



ENVIRONNEMENT QUÉBEC

---

**Campagne d'échantillonnage  
des  
puits individuels et des petits réseaux d'aqueduc**

**Présentation des résultats**

**Campagne 2002**

---

**DIRECTION DES POLITIQUES DU SECTEUR MUNICIPAL**

*SERVICE DE L'EXPERTISE TECHNIQUE EN EAU*

Mise à jour  
18 février 2003

## CONTEXTE

Le ministère de l'Environnement (MENV) a instauré un programme d'échantillonnage de l'eau captée par des puits dans tout le Québec, au cours de l'été 2001. Environ 33 000 trousse ont été demandées et près de 18 000 échantillons d'eau ont été retournés aux laboratoires lors de cette première campagne. Une deuxième campagne a été lancée vendredi le 12 juillet 2002, et a pris fin le 31 octobre 2002. Cette nouvelle campagne vise, comme l'an dernier, les propriétaires de puits individuels et de réseaux d'aqueduc desservant moins de 21 personnes.

Les propriétaires intéressés ont fait une demande de trousse d'échantillonnage d'eau au Centre d'information du ministère, accompagnée d'un montant de trente dollars (30 \$). Cette somme servait à couvrir les frais de manutention et d'analyses. Par la suite, les propriétaires envoyaient leurs échantillons d'eau pour être analysés par des laboratoires accrédités par le MENV. Les laboratoires accrédités envoient les résultats aux propriétaires ainsi qu'au Centre d'expertise en analyse environnementale (CEAE). Sur réception des analyses, le CEAE a payé les laboratoires concernés.

Chaque trousse contient deux bouteilles et les instructions permettant de faire un prélèvement d'eau de façon adéquate. Les résultats transmis au MENV par le CEAE portent sur trois éléments, le nom de la municipalité dans laquelle fut prélevé l'échantillon, le dénombrement de *E. coli* et la concentration de nitrates. Ainsi, il n'est pas possible de relier les résultats avec le propriétaire concerné.

Le programme a pour objectif de sensibiliser les propriétaires de puits à l'importance de procéder périodiquement<sup>1</sup> à l'analyse de l'eau souterraine qu'ils captent, afin de vérifier la qualité de l'eau qu'ils consomment. Ce programme n'a pas été conçu de manière à pouvoir caractériser, avec rigueur, la qualité de l'eau souterraine des aquifères exploités à des fins d'alimentation en eau potable.

En date du 17 février 2003, 6 266 trousse ont été demandées. Le nombre de résultats compilé pour les *E. coli* s'élève à 4 882 et il est de 4 841 pour les nitrates<sup>2</sup>. La différence de l'ordre de 1 400 entre le nombre de résultats compilés et le nombre de trousse demandées peut provenir de l'interdiction de transmettre les résultats du CEAE au MENV par des propriétaires qui ont demandé des trousse d'échantillonnage.

Mentionnons qu'une campagne de caractérisation de la qualité des eaux souterraines, exploitées à des fins d'alimentation en eau potable, a également été réalisée en 2002. Elle couvre les bassins versants des rivières Chaudière, Etchemin, Boyer, Bayonne, L'Assomption, Yamaska et Nicolet. Ces bassins versants présentent un problème de

---

<sup>1</sup> Normalement, tout propriétaire d'un puits à usage domestique devrait effectuer l'analyse de l'eau deux fois par année.

<sup>2</sup> L'écart entre le nombre de résultats compilé pour les *E. coli* et les nitrates résulte du fait que certains laboratoires ont confié à un second laboratoire l'analyse de l'un ou l'autre de ces paramètres. La transmission des deux résultats au Centre d'expertise en analyse environnementale n'a donc pas été systématique.

surplus de fumier. Des représentants du MAPAQ et du MENV ont prélevé 2 094 échantillons d'eau, qui ont été analysés par leur laboratoire respectif. Les paramètres analysés sont les *E.coli*, les entérocoques, les virus de type coliphage ainsi que les nitrates.

### **RÉSULTATS DE LA CAMPAGNE D'ÉCHANTILLONNAGE (17 FÉVRIER 2003)**

Pour aider la compréhension du texte, un tableau<sup>3</sup> annexé représente les statistiques générales et celles région par région. De même se trouve en annexe 2 figures représentant les *E.coli* et les nitrates-nitrites.

#### *Escherichia coli*

Sur les 4 882 résultats compilés, il y en a 405, soit 8,3 % des analyses en *E. coli*, qui sont hors normes, ou en d'autres termes, ces bactéries sont présentes mais à différentes concentrations.

Toutefois, le portrait région par région est variable. Sur les 17 régions, cinq d'entre-elles ont un pourcentage de résultats hors normes supérieur à 10 %, soit Bas-Saint-Laurent (11,9 %), Outaouais (12,9 %), Abitibi-Témiscamingue (10,9 %), Gaspésie-Îles-de-la-Madeleine (21,7 %), Chaudière-Appalaches (14,0 %). Cinq régions ont des résultats qui sont inférieurs à 5 % soit : Saguenay-Lac-St-Jean, Mauricie, Côte-Nord, Nord-du-Québec, , Laval. Sept régions ont des résultats qui se situent entre 5 % et 10 % : Capitale Nationale, Estrie, Montréal, Lanaudière, Laurentides, Montérégie, Centre-du-Québec.

#### *Nitrates*

Les nitrates dans l'eau souterraine existent à l'état naturel partout au Québec, mais à des concentrations souvent inférieures à la limite de détection (0,02 mg/L). Le milieu naturel n'est toutefois pas homogène, de sorte que la concentration naturelle peut varier d'une région à l'autre et être supérieure à la limite de détection. Au Québec, une concentration naturelle en nitrate est, règle générale, inférieure à 1 ppm. Afin de distinguer les concentrations d'origine naturelle de celles influencées par une activité anthropique, un seuil de 3 ppm est utilisé. Le dépassement d'une telle concentration signifie que des activités humaines affectent nécessairement la qualité de l'eau souterraine. Ce seuil de 3 ppm est mentionné dans le *Règlement sur le captage de l'eau souterraine*, en vigueur depuis le 14 juin 2002. Il faut toutefois noter que la norme pour l'eau potable est de 10 ppm.

Sur les 4 841 échantillons de compilés, il y en a 82 (1,7 %) qui dépassent la norme de l'eau potable et 338 échantillons (7,0 %), dont les concentrations varient entre 3 et 10 ppm. Pour cette raison, il est plus instructif de procéder à l'analyse en termes de pourcentage de résultats dépassant le seuil de 3 mg/L, lequel démontre l'influence d'activités humaines.

---

<sup>3</sup> Le tableau de 3 pages est intitulé Programme d'échantillonnage de l'eau de puits individuels ou de petits réseaux

Cinq régions ont un pourcentage de résultats influencés inférieur à 5 % : Montréal, Abitibi-Témiscamingue, Côte-Nord, Nord-du-Québec, Laval. Six régions se situent entre 5 % et 10 % : Saguenay-Lac-Saint-Jean, Estrie, Outaouais, Laurentides, Montérégie, Centre-du-Québec. Six régions ont un pourcentage supérieur à 10%, Bas-Saint-Laurent, Capitale Nationale, Mauricie, Gaspésie-Îles-de-la-Madeleine, Chaudière-Appalaches, Lanaudière.

### **Comparaison entre la campagne d'échantillonnage de 2001 et celle de 2002**

Toute proportion gardée, le pourcentage général des résultats hors normes pour les *E. coli* et les nitrates-nitrites sont dans le même ordre de grandeur. En 2001, il y avait 7,7 % d'échantillons en *E. coli* hors norme, par rapport à 8,3 % en 2002. En 2001, il y avait 1,4 % d'échantillons hors norme en nitrates-nitrites, par rapport à 1,7 % en 2002. Toutefois on note une contamination ayant une tendance à la hausse.

### **Représentativité statistique des résultats**

*Est-ce que les résultats disponibles sont représentatifs de l'état de la qualité de l'eau souterraine captée par l'ensemble des puits individuels du Québec ?*

L'avis méthodologique produit en date du 4 février 2002, par la *Direction de la méthodologie et des enquêtes spéciales* (DMES), de l'*Institut de la Statistique du Québec*, fournit une réponse détaillée à cette question. Bien que le lecteur soit invité à lire cet avis, joint en annexe, il convient d'en citer ici la conclusion :

*La base de données ne peut être considérée pour réaliser des estimations objectives sur la qualité de l'eau des puits et des petits réseaux d'aqueduc québécois. Le profil des données fournies par les participants au programme risquant de différer sensiblement de celui qu'aurait pu nous fournir les non-participants au programme, il serait même périlleux d'utiliser cette base de données, car ceci pourrait mener à des constats qui ne se vérifieraient pas dans la population.*

*Les compilations qui seront produites à partir de cette base de données devront être accompagnées de sérieuses mises en garde pour éviter une confusion quant à l'interprétation de ces résultats. On pourra tout de même dégager des conclusions d'ordre général quant à la campagne de sensibilisation et quelques résultats approximatifs tirés de la base de données.*

### **Interprétation des résultats**

L'avis méthodologique produit par la DMES de l'*Institut de la Statistique du Québec* est sans équivoque. Pour cette raison, l'interprétation qui suit se limite à une compilation des

résultats, sans tirer de conclusions précises sur les caractéristiques de la population des puits individuels du Québec. Compte tenu qu'il n'y a pas eu de changement au programme, la teneur de cet avis est encore d'actualité.

### *Escherichia coli*

L'information disponible ne permet pas d'attribuer la contamination des puits par les *E. coli* à une source ou une activité humaine particulière. **Il existe néanmoins un pourcentage élevé de puits contenant des bactéries d'origine fécale, donc de gens exposés s'ils consomment l'eau de leur puits.**

Bien qu'il y ait peu d'informations disponibles, peut-on considérer le pourcentage global de puits contaminés par des *E. coli*, observé en date du 17 février 2003, comme un maximum?

Pour répondre à une telle question, il est requis de poser une hypothèse qui, il est important de le souligner, ne peut être vérifiée. Si on pose l'hypothèse que plusieurs participants au programme d'échantillonnage étaient motivés par des doutes sur la qualité de l'eau consommée, donc qu'une majorité des résultats obtenus découle de l'analyse de puits faisant partie de l'ensemble des puits individuels du Québec qui présente un risque de contamination, alors il est permis d'avancer que globalement le pourcentage de puits contaminés par les bactéries *E. coli* peut s'apparenter à un maximum. C'est-à-dire que ce pourcentage pourra diminuer avec l'augmentation du nombre de puits échantillonnés.

### *Nitrates*

Très peu de résultats sont hors normes, mais **les données disponibles démontrent qu'une influence significative est exercée par les activités humaines en surface, c'est-à-dire qu'elles contaminent l'eau souterraine en nitrates**. Tout comme pour les *E. coli*, il n'est pas possible d'identifier, dans chacun des cas, une activité humaine particulière pour expliquer cette augmentation des nitrates dans les eaux souterraines, en raison de l'absence d'information sur les caractéristiques des puits et leur environnement.

Comme dans le cas des *E. coli*, il est requis de poser des hypothèses, impossibles à vérifier, pour tenter de dégager une tendance quant aux résultats de la campagne d'échantillonnage. Pour la région administrative de Montréal, le nombre de demandes de trousse (496) a été très forte considérant les 600 puits potentiels de la région; le taux de participation est de 82,7%! Dans la région de la Capitale Nationale le taux de participation a été de 4,8 %, à Laval de 8,1 %, et la moyenne pour les autres régions est de 2,4 %. Les taux de participation, en milieu fortement urbanisé ayant des réseaux d'aqueducs, suggèrent l'hypothèse qu'un nombre significatif des participants a demandé des trousse afin de vérifier la qualité de l'eau du puits qui alimente leur résidence secondaire, tel un chalet. À cette hypothèse peut s'en ajouter une seconde, à l'effet que la plupart de ces résidences secondaires ne sont pas situées en zone agricole (les activités agricoles sont les plus susceptibles de générer une contamination des eaux souterraines par les nitrates). En posant ces deux hypothèses, il est permis de croire que globalement

le pourcentage de puits contaminés par les nitrates est susceptible d'augmenter avec l'ajout de nouveaux résultats.

### **Lien entre les nitrates et les *E. coli***

Les résultats transmis par le Centre d'expertise en analyse environnementale du Québec indiquent que globalement, il y a 8,3 % des puits qui sont contaminés par des *E. coli* et 8,7 % par des nitrates.

Il n'existe pas de lien apparent entre les fortes concentrations en nitrates (supérieures à 3 mg/L) et la détection de contamination au *E. coli* dans les résultats reçus. Seulement 0,9 % des puits présentent une « double » contamination ; 7,45 % des puits dépassent 3 mg/L de nitrates mais n'ont pas de *E. coli*, alors que 7,15 % révèlent une détection de *E. coli*, mais peu de nitrates. Cette constatation apporte un argument supplémentaire en faveur de la difficulté d'identifier une source de contamination. Les sources possibles sont sans doute de nature multiple et probablement, à la fois, d'envergure locale (ex : système de traitement d'eaux usées domestique) et régionale (ex : fertilisation de terres agricoles).

**PROGRAMME D'ÉCHANTILLONNAGE**  
DE L'EAU DE PUIITS INDIVIDUELS OU DE PETITS RÉSEAUX

<b>Statistiques générales</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels	246 790	---
demande de trousse au 31.10.2002 (% vs puits potentiels)	6 266	2.5%
résultats compilés pour les E. Coli	4 882	77.9%
résultats "hors norme" pour les E. Coli (% vs résultats compilés)	405	8.3%
résultats compilés pour les nitrates & nitrites	4 841	77.3%
résultats "hors norme" pour les nitrates & nitrites (% vs résultats compilés)	82	1.7%
résultats (>3 et < 10) pour les nitrates & nitrites (% vs résultats compilés)	338	7.0%

<b>Statistiques - Bas-Saint-Laurent (01)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	17 900	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	199	1.1%
résultats compilés pour les E. Coli	194	97.5%
résultats "hors norme" pour les E. Coli	23	11.9%
résultats compilés pour les nitrates & nitrites	194	97.5%
résultats "hors norme" pour les nitrates & nitrites	4	2.1%
résultats (>3 et < 10) pour les nitrates & nitrites	17	8.8%

<b>Statistiques - Saguenay-Lac-Saint-Jean (02)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	6 400	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	222	3.5%
résultats compilés pour les E. Coli	185	83.3%
résultats "hors norme" pour les E. Coli	5	2.7%
résultats compilés pour les nitrates & nitrites	184	82.9%
résultats "hors norme" pour les nitrates & nitrites	1	0.5%
résultats (>3 et < 10) pour les nitrates & nitrites	9	4.9%

<b>Statistiques - Capitale Nationale (03)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	14 700	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	706	4.8%
résultats compilés pour les E. Coli	470	66.6%
résultats "hors norme" pour les E. Coli	41	8.7%
résultats compilés pour les nitrates & nitrites	466	66.0%
résultats "hors norme" pour les nitrates & nitrites	21	4.5%
résultats (>3 et < 10) pour les nitrates & nitrites	61	13.1%

<b>Statistiques - Mauricie (04)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	4 700	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	170	3.6%
résultats compilés pour les E. Coli	130	76.5%
résultats "hors norme" pour les E. Coli	4	3.1%
résultats compilés pour les nitrates & nitrites	129	75.9%
résultats "hors norme" pour les nitrates & nitrites	7	5.4%
résultats (>3 et < 10) pour les nitrates & nitrites	9	7.0%

<b>Statistiques - Estrie (05)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	20 900	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	577	2.8%
résultats compilés pour les E. Coli	542	93.9%
résultats "hors norme" pour les E. Coli	41	7.6%
résultats compilés pour les nitrates & nitrites	534	92.5%
résultats "hors norme" pour les nitrates & nitrites	5	0.9%
résultats (>3 et < 10) pour les nitrates & nitrites	23	4.3%

<b>Statistiques - Montréal (06)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	600	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	496	82.7%
résultats compilés pour les E. Coli	15	3.0%
résultats "hors norme" pour les E. Coli	1	6.7%
résultats compilés pour les nitrates & nitrites	15	3.0%
résultats "hors norme" pour les nitrates & nitrites	0	0.0%
résultats (>3 et < 10) pour les nitrates & nitrites	0	0.0%

<b>Statistiques - Outaouais (07)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	20 600	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	294	1.4%
résultats compilés pour les E. Coli	248	84.4%
résultats "hors norme" pour les E. Coli	32	12.9%
résultats compilés pour les nitrates & nitrites	248	84.4%
résultats "hors norme" pour les nitrates & nitrites	4	1.6%
résultats (>3 et < 10) pour les nitrates & nitrites	16	6.5%

<b>Statistiques - Abitibi-Témiscamingue (08)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	15 000	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	184	1.2%
résultats compilés pour les E. Coli	138	75.0%
résultats "hors norme" pour les E. Coli	15	10.9%
résultats compilés pour les nitrates & nitrites	137	74.5%
résultats "hors norme" pour les nitrates & nitrites	0	0.0%
résultats (>3 et < 10) pour les nitrates & nitrites	3	2.2%

<b>Statistiques - Côte-Nord (09)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	1 700	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	53	3.1%
résultats compilés pour les E. Coli	49	92.5%
résultats "hors norme" pour les E. Coli	2	4.1%
résultats compilés pour les nitrates & nitrites	49	92.5%
résultats "hors norme" pour les nitrates & nitrites	0	0.0%
résultats (>3 et < 10) pour les nitrates & nitrites	1	2.0%

<b>Statistiques - Nord-du-Québec (10)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	290	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	11	3.8%
résultats compilés pour les E. Coli	7	63.6%
résultats "hors norme" pour les E. Coli	0	0.0%
résultats compilés pour les nitrates & nitrites	7	63.6%
résultats "hors norme" pour les nitrates & nitrites	0	0.0%
résultats (>3 et < 10) pour les nitrates & nitrites	0	0.0%

<b>Statistiques - Gaspésie-Iles-de-la-Madeleine (11)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	5 200	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	87	1.7%
résultats compilés pour les E. Coli	60	69.0%
résultats "hors norme" pour les E. Coli	13	21.7%
résultats compilés pour les nitrates & nitrites	60	69.0%
résultats "hors norme" pour les nitrates & nitrites	0	0.0%
résultats (>3 et < 10) pour les nitrates & nitrites	7	11.7%

<b>Statistiques - Chaudière-Appalaches (12)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	25 100	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	721	2.9%
résultats compilés pour les E. Coli	694	96.3%
résultats "hors norme" pour les E. Coli	97	14.0%
résultats compilés pour les nitrates & nitrites	672	93.2%
résultats "hors norme" pour les nitrates & nitrites	13	1.9%
résultats (>3 et < 10) pour les nitrates & nitrites	58	8.6%

<b>Statistiques - Laval (13)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	1 000	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	81	8.1%
résultats compilés pour les E. Coli	3	3.7%
résultats "hors norme" pour les E. Coli	0	0.0%
résultats compilés pour les nitrates & nitrites	3	3.7%
résultats "hors norme" pour les nitrates & nitrites	0	0.0%
résultats (>3 et < 10) pour les nitrates & nitrites	0	0.0%

<b>Statistiques - Lanaudière (14)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	21 600	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	520	2.4%
résultats compilés pour les E. Coli	536	103.1%
résultats "hors norme" pour les E. Coli	40	7.5%
résultats compilés pour les nitrates & nitrites	535	102.9%
résultats "hors norme" pour les nitrates & nitrites	16	3.0%
résultats (>3 et < 10) pour les nitrates & nitrites	46	8.6%



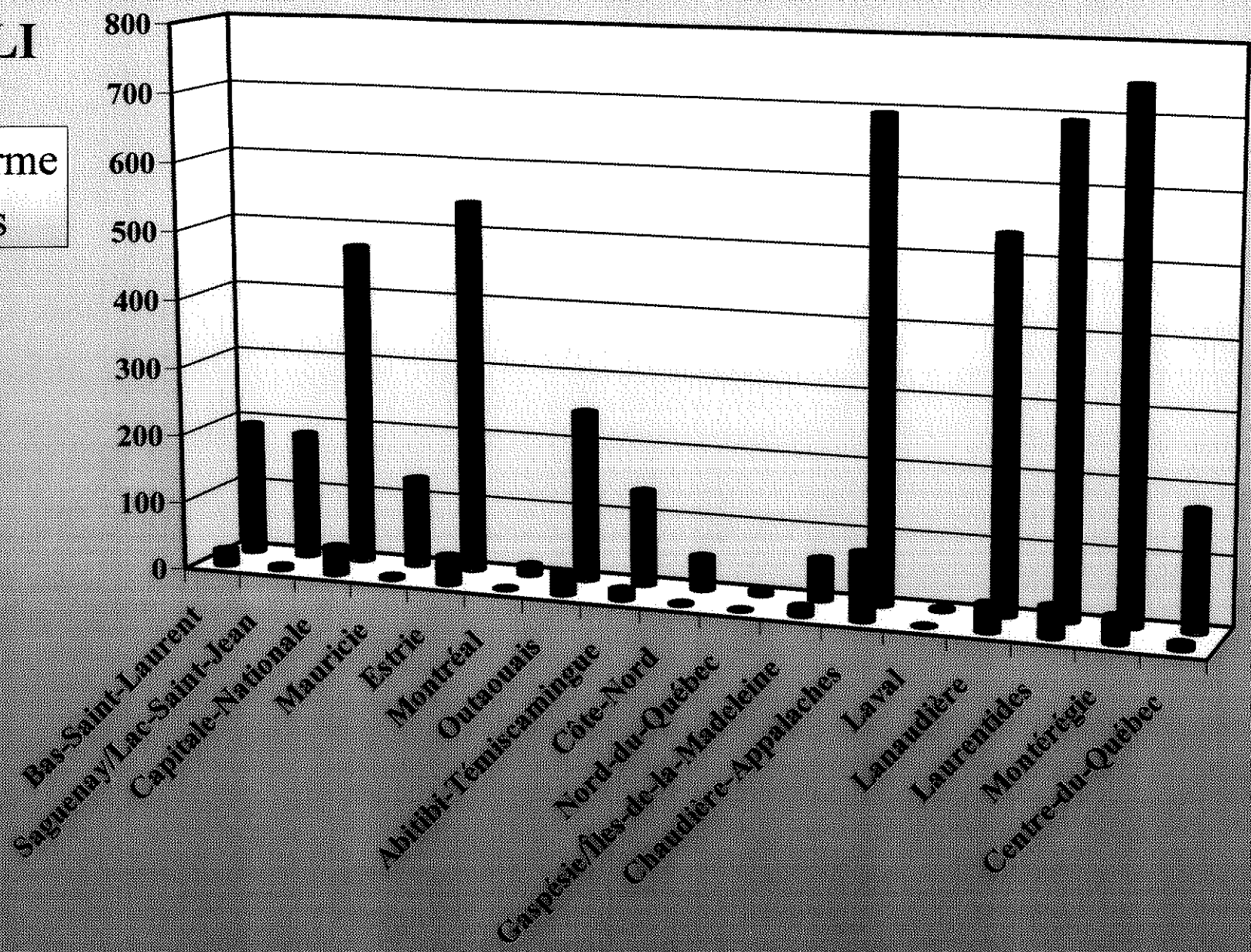
<b>Statistiques - Laurentides (15)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	23 100	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	709	3.1%
résultats compilés pour les E. Coli	693	97.7%
résultats "hors norme" pour les E. Coli	44	6.3%
résultats compilés pour les nitrates & nitrites	691	97.5%
résultats "hors norme" pour les nitrates & nitrites	5	0.7%
résultats (>3 et < 10) pour les nitrates & nitrites	42	6.1%

<b>Statistiques - Montérégie (16)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	55 000	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	1 049	1.9%
résultats compilés pour les E. Coli	746	71.1%
résultats "hors norme" pour les E. Coli	38	5.1%
résultats compilés pour les nitrates & nitrites	745	71.0%
résultats "hors norme" pour les nitrates & nitrites	4	0.5%
résultats (>3 et < 10) pour les nitrates & nitrites	35	4.7%

<b>Statistiques - Centre-du-Québec (17)</b>	<b>nombre</b>	<b>%</b>
nombre de puits potentiels de la région	13 000	---
demande de trousse pour cette région au 31.10.2002 (% vs puits potentiels)	176	1.4%
résultats compilés pour les E. Coli	172	97.7%
résultats "hors norme" pour les E. Coli	9	5.2%
résultats compilés pour les nitrates & nitrites	172	97.7%
résultats "hors norme" pour les nitrates & nitrites	2	1.2%
résultats (>3 et < 10) pour les nitrates & nitrites	11	6.4%

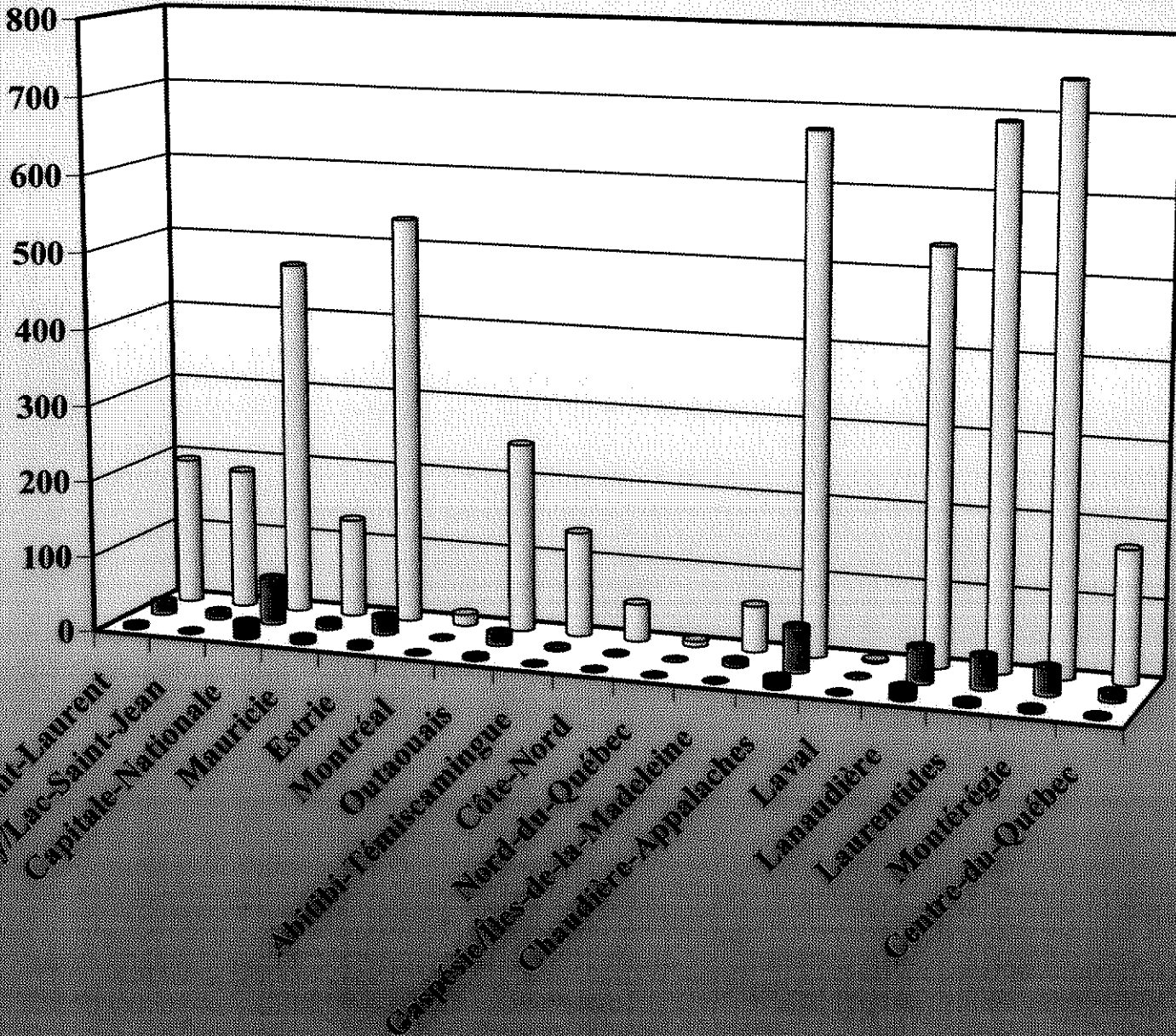
# E. COLI

■ hors norme  
■ analysés



FÉVRIER 2003

# NITRATES & NITRITES



FÉVRIER 2003

**ANNEXE**

**Campagne d'échantillonnage des puits individuels et des petits réseaux d'aqueduc**

**Avis méthodologique**

**Préparé par**

**DIRECTION DE LA MÉTHODOLOGIE ET DES ENQUÊTES SPÉCIALES  
INSTITUT DE LA STATISTIQUE DU QUÉBEC**

**4 FÉVRIER 2002**

## AVIS MÉTHODOLOGIQUE

### CAMPAGNE D'ÉCHANTILLONNAGE DES PUIITS INDIVIDUELS ET DES PETITS RÉSEAUX D'AQUEDUC

#### Section 1 : Contexte

L'Institut de la statistique du Québec (ISQ) a été sollicité par le Service de l'expertise en eau du ministère de l'Environnement du Québec (MENV) en vue d'obtenir un avis méthodologique concernant la représentativité statistique de la base de données constituée en 2001 lors de la « Campagne d'échantillonnage des puits individuels et des petits réseaux d'aqueduc ». De plus, ce même Service désirait que l'ISQ précise les possibilités et limites en vue de l'utilisation de cette base de données. Enfin, quelques recommandations ont été émises en vue d'une campagne similaire qui pourrait être menée ultérieurement.

Le mandat de l'ISQ consistait à répondre aux questions suivantes :

Est-ce que l'échantillon de données constituant cette base est représentatif de l'ensemble de la population, à savoir l'ensemble des puits individuels et des petits réseaux d'aqueduc au Québec? Par exemple, le calcul de concentrations moyennes à l'aide de cette base de données serait-il applicable à l'ensemble des puits utilisés actuellement au Québec?

Quelles sont les possibilités et les limites en ce qui concerne l'utilisation de cette base de données?

Qu'est-ce qui devrait être amélioré lors d'une campagne similaire ultérieure afin d'augmenter le potentiel d'utilisation de cette base de données?

#### Section 2 : La représentativité statistique

##### 2.1 Description de la base de données du MENV

La base de données ayant été créée à la suite du programme d'échantillonnage mis sur pied par le ministère de l'Environnement au cours de l'été 2001, il apparaît pertinent de s'attarder au mode de fonctionnement de ce programme de manière à mieux comprendre la provenance et les caractéristiques des données qu'elle renferme.

Le programme visait tous les propriétaires de puits individuels et de réseaux d'aqueduc desservant moins de 21 personnes situés au Québec et avait pour objectif de sensibiliser les propriétaires de puits à l'importance de procéder périodiquement à l'analyse de l'eau souterraine qu'ils captent, afin de vérifier la qualité de l'eau qu'ils consomment. Les analystes du MENV mentionnent dans un document interne que ce programme n'a pas

été conçu de manière à pouvoir caractériser, avec rigueur, la qualité de l'eau souterraine des aquifères exploités à des fins d'alimentation en eau potable. De fait, une campagne de caractérisation rigoureuse de la qualité des eaux souterraines, exploitées à des fins d'alimentation en eau potable, est en cours de préparation et sera réalisée au cours de l'année 2002 pour les bassins versants des rivières Chaudière, Etchemin, Boyer, Bayonne, L'Assomption, Yamaska et Nicolet.

Concrètement, les propriétaires intéressés pouvaient s'adresser au ministère et faire la demande d'une trousse contenant deux bouteilles et les instructions permettant de faire le prélèvement de façon adéquate. Ces échantillons étaient ensuite envoyés à un laboratoire accrédité pour y faire l'analyse, les coûts étant partagés entre le ministère et le propriétaire du puits. Les analyses touchaient deux aspects de la qualité de l'eau, soit le dénombrement de la bactérie E. coli et la concentration de nitrates. Un espace prévu sur le formulaire permettait au propriétaire d'autoriser le ministère de l'Environnement à utiliser les résultats d'analyse de l'eau de son puits à des fins statistiques. Les données transmises au MENV ne comportaient que trois éléments, soit le nom de la municipalité dans laquelle fut prélevé l'échantillon, le dénombrement de E. coli et la concentration de nitrates. La localisation précise du point de prélèvement n'est donc pas possible avec les données présentes dans la base.

En date du 17 octobre 2001, 32 033 troussees ont été demandées. De ce nombre, on estime à 12 000 le nombre de troussees envoyées à un laboratoire pour l'analyse. Le nombre de résultats compilés pour les E. coli s'élève à 7 806, pour les nitrates, il est de 7 025. Environ 80 % des participants ont consenti à ce que leurs résultats soient transmis au MENV pour faire partie de la base de données.

Puisque le MENV souhaiterait éventuellement utiliser sa base de données pour décrire la situation de la qualité de l'eau des puits au Québec, on doit considérer, d'un point de vue statistique, que les informations contenues dans cette base constituent un échantillon censé représenter la population visée. La section 2.3 tente de définir la population que l'on aimerait idéalement voir représenter par l'échantillon.

## **2.2 Qu'est-ce que la représentativité statistique?**

Le problème de la mesure de la représentativité d'un échantillon n'est pas chose facile! Mais avant de penser à mesurer une telle chose, encore faut-il pouvoir la définir. Qu'entend-on au juste par représentativité statistique? Les quelques lignes qui suivent tentent de cerner cette notion.

Plutôt que de considérer au départ un échantillon et de se demander par la suite à quelle(s) population(s) il pourrait bien appartenir, la démarche préconisée par les statisticiens pour s'assurer de la représentativité d'un échantillon se présente en fait dans l'ordre inverse! Ainsi, au départ, les unités de la population doivent être convenablement énumérées, le plus souvent dans un fichier appelé « base de sondage ». Par la suite, un protocole d'échantillonnage au hasard des unités est appliqué et l'échantillon est extrait de

la population. Puisque les probabilités de sélection<sup>1</sup> des unités sont définies dans le protocole d'échantillonnage, l'échantillon obtenu est qualifié de « probabiliste »<sup>2</sup>. Dans tous les autres cas où les probabilités de sélection sont mal connues, voire inconnues, on parle plutôt d'échantillon non probabiliste. Des formules mathématiques sont ensuite établies pour permettre le calcul des estimations obtenues avec l'échantillon. Ces estimations sont le plus souvent accompagnées d'une cote indiquant leur degré de fiabilité (marge d'erreur).

Règle générale, les statisticiens, tout comme la majorité de leurs clients d'ailleurs, préfèrent appliquer des procédures impliquant des estimateurs non biaisés. Ces estimateurs ont comme propriété qu'en moyenne, les estimations que l'on pourrait théoriquement obtenir en employant le même estimateur sur d'autres échantillons semblables prélevés dans la même population soient centrées autour de la vraie valeur que l'on cherche à estimer, cette dernière valeur pouvant être connue si on disposait des ressources requises pour effectuer un recensement complet de toutes les unités de la population. Les estimateurs obtenus avec ce type d'échantillon sont qualifiés de « non biaisés ».

En plus d'être non biaisé, l'autre grande qualité recherchée pour un estimateur est qu'il possède une grande probabilité de se trouver à proximité de la vraie valeur que l'on tente d'estimer. En appliquant des procédures d'échantillonnage probabilistes telles que celles décrites plus haut, il est possible d'obtenir, avec les données de l'échantillon, une mesure moyenne reflétant cette proximité, qui est appelée variance d'échantillonnage. Plus la variance d'échantillonnage est petite, plus les estimations sont qualifiées de précises.

Ceci étant dit, une définition de la représentativité statistique d'un échantillon pourrait être la suivante :

**« Un échantillon est dit statistiquement représentatif d'une population s'il est extrait de celle-ci par un tirage probabiliste des unités et qu'il offre la possibilité de recourir à des estimateurs non biaisés dont la variance d'échantillonnage est suffisamment petite. »**

Cette définition étant établie, la prochaine section vise à présenter les caractéristiques de la base de données du MENV.

### **2.3 Représentativité statistique de la base de données**

Les éléments clés pour statuer sur la base de données étant maintenant en place, cette section s'attaque au vif du sujet, à savoir la capacité de la base de données à jouer le rôle d'échantillon représentatif de la population. Si on retourne aux objectifs de la campagne d'échantillonnage, la population à l'étude pourrait être l'ensemble des prises d'eau

- 
1. La probabilité de sélection est définie pour chaque unité de la population et correspond à la chance qu'a cette unité d'être sélectionnée pour faire partie de l'échantillon.
  2. Plus répandue, l'appellation « sondage scientifique » fait essentiellement référence au même concept.

connectées soit à un puits individuel ou à un réseau d'aqueduc desservant moins de 21 personnes en opération en 2001 au Québec.

Un premier constat s'impose, et il est de taille : les informations de la base de données ne constituent pas un échantillon probabiliste. En effet, les propriétaires ayant profité du programme ont agi sur une base volontaire, et non parce qu'ils avaient été sélectionnés pour faire partie d'un échantillon. Qui plus est, le consentement donné par le propriétaire pour faire partie de la base de donnée est une autre étape qui s'est faite aussi sur une base purement volontaire. Nul ne peut donc dire avec précision quelles étaient les probabilités de sélection des propriétaires qui se retrouvent actuellement dans la base de données, puisqu'il n'y a pas eu de sélection proprement dite. En fait, puisque l'on est en présence d'un échantillon de volontaires, on peut dire que ce n'est pas le MENV qui a sélectionné l'échantillon, mais plutôt les propriétaires qui se sont sélectionnés eux-mêmes pour faire partie de l'échantillon en participant au programme et en donnant leur consentement.

Autrement dit, le mécanisme ayant permis d'obtenir l'échantillon ne garantit pas à lui seul la représentativité de celui-ci. Il s'agit là d'une faille majeure, qui fait en sorte que les procédures habituelles d'estimation ne peuvent être utilisées sans faire des hypothèses sur les données présentes dans la base. Par exemple, aurait-on l'idée de juger de l'état de santé des québécois par l'utilisation de tests sanguins effectués lors de collectes de sang? On pourrait croire que l'état de santé des donneurs de sang est sensiblement meilleur que dans la population en général, d'où un biais si on utilisait ces données telles quelles. Dans cet exemple, tout comme dans le cas qui nous intéresse, il s'agit d'échantillons de volontaires qui se sélectionnent par eux-mêmes, ils peuvent donc avoir des caractéristiques différentes de la population en général.

Dans le cas des puits et des petits réseaux d'aqueduc, plusieurs raisons ont pu conduire un propriétaire à participer au programme, et ces raisons peuvent avoir un lien avec la qualité de l'eau qu'il consomme. Par exemple, quelqu'un ayant été indisposé dernièrement par un malaise qu'il croit relié à la qualité de l'eau sera probablement plus tenté qu'un autre à participer au programme. Ou encore, un propriétaire n'ayant pas fait analyser son eau depuis plusieurs années sera possiblement davantage curieux d'en connaître la qualité par un test de laboratoire.

Dans certains cas, on peut envisager l'utilisation d'un échantillon non probabiliste. Par exemple, si la base de données contenait plusieurs variables, et que certaines d'entre elles étaient possiblement en lien avec la qualité de l'eau, des analyses pourraient être effectuées en tenant compte de ces variables. Malheureusement, la base de données ne contenant que la municipalité où se situe le point d'eau échantillonné, ce genre d'analyse ne peut être envisagé.



## **Section 3 : Les limites à l'utilisation de la base de données**

### **3.1 Pas d'inférence à l'ensemble des puits de la population**

Lorsque des estimations seront produites (proportion, moyenne, etc.), il ne sera pas possible d'inférer ces résultats à l'ensemble de la population, i.e. que l'estimation ne représentera pas la vraie valeur dans la population de puits individuels et des petits réseaux d'aqueduc desservant moins de 21 personnes. Pour que cela ait été possible, il aurait été nécessaire que l'échantillon ait été sélectionné de façon à être représentatif de la population. Et il a été établi à la section 2 que ce n'est pas le cas. Ainsi, les estimations sont davantage des compilations puisqu'elles ne reflètent que l'état de la base de données.

### **3.2 Biais des estimations**

Si l'on regarde la façon dont la base de données a été constituée, on constate qu'il y a un grand risque de biais des estimations calculées à partir de cette base. En effet, il y a certains doutes raisonnables de penser que certains propriétaires qui ont fait une demande de trousse ne font pas partie de la population à l'étude ou sont plus enclins à faire une demande de trousse. Par exemple :

- a) Est-ce que les gens qui font une demande de trousse ont des caractéristiques particulières? Par exemple : on peut penser que des personnes qui ont des soupçons sur la qualité de l'eau consommée ou de l'existence d'un risque de contamination.
- b) Est-ce que ceux qui font parvenir la (ou les) bouteille(s) d'eau au laboratoire diffèrent de ceux qui ne l'envoient pas? Par exemple : Ont-ils des raisons très fortes de croire qu'il peut y avoir un risque de contamination?
- c) Est-ce que ceux qui accordent la permission au MENV d'utiliser les résultats présentent des caractéristiques différentes de ceux qui refusent d'accorder cette permission?

La base de données représente donc un sous-ensemble de l'ensemble des puits individuels et des petits réseaux d'aqueduc desservant moins de 21 personnes mais on ne peut étudier les caractéristiques de ceux qui ne s'y trouvent pas versus ceux qui s'y trouvent car on ne dispose pas d'information suffisante pour cela. De même, on ne peut pas évaluer l'incidence des puits qui font partie de la base de données alors qu'ils ne devraient pas s'y retrouver (par exemple : les personnes qui auraient fait analyser l'eau du robinet).

### **3.3 Validation**

Une étape importante avant de procéder à l'échantillonnage est la validation de la base de sondage (à partir de laquelle un échantillon sera sélectionné). Une première vérification est de s'assurer que la base ne contient qu'un enregistrement par unité. Dans le cas présent, il n'existe pas de base de sondage mais il aurait été important d'effectuer cette première vérification auprès de la base de données. Si, par exemple, un propriétaire avait demandé cinq fois durant l'année une trousse, il pourrait se retrouver jusqu'à cinq fois sur

la base de données. Cela pourrait fausser les estimations qui sont produites à partir de la base de données. Par exemple : si la concentration de nitrates d'un certain nombre de puits est très élevée et que ces puits se retrouvent plusieurs fois sur la base de données, la proportion de puits hors normes (pour ce qui est de la concentration de nitrates) pourrait être faussée.

Dans le cas présent, la base de données ne permet pas de détecter les puits qui se retrouvent plusieurs fois, puisqu'on ne retrouve pas de variable qui identifie de façon unique chaque puits de la base de données.

De plus, d'autres vérifications auraient normalement été effectuées, dans le cadre d'une enquête statistique. On aurait, par exemple, effectué un meilleur contrôle sur l'échantillonnage des bouteilles d'eau en envoyant un technicien sur place. Cela aurait permis de s'assurer que les échantillons d'eau étaient toujours recueillis de la même façon. De plus, on aurait peut-être effectué un meilleur contrôle (tel un rappel téléphonique) auprès des propriétaires de puits dont l'échantillon d'eau aurait été recueilli dans une région où le nombre de puits est faible. On aurait pu ainsi détecter les échantillons d'eau qui ne provenaient pas de la population visée.

En bref, des validations nécessaires des données n'ont pu être effectuées, ce qui fait que leur fiabilité ne peut être assurée.

### **3.4 Tests statistiques**

Lorsque l'on produit plusieurs estimations, il peut être intéressant de comparer certaines d'entre elles. Par exemple : on peut vouloir comparer la proportion de puits hors normes par rapport à la quantité de nitrates/nitrites pour deux régions, et faire un test statistique pour détecter une différence significative entre ces deux proportions. Pour effectuer un tel test, on doit pouvoir estimer l'erreur due au fait qu'un échantillon a été sélectionné (et non toute la population). Cependant, cela n'est pas possible dans le cas présent puisqu'on n'a pas sélectionné un échantillon aléatoire de puits dans la population. Donc lorsque l'on fera des comparaisons de proportions, on ne pourra affirmer qu'une proportion diffère significativement d'une autre. Ainsi, la base de données ne permet pas d'effectuer des tests statistiques, tels des comparaisons de proportions.

## **Section 4 : Les possibilités et les recommandations**

### **4.1 Les possibilités d'utilisation de la base de données**

Tout en tenant compte des limites énumérées à la section précédente, il est possible d'utiliser la base de données de différentes façons.

#### ➤ Évaluation du programme

Cette base de données peut fournir des informations utiles pour une campagne similaire qui pourrait être menée ultérieurement. Par exemple : on peut calculer le taux de demande de trousse par région et s'en servir comme une indication (pour les régions dont le taux est plus bas) des régions qui nécessiteraient un plus grand effort de sensibilisation. Le taux de transmission des trousseaux aux laboratoires d'analyse est également une information utile pour une campagne ultérieure. On pourra ainsi tenter d'évaluer si l'objectif de sensibilisation auprès des propriétaires de puits a été atteint.

#### ➤ La portée des estimations

Lorsque des estimations (comme des proportions) seront estimées à partir de cette base de données, il faudra bien en préciser la portée. On pourra spécifier que les estimations sont faites à partir d'un sous-ensemble de données sur des puits individuels et des petits réseaux d'aqueduc obtenus de propriétaires qui ont fait une demande de trousse, ont fait parvenir leur trousse à un laboratoire et qui ont accordé la permission au MENV d'utiliser leurs données. En fait, on pourra plutôt mentionner que des compilations ont été obtenues et non des estimations. En effet, les compilations proviennent de la base de données qui elle-même ne constitue pas un échantillon représentatif.

On pourra aussi mentionner que ces résultats doivent être utilisés à titre indicatif seulement, et qu'ils peuvent permettre de juger de l'importance de certains phénomènes, selon la région par exemple.

Par exemple : Il semble que la proportion de résultats hors normes pour les *E.coli* varie de 0,0 % à près de 16 %. C'est au Bas-Saint-Laurent que la proportion de résultats hors normes est d'environ 16 %.

#### ➤ Les hypothèses sur lesquelles se fonde l'interprétation

En plus de préciser la portée des estimations, on peut en faire l'interprétation tout en spécifiant les hypothèses sur lesquelles elles reposent. En effet, si on fait l'hypothèse que les gens ayant fait une demande de trousse ont des doutes quant à la qualité de l'eau de leur puits, on peut penser que la proportion de puits hors normes (quant aux bactéries *E. coli*) est un maximum. Cependant, comme beaucoup de facteurs n'ont pas été contrôlés lors de « l'échantillonnage » des puits, on ne peut en être assuré puisque les biais potentiels n'ont pu être étudiés. Certains facteurs non contrôlés auraient pu faire

augmenter et d'autres diminuer cette proportion, comme par exemple l'activité humaine près du puits ou encore le type de puits.

De façon similaire, en ce qui concerne la concentration en nitrates, il est possible, de poser les hypothèses suivantes :

- a) Un grand nombre de participants ont demandé des trousse afin de vérifier la qualité de l'eau du puits qui alimente leur résidence secondaire;
- b) La plupart des résidences secondaires ne sont pas situées en zone agricole (donc ces puits sont moins susceptibles d'être contaminés par des activités agricoles).

En posant ces deux hypothèses, on peut penser que la proportion de puits contaminés par les nitrates calculée à partir de la base de données est un minimum. Mais encore une fois, il aurait fallu procéder à une étude de biais pour être rigoureux dans cette affirmation.

#### **4.2 Recommandations pour une campagne ultérieure**

À partir des limites énumérées à la section 3, il est possible d'émettre quelques recommandations à propos d'une campagne de sensibilisation ultérieure. Ces recommandations visent à augmenter le potentiel d'analyse des données recueillies.

➤ Ajout de variables disponibles sur le fichier

On pourrait faire remplir un petit questionnaire aux gens qui font une demande de trousse afin de pouvoir mieux caractériser leurs puits. On pourrait inclure des variables qui sont reliées aux estimations que l'on souhaite obtenir de cette base de données. Par exemple : activité humaine ou agricole près du puits, nature du sol, le type de puits, etc.

Ces informations permettraient d'étudier les biais d'estimations de ceux qui font parvenir leur échantillon d'eau au laboratoire et accordent la permission au MENV d'utiliser les données versus ceux qui font une demande de trousse sans en faire l'analyse.

On pourrait également attribuer un numéro unique à chaque puits dont le propriétaire souhaite en faire l'analyse. Pour cela, il serait nécessaire d'obtenir soit l'adresse du propriétaire ou la localisation exacte de son puits. Cette information permettrait d'aider à effectuer la validation des données (et éviter les doublons).

➤ Définir la population des puits

Si l'on pouvait obtenir un inventaire québécois de tous les puits individuels et des petits réseaux d'aqueduc desservant moins de 21 personnes ainsi que certaines informations pour tous ces puits, on pourrait augmenter grandement le potentiel d'utilisation des données. Les informations pourraient être les suivantes : localisation du puits (région rurale ou urbaine), l'âge du puits, etc. Ces informations devraient être disponibles pour tous ces puits et elles permettraient d'étudier les caractéristiques des puits de la base de

données versus ceux de la population et vérifier s'il y a des biais qui pourraient affecter les estimations.

Toutefois, même en possédant la population des puits, il ne sera pas possible d'inférer à la population des puits, ni d'effectuer des tests statistiques puisque la campagne se fait sur une base volontaire et que ce n'est pas un échantillon aléatoire qui serait sélectionné à partir de cette liste. Donc on ne pourrait tout de même pas déclarer que la base de données constitue un échantillon représentatif de la population.

## **Section 5 : Conclusions**

La base de données ne peut être considérée pour réaliser des estimations objectives sur la qualité de l'eau des puits et des petits réseaux d'aqueduc québécois. Le profil des données fournies par les participants au programme risquant de différer sensiblement de celui qu'aurait pu nous fournir les non-participants au programme, il serait même périlleux d'utiliser cette base de données, car ceci pourrait mener à des constats qui ne se vérifieraient pas dans la population

Les compilations qui seront produites à partir de cette base de données devront être accompagnées de sérieuses mises en garde pour éviter une confusion quant à l'interprétation de ces résultats. On pourra tout de même dégager des conclusions d'ordre général quant à la campagne de sensibilisation et quelques résultats approximatifs tirés de la base de données.

par  
Catherine Fontaine  
Statisticienne-méthodologiste  
et Sylvain Végiard  
Coordonnateur des enquêtes spéciales

Institut de la statistique du Québec  
Le 4 février 2002