

# Protocole d'étude

## Evaluation des résultats des laboratoires d'insémination

### -Activité 2018-

#### Contexte

Depuis 2005, l'Agence de la biomédecine a pour mission de suivre et d'évaluer les activités cliniques et biologiques d'assistance médicale à la procréation (AMP). La loi de 2011 prévoit que l'Agence publie régulièrement les résultats des activités d'AMP en tenant compte des caractéristiques de leur patientèle et en particulier de l'âge des femmes.

Les laboratoires qui assurent la préparation de sperme en vue d'insémination envoient chaque année un rapport faisant le bilan de leur activité à l'Agence de la biomédecine, sous forme de données agrégées.

Ce rapport présente les résultats de l'analyse statistique portant sur l'évaluation des résultats des inséminations artificielles intra-utérines intraconjugales de l'année 2018 en France, en tenant compte des caractéristiques de leur patientèle disponible dans leur rapport d'activité et du nombre de cycles effectués. La méthodologie a été élaborée en collaboration avec un groupe d'experts (cliniciens, biologistes, épidémiologistes) et validée par le groupe de travail « Stratégie AMP » au sein de l'Agence de la biomédecine.

#### Objectif

L'objectif général de l'étude est de comparer les résultats des inséminations en 2018 des laboratoires à la moyenne nationale.

L'objectif de l'analyse statistique est de mettre en place un test qui permet de comparer le nombre d'accouchements observés par laboratoire au nombre attendu sous l'hypothèse de la moyenne nationale, en tenant compte des caractéristiques des patientèles dans les laboratoires.

#### Méthodes

##### Population d'étude

###### Critères d'inclusion

- Les cycles d'insémination intraconjugaux effectués en 2018

###### Critères d'exclusion

- Les cycles d'insémination avec don de spermatozoïdes
- Les laboratoires ayant réalisé moins de 50 inséminations

## Sources d'information – Variables étudiées

Les données sont issues des rapports d'activité agrégés 2018 dont la base a été figée en juillet 2020.

### Processus étudié

- Nombre de cycles en vue d'insémination

### Résultat des inséminations

- Nombre d'accouchements (enfant  $\geq 22$  SA ou  $\geq 500$ g), quel que soit le statut vital des enfants à la naissance, hors IMG
- Nombre d'accouchements multiples (naissance de 2 enfants ou plus)

### Variables d'ajustement

Les variables d'ajustement des résultats ne peuvent concerner que les caractéristiques de la patientèle des laboratoires. Les différentes techniques et types de prise en charge des couples, connus pour améliorer les résultats, ne doivent pas être utilisés dans les modèles d'ajustement dans la mesure où ils sont partie intégrante des stratégies mises en place par l'équipe. Les variables d'ajustement retenues sont :

- Répartition de l'âge des femmes : les classes d'âge utilisées sont celles qui sont recueillies dans le rapport annuel d'activité des laboratoires.

### Validation des données

Une validation de la cohérence interne des déclarations des rapports d'activité annuelle a été réalisée, afin d'exclure les risques d'erreurs de déclaration ou de saisie.

## Analyse statistique

### Indicateurs de résultat

Deux indicateurs de résultats ont été étudiés :

- **Indicateur 1** : taux d'accouchement, rapporté au nombre de cycles d'insémination. Il s'agit du nombre d'accouchements survenant après 22 semaines d'aménorrhée quel que soit le statut vital des enfants à la naissance, rapporté au nombre de cycles d'insémination réalisés en 2018.
- **Indicateur 2** : fréquence des accouchements multiples. Il s'agit du nombre d'accouchements de 2 enfants ou plus parmi l'ensemble des accouchements tels que définis ci-dessus.

### Modèles d'analyse

L'analyse des facteurs prédictifs a été réalisée avec la méthode de régression logistique pour données agrégées. Cette méthode permet de modéliser la relation entre les indicateurs de résultat et les variables d'ajustement. Une pondération normalisée de la régression logistique a été utilisée pour tenir compte de la taille des laboratoires (nombre de cycles) dans l'identification des facteurs prédictifs.

- Concernant l'analyse du taux d'accouchement, comme il n'y avait qu'un seul facteur d'ajustement, un modèle logistique univarié a été réalisé afin d'estimer l'effet prédictif sur le taux d'accouchement. Les laboratoires ayant effectués moins de 50 inséminations en 2018 ont été exclus.
- Aucun facteur d'ajustement n'a été utilisé pour l'analyse de la fréquence des accouchements multiples parce que ces facteurs ne sont pas disponibles. La fréquence des accouchements multiples parmi l'ensemble des accouchements est un indicateur qui vient compléter l'interprétation de l'indicateur précédent. Pour cet indicateur, les laboratoires ayant effectué moins de 50 inséminations ont été exclus.

Ces pondérations sont estimées au niveau national et sont utilisées pour standardiser les taux d'accouchement des laboratoires de manière à les rendre comparables pour les caractéristiques étudiées.

La relation entre les indicateurs de résultat et les facteurs prédictifs est quantifiée sous forme d'Odds Ratio (OR). Ils mesurent la variation du taux d'accouchement pour chaque unité supplémentaire du facteur étudié : par exemple, si l'OR est supérieur à 1 pour les moins de 35 ans, alors on attend plus d'accouchements pour le laboratoire qui a plus de femmes de cet âge. Inversement, si l'OR est inférieur à 1 pour la classe d'âge des plus de 40 ans, on attend moins d'accouchements pour les laboratoires ayant plus de femmes âgées.

### **Le taux d'accouchement standardisé**

Le taux d'accouchement standardisé d'un laboratoire est le taux que l'on observerait dans ce laboratoire s'il avait la même patientèle que celle observée au niveau national. La standardisation a pour objectif de permettre une comparaison avec la moyenne nationale. Le calcul du taux standardisé est déduit du taux attendu d'accouchement dans le laboratoire au regard des caractéristiques de sa patientèle.

Le taux d'accouchement attendu d'un laboratoire est le taux d'accouchement prédit pour le laboratoire lorsqu'on tient compte des caractéristiques de sa patientèle. Il est estimé en appliquant à la patientèle du laboratoire les pondérations prédictives observées au niveau national pour chacune des caractéristiques étudiées. Ces pondérations sont issues du modèle de régression logistique.

Pour chaque laboratoire, le taux d'accouchement observé est rapporté au taux d'accouchement attendu pour mesurer l'écart du laboratoire à la moyenne nationale. Le taux standardisé est estimé par le produit de ce rapport avec le taux d'accouchement national observé :

$$\text{Taux standardisé du laboratoire} = \frac{\text{taux observé dans le laboratoire}}{\text{taux attendu dans le laboratoire}} \times \text{taux observé national}$$

Un exemple de calcul est présenté en annexe I.

### Test statistique d'écart à la moyenne nationale

La méthode du « **funnel plot** » (Spiegelhalter, DJ. 2005)<sup>1</sup> permet de comparer le taux standardisé d'accouchements de chaque laboratoire à l'intervalle de confiance de la moyenne nationale en fonction du volume d'activité du laboratoire (nombre de cycles).

La construction de l'intervalle de confiance de la moyenne nationale en fonction du nombre de cycles est réalisée avec la méthode du logit (Collett D., 2003)<sup>2</sup>.

Deux intervalles de confiance de la moyenne nationale sont présentés dans les funnel plots : à 95% (pointillés) et à 99% (trait plein). Ces deux intervalles de confiance correspondent à des risques différents de conclure à tort à une différence d'avec la moyenne nationale. L'intervalle de confiance à 99% est utilisé pour tenir compte du fait que la répétition des tests statistiques par rapport à la moyenne nationale (un test statistique pour chaque laboratoire) augmente le risque de conclure à tort à une différence. L'intervalle de confiance à 95% est présenté à titre indicatif, il constitue un niveau d'alerte pour les laboratoires lorsque leur taux standardisé se trouve au-delà de cette limite.

Si le taux d'accouchement standardisé d'un laboratoire est inférieur à la borne inférieure de l'intervalle de confiance à 99%, cela indique que le taux d'accouchement du laboratoire est considéré comme significativement inférieur au taux d'accouchement observé au niveau national ; inversement lorsqu'il est supérieur à la borne supérieure, le taux d'accouchement du laboratoire est significativement supérieur au taux observé au niveau national.

---

<sup>1</sup> Spiegelhalter DJ. Funnel plots for comparing institutional performance. Stat Med. 2005 Apr 30; 24(8):1185-202.

<sup>2</sup> Collett D. Modelling Binary Data, Second edition, 2003

## Annexe 1 : Exemple de calcul du taux d'accouchement standardisé

Le *taux d'accouchement standardisé* d'un laboratoire est le taux que l'on observerait dans ce laboratoire s'il avait la même patientèle que celle observée au niveau national. La standardisation a pour objectif de permettre une comparaison avec la moyenne nationale. Le calcul du taux standardisé est déduit du taux attendu d'accouchement dans le laboratoire.

Le *taux d'accouchement attendu* d'un laboratoire est le taux d'accouchement prédit dans le laboratoire lorsqu'on tient compte des caractéristiques de sa patientèle. Il est estimé en appliquant à la patientèle du laboratoire les pondérations prédictives observées au niveau national pour chacune des caractéristiques étudiées. Ces pondérations sont issues du modèle de régression logistique d'analyse multivariée. On nommera cette fonction de pondération la fonction prédictive (FP).

- Equation de la fonction prédictive (FP) issue du modèle de régression logistique :

$$FP = \alpha + \beta_1 * (\% \text{ «30-34 ans »}) + \beta_2 * (\% \text{ «35-37 ans »}) + \beta_3 * (\% \text{ «38-39 ans »}) + \beta_4 * (\% \text{ «40-42 ans »}) + \beta_5 * (\% \text{ «>=43 ans »})$$

- Taux attendu =  $\frac{e^{FP}}{1 + e^{FP}}$

- Taux standardisé du laboratoire =  $\frac{\text{taux observé dans le laboratoire}}{\text{taux attendu dans le laboratoire}} \times \text{taux observé national}$

Considérons le taux d'accouchement par cycle estimé sur l'activité 2018. Les pondérations prédictives du modèle de régression logistique sont présentées dans le tableau suivant :

Variables	Modalité	$\beta$	Valeur des $\beta$	Odds Ratio ( $\exp(\beta)$ )
Intercept		$\alpha$	-2,1444	-
Age des femmes	% <30 ans			1
	% [30-34 ans]	$\beta_1$	0,00629	1,006
	% [35-37 ans]	$\beta_2$	-0,00071	0,999
	% [38-39 ans]	$\beta_3$	-0,00129	0,999
	% [40-42 ans]	$\beta_4$	-0,00836	0,992
	% >=43 ans	$\beta_5$	-0,0389	0,962

Prenons comme exemple un laboratoire ayant :

- 40,6% de femmes de 30-34 ans
- 11,1% de femmes de 35-37 ans
- 2,2% de femmes de 38-39 ans
- 2,2% de femmes de 40-42 ans
- 0% de femmes de plus de 43 ans
- Taux d'accouchement observé : 0,104
- Taux d'accouchement observé au niveau national : 0,105

L'estimation du taux standardisé pour ce laboratoire est :

- $FP = -2,1444 + 0,00629 * 40,6 - 0,00071 * 11,1 - 0,00129 * 2,2 - 0,00836 * 2,2 - 0,0389 * 0$
- Taux d'accouchement attendu =  $\frac{e^{FP}}{1 + e^{FP}} \approx 12,8\%$
- Taux d'accouchement standardisé =  $(0,104 / 0,128) * 0,105 = 8,5\%$