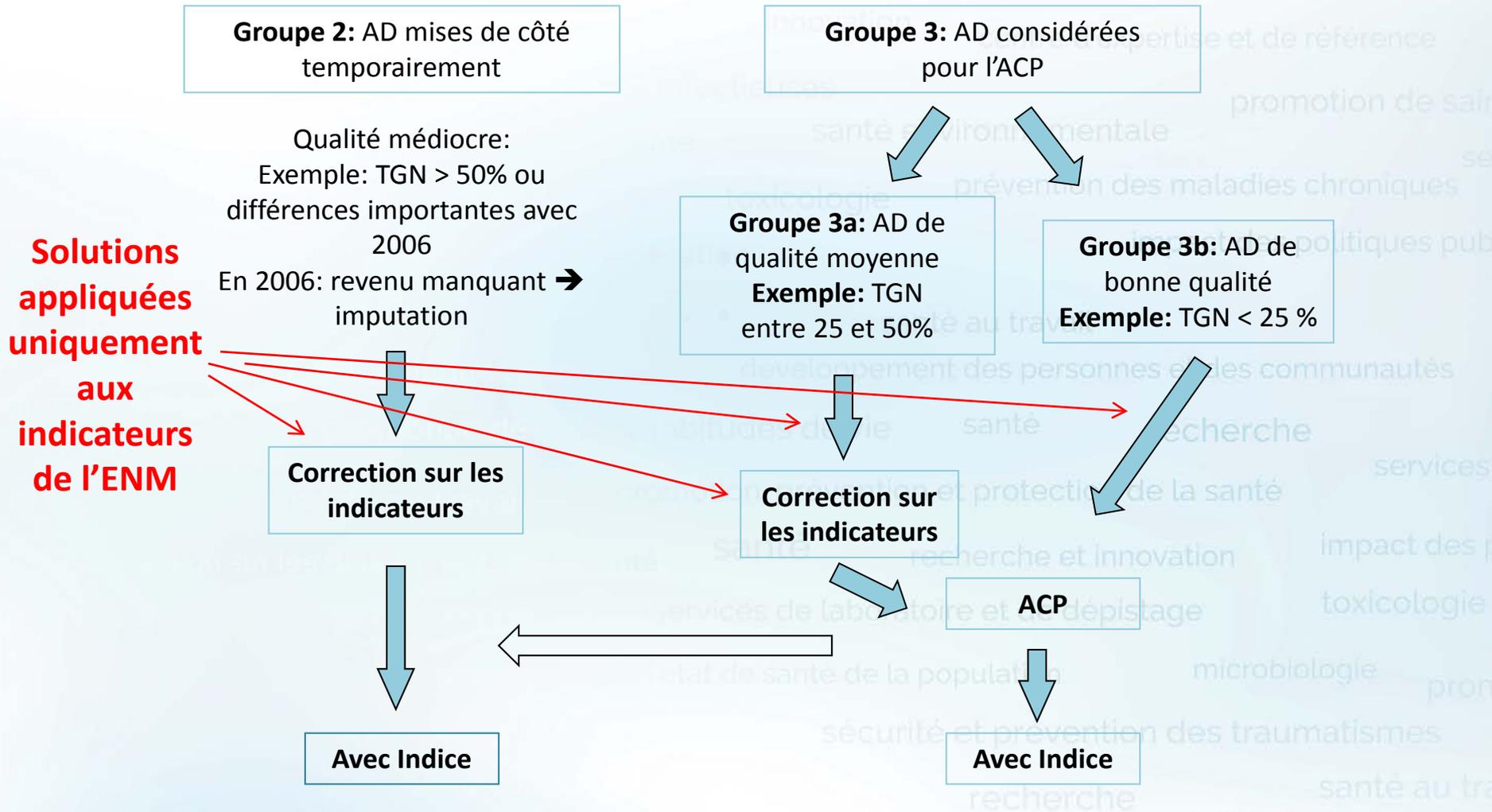
Peu peuplées  
Ménages collectifs  
Sans information SE  
En 2006: 622 AD (3,1% Population)



Sans indice



# Classification des AD



# Solution 1: Standardisation indirecte

- Mesure retenue du type ratio: observé sur attendu (semblable au SMR)
- Appliquée à toutes les AD
- Implique uniquement la connaissance de la structure d'âge de la population vivant en ménages privés de l'AD  
→ recensement
- Inconvénient: lacune connue pour la comparaison directe entre AD.

# Solution 1: Standardisation indirecte

- Possibilité d'appliquer une correction supplémentaire aux groupes d'AD dont le TGN > 25% (groupes 2 et 3a):
  - Améliorer la qualité en utilisant de l'information auxiliaire,  $G$ 
    - Estimations du recensement précédent
    - Estimations d'un territoire plus grand (par exemple un ensemble d'ADs contigus ou d'un secteur de recensement) avec de meilleurs TGN
  - Suggérer un estimateur composite empirique de Bayes de la forme suivante:

$$Est_{AD}^{EB} = B \times Est_{AD} + (1 - B) \times Est_G$$

# Solution 2: Estimations par le quotient: taux ajustés de 2006

- Pour les AD du groupe 3b (TGN < 25%), lorsque possible, calcul des proportions ajustées;
- Sinon et pour les autres AD (groupes 2 et 3a), approximer par les valeurs ajustées de 2006 et des proportions brutes de 2006 et 2011:

$$Prop_{aj,2011} = \alpha \times \frac{Prop_{brut,2011}}{Prop_{brut,2006}} \times Prop_{aj,2006}$$

où  $\alpha$  est un facteur de correction égal à 1 si la distribution âge-sexe pour l'AD est semblable entre 2006 et 2011.

$$\alpha = \frac{Prop_{brut,2006}}{\sum_i w_{i,2011} Prop_{i,2006}}$$

# Solution 2: Estimations par le quotient: taux ajustés de 2006

- Possibilité d'appliquer une correction supplémentaire aux groupes d'AD 2 et 3a:
  - Mais cette fois-ci en utilisant uniquement de l'information auxiliaire d'un territoire plus grand (par exemple un ensemble d'ADs contigus ou d'un secteur de recensement) avec de meilleurs TGN

# Aperçu des méthodes considérées

Type d'ajustement:

Standardisation  
indirecte (ratio)

Estimation par le quotient  
Taux ajustés de 2006

Correction Bayes:

Aucune

Recensement  
2006

Territoires

Aucune

Territoires

Méthode 1

Méthode 2

Méthode 3

Méthode 4

Méthode 5

# Conclusion: suite des travaux

[www.inspq.qc.ca](http://www.inspq.qc.ca)

# ENM: source de problème pour l'indice de défavorisation

- Mise à niveau de l'indice de défavorisation pour 2011 nécessaire
  - Besoins demeurent importants: pour continuer de mesurer les inégalités sociales de santé

## ... MAIS

- Il n'existe pas de solution parfaite pour minimiser les risques de biais « potentiels » liés à la présence des TGN élevés au niveau des AD;
- Différentes méthodes seront évaluées et testées sur les données de 2006 et 2011 pour vérifier leur propension à détecter les inégalités sociales de santé sur quelques indicateurs de santé.

# Quel avenir pour l'indice de défavorisation ?

## D'autres solutions ???

- Produire un indice à une échelle géographique plus élevée: secteurs de recensement ou communautés.
- Maintien de l'abolition du questionnaire long du recensement dans l'avenir
  - Considérer d'autres sources de données
  - Innover: nouvelle vision d'un indice
    - Passer de la défavorisation à la vulnérabilité

Merci de votre attention

Contact: [denis.hamel@inspq.qc.ca](mailto:denis.hamel@inspq.qc.ca)

[www.inspq.qc.ca](http://www.inspq.qc.ca)

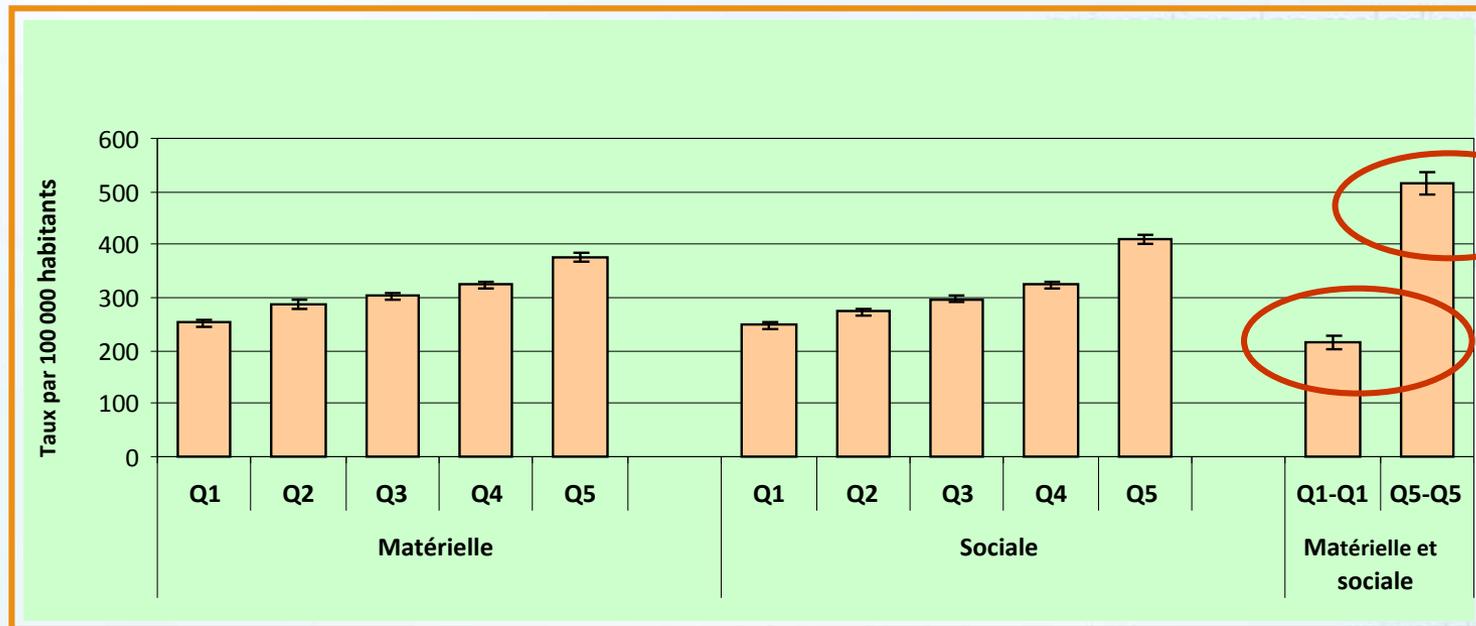
# Ventilation des TGN

TGN	Municipalités		Aire de diffusion	
	Nombre (%)	Population moyenne	Nombre (%)	Population moyenne
0%-25%	233 (20%)	25 749	4724 (35%)	607
25%-40%	448 (39%)	2938	5790 (43%)	599
40%-50%	298 (26%)	1400	1902 (14%)	554
50%-60%	143 (12%)	1034	782 (6%)	544
60%-100%	40 (3%)	530	168 (1%)	500
Inconnu	0	0	256 (2%)	23

# Exemples de mesures d'inégalités sociales de santé

## Mortalité prématurée (avant 75 ans) au Canada, 2001

Taux de mortalité par quintile de défavorisation matérielle et sociale\*



\* Les taux sont ajustés selon l'âge et le sexe, la zone géographique et l'autre forme de défavorisation.