**Statistiques inférentielles sur les variables de poids ; âge ; taille et sexe sur les enfants de 0-3 ans et de 8-12 ans**

**Analyser s’il y a des différences significatives entre les villes/altitudes pour ces variables.**

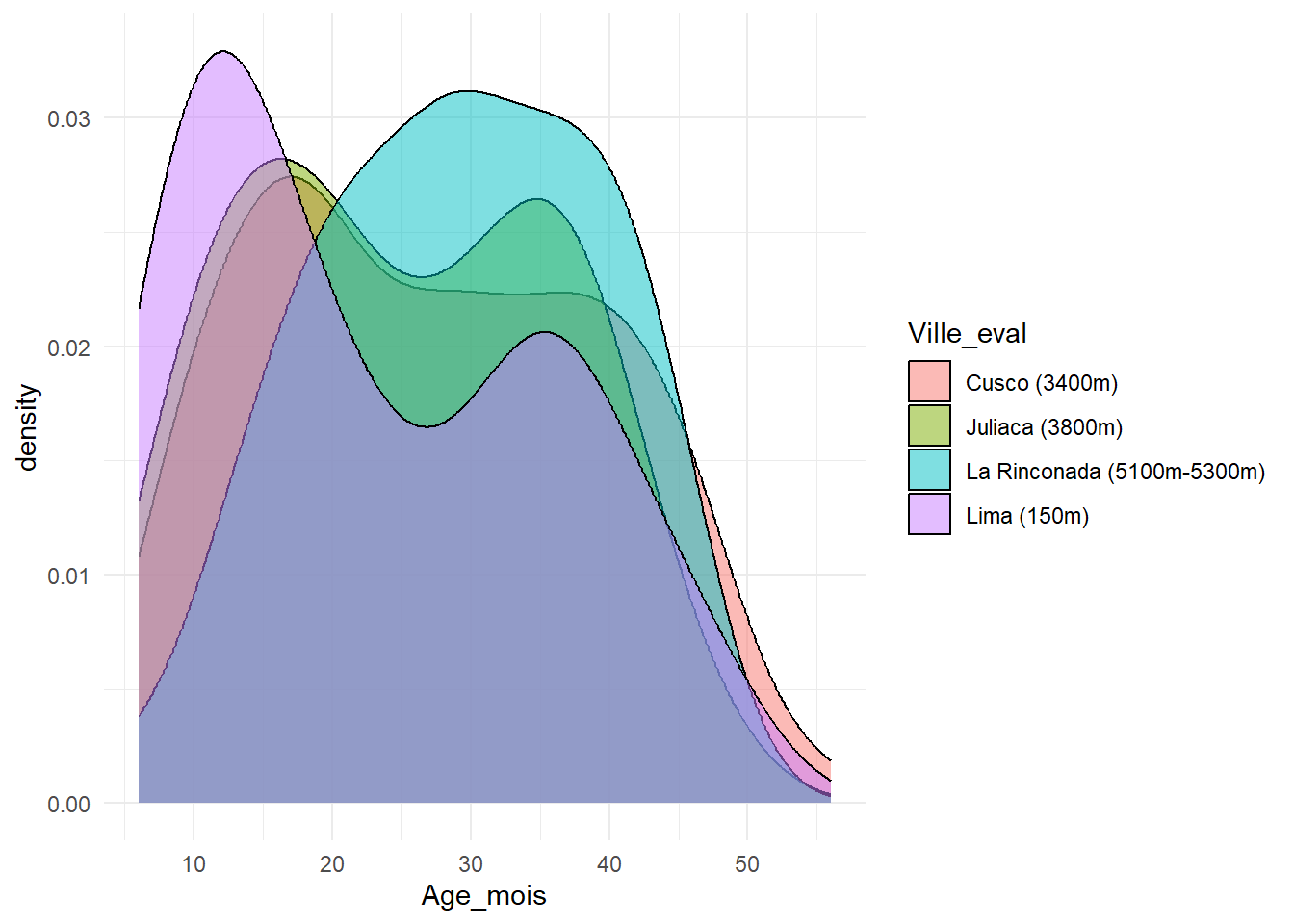
Pour les données qui ne sont pas normalement distribuées, la représentation graphique se fera sous la forme d’un box plot.

A cet égard les données qui ne sont pas normalement distribuées et qui auront donné lieu à des tests non paramétriques (donc des tests de rang) seront exprimées dans le tableau en médiane et 25ème et 75ème percentile. Les données normalement distribuées seront quant à elles exprimées sous la forme de moyenne ± écart type.

1. **AGE en mois**

**Evaluation de la normalité des données d’âge en mois de chaque ville.**

Une image contenant texte, diagramme, Tracé

Le contenu généré par l’IA peut être incorrect.

**0-3 ans 8-12 ans**

Graphiques représentant la distribution des données avec la fonction de densité.

Mise en place d’un test de Shapiro-Wilk pour confirmer la non-normalité visuelle : test de Shapiro-Wilk significatif pour au moins un groupe que ce soit pour les 0-3 ans ou les 8-12 ans donc test non paramétrique pour les deux. Le test utilisé sera celui de Kruskal-Wallis.

**Pour les 0-3 ans, le modèle global est significatif et les tests post-hoc mettent évidence la différence significative suivante d’âge :**

**Une image contenant texte, diagramme, capture d’écran, ligne

Le contenu généré par l’IA peut être incorrect.**

**Pour les 8-12, pas de différences significatives :**

**Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.**

1. **Poids**

**Une image contenant texte, diagramme, Tracé

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, diagramme, Tracé, ligne

Le contenu généré par l’IA peut être incorrect.Evaluation de la normalité des données de poids de chaque ville.**

**0-3 ans 8-12 ans**

Graphiques représentant la distribution des données avec la fonction de densité.

Mise en place d’un test de Shapiro-Wilk pour s’assurer de la nature de la distribution des données de poids. Test de Shapiro-Wilk significatif pour au moins un groupe que ce soit pour les 0-3 ans ou les 8-12 ans donc test non paramétrique pour les deux. Le test utilisé sera celui de Kruskal-Wallis.

**Pour les 0-3 ans, le modèle global n’est pas significatif :**

**Une image contenant texte, diagramme, capture d’écran, ligne

Le contenu généré par l’IA peut être incorrect.**

**Pour les 8-12 ans, le modèle global est significatif et les tests post-hoc mettent évidence les différences significatives suivantes de poids en fonction des villes :**

**Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.**

1. **Taille**

**Evaluation de la normalité des données de taille de chaque ville.**

Une image contenant texte, diagramme, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, diagramme, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.

**0-3 ans 8-12 ans**

Graphiques représentant la distribution des données avec la fonction de densité.

Mise en place d’un test de Shapiro-Wilk pour s’assurer de la nature de la distribution des données de poids.

Pour les 0-3 ans, au moins un des quatre groupes à une p-value du shapiro-wilk significative ce qui signifie des données non normales et l’utilisation du test non paramétrique Kruskal-Wallis.

Pour les 8-12 ans en revanche, aucune des p-value n’est significative, ce qui signifie que les données de taille sont normalement distribuées. L’homogénéité des variances du poids dans chaque ville du groupe 8-12 ans a été vérifiée grâce à un test de Levene non significatif. On effectuera donc une ANOVA à une voie pour ce groupe.

**Pour les 0-3 ans, le modèle global n’est pas significatif :**

**Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.**

**Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.Pour les 8-12 ans, le modèle global est significatif et les tests post-hoc mettent évidence les différences significatives suivantes de poids en fonction des villes :**

1. **SEXE**

Pour le sexe, utilisation du test du chi 2 pour déterminer si les proportions de filles ou de garçon diffèrent significativement entre les villes des 0-3 ans et des 8-12 ans.

**Pour les 0-3 ans, le modèle global n’est pas significatif :**

Une image contenant texte, capture d’écran, diagramme, Police

Le contenu généré par l’IA peut être incorrect.

**Pour les 8-12 ans, le modèle global est significatif et les tests post-hoc mettent évidence les différences significatives suivantes de sexe en fonction des villes :**

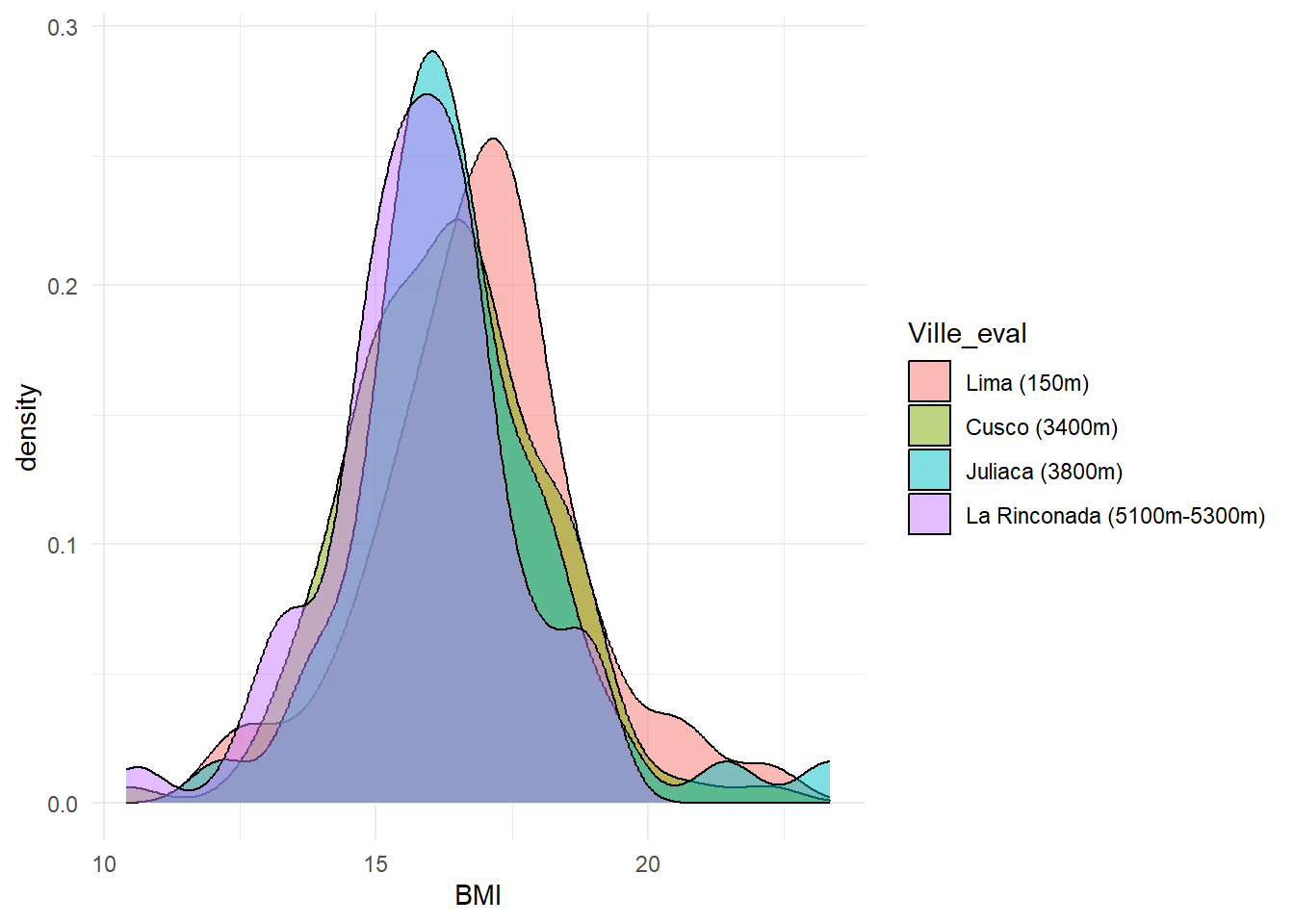
**Une image contenant texte, capture d’écran, diagramme, Police

Le contenu généré par l’IA peut être incorrect.**

1. **BMI**

Vérification distribution des données :

Une image contenant texte, diagramme, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.

**0-3 ans 8-12 ans**

Des tests de Shapiro-wilk ont été réalisé pour connaître la nature précise de la distribution des données. Dans les 0-3 ans et 8-12 ans, au moins un groupe avait un test significatif (p < 0.05) donc au moins un groupe avait une distribution non normale donc tests non-paramétriques.

**0-3 ans : comparaison significative**

**Une image contenant texte, capture d’écran, diagramme, ligne

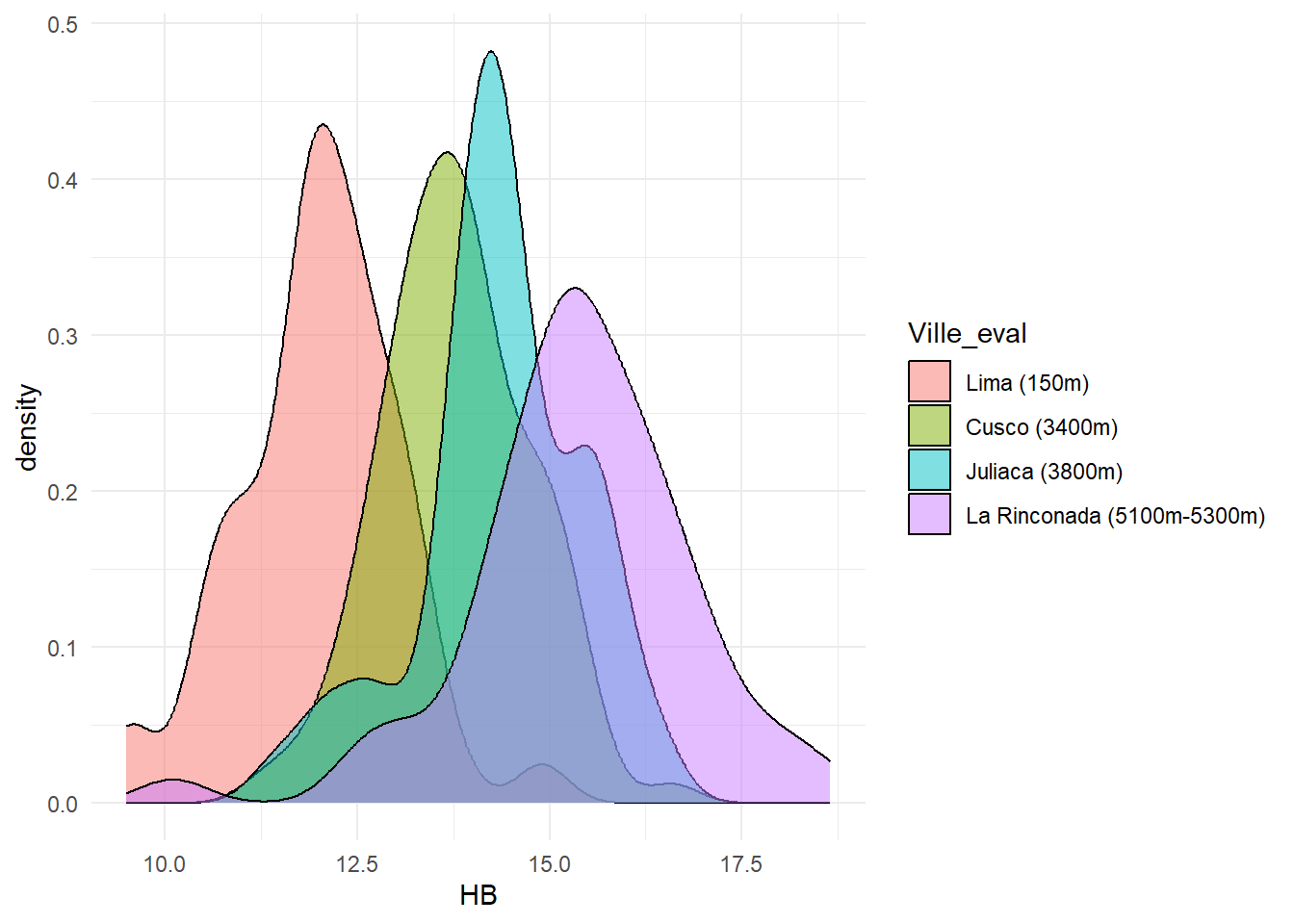
Le contenu généré par l’IA peut être incorrect.**

Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.**8-12 ans : modèle significatif**

1. **Concentration en hémoglobine [HB]**

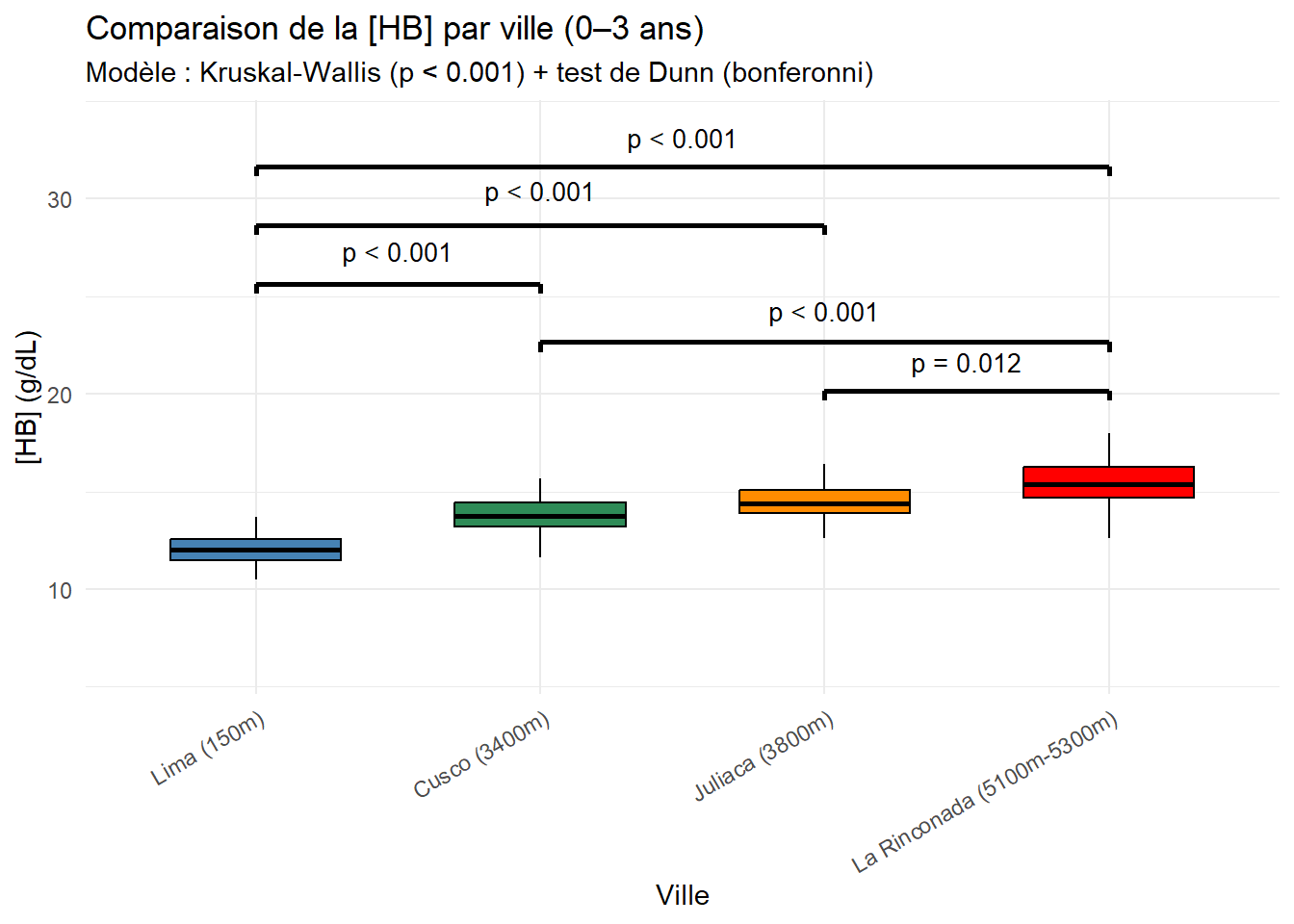
Une image contenant texte, diagramme, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.****Distribution des données :

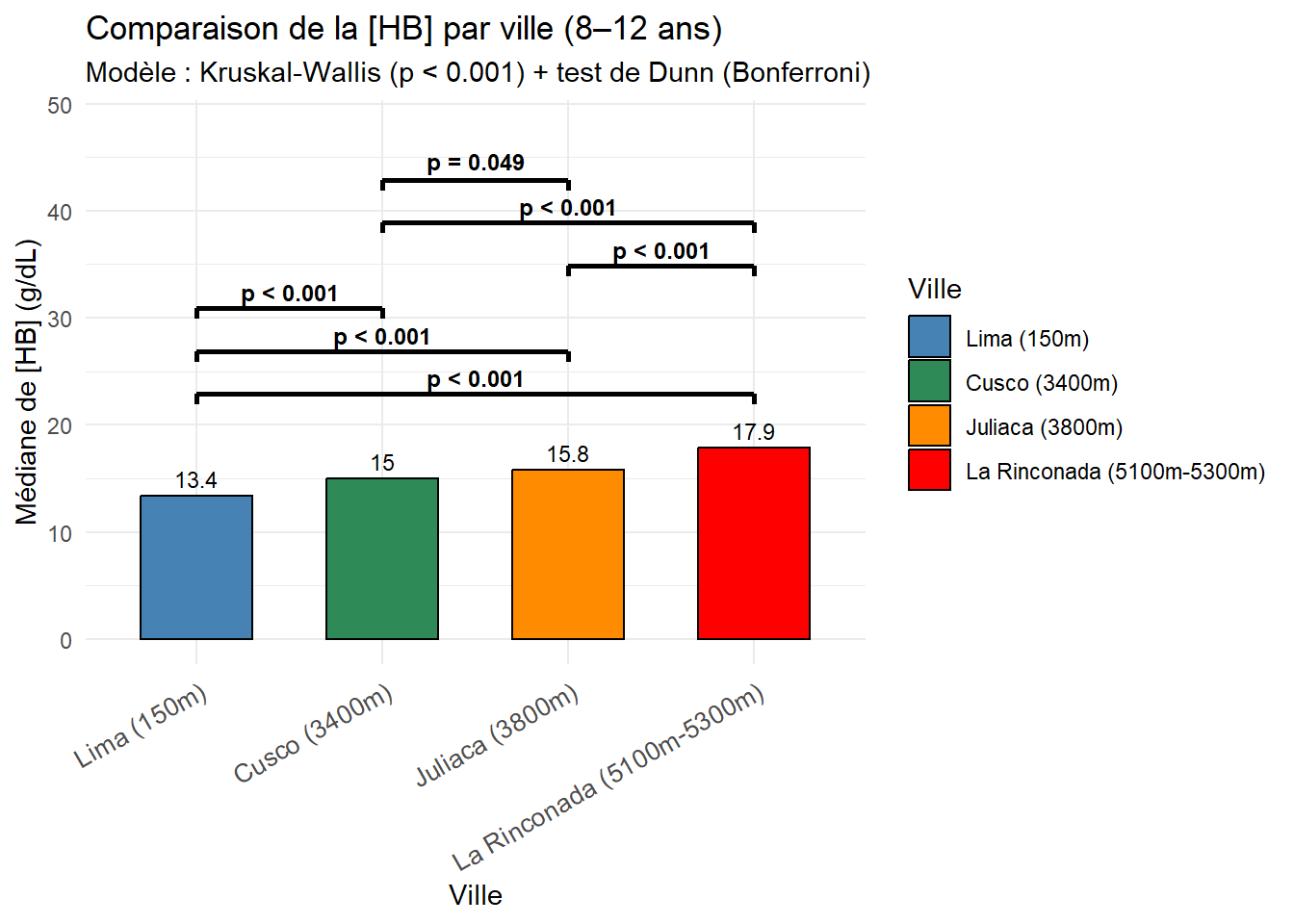
**0-3 ans 8-12 ans**

Au moins 1 ville dans les deux groupes a un test de Shapiro wilk significatif donc test non paramétrique pour les deux.

**[HB] 0-3 ans : modèle significatif :**

****

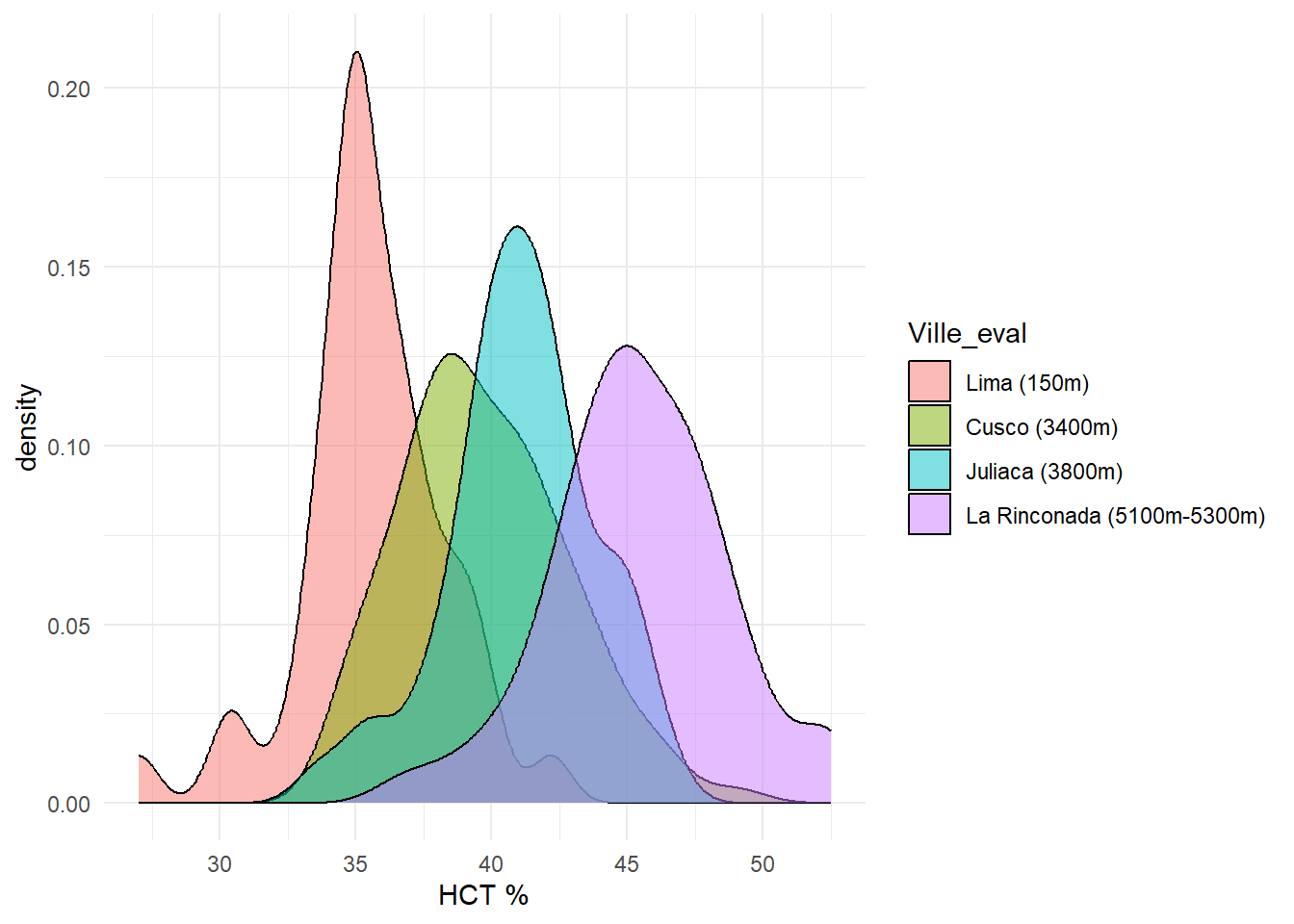
**[HB] 8-12 ans : modèle significatif :**

****

Les bar-plots représentent bien la médiane en accord avec les tests de rang utilisés mais avec toutes les comparaisons significatives à afficher, j’ai eu du mal à créer un graphique avec des box plots clairs.

1. **Taux d’Hématocrite**

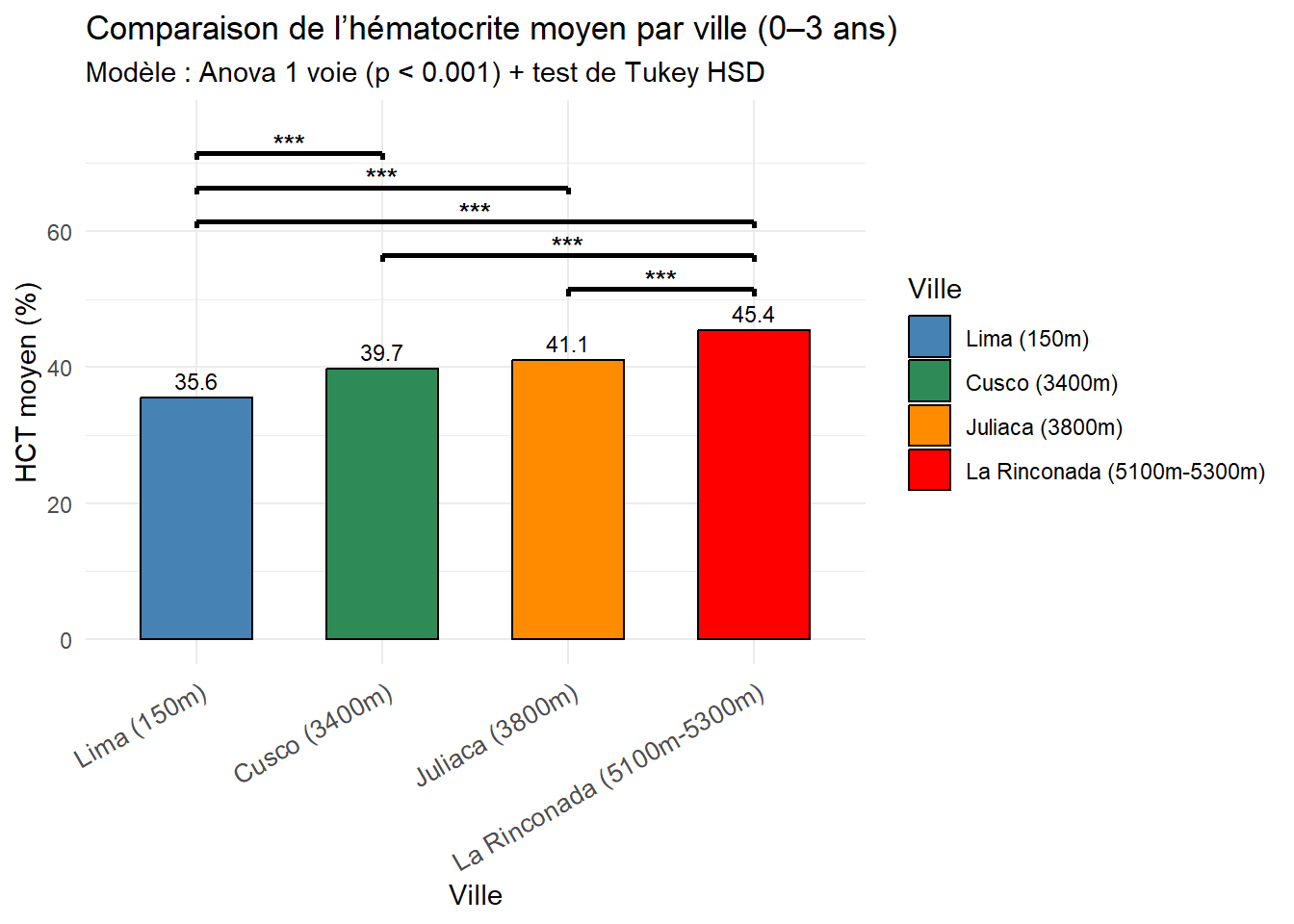
Une image contenant texte, diagramme, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.Distribution des données :

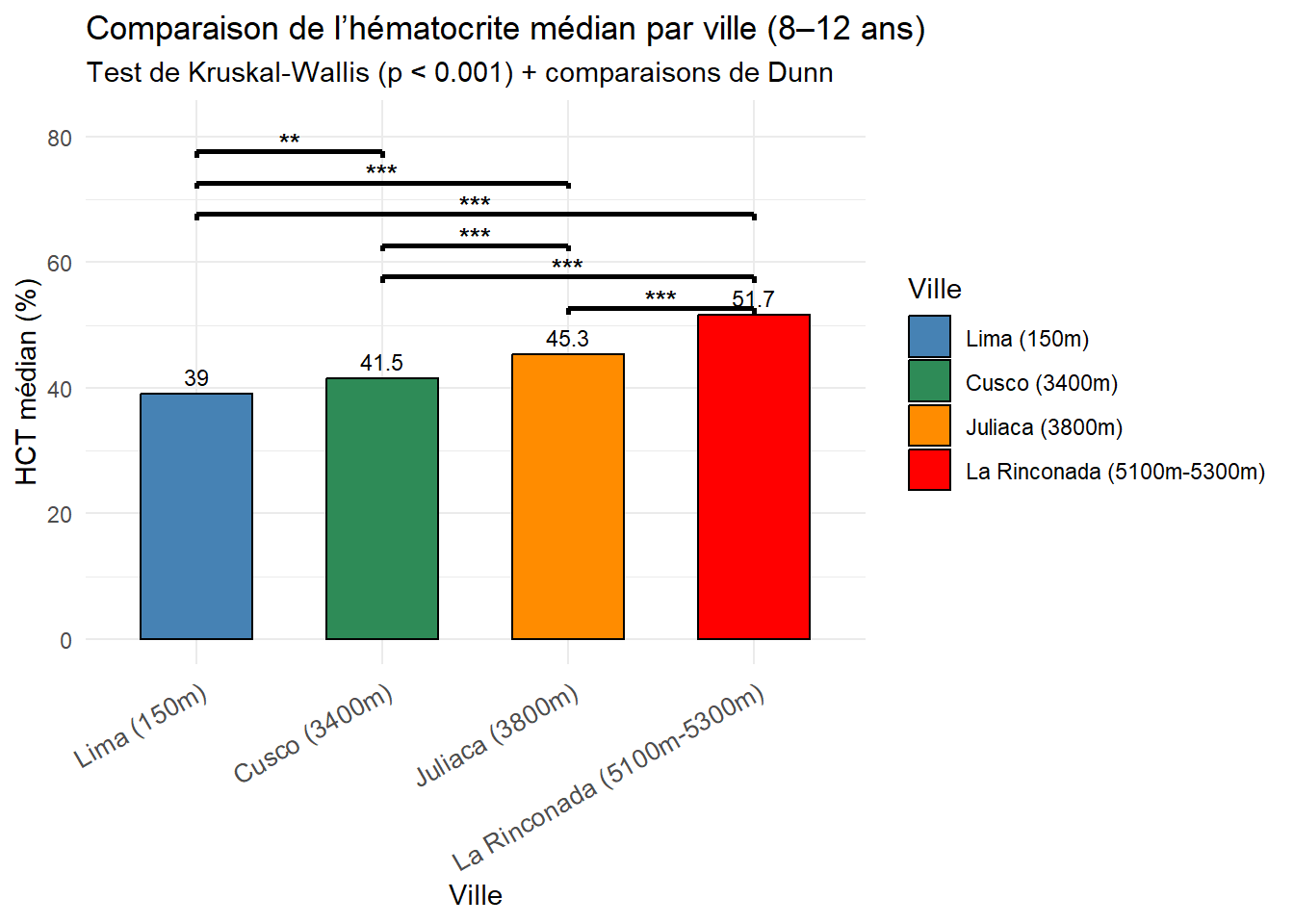
**0-3 ans 8-12 ans**

Pour les 0-3 ans le Shapiro-Wilk non significatif indique que les données sont normalement distribuées pour toutes les villes, ce qui n’est pas le cas des 8-12 ans. Nous réaliserons donc un test paramétrique pour les 0-3 ans et non paramétrique pour les 8-12 ans.

**0-3 ans : modèle significatif :**

\*\*\* = p < 0.001

**8-12 ans : modèle significatif :**



\*\*\* = p < 0.001

1. **Masse totale d’hémoglobine (HB mass), uniquement pour les 8-12 ans**

Une image contenant texte, diagramme, Tracé

Le contenu généré par l’IA peut être incorrect.Une image contenant texte, diagramme, Tracé, capture d’écran

Le contenu généré par l’IA peut être incorrect.Distribution des données :

**HB mass absolue (g) HB mass relative (g / kg)**

Pour les deux variables, au moins une ville n’a pas ses données normalement distribuées donc test non paramétrique pour les deux.

**HB mass absolue, modèle significatif :**

**Une image contenant texte, capture d’écran, diagramme, Parallèle

Le contenu généré par l’IA peut être incorrect.**

**HB mass relative rapportée au poids, modèle significatif également :**

Une image contenant texte, capture d’écran, diagramme, ligne

Le contenu généré par l’IA peut être incorrect.

\*\*\* = p < 0.001

\*\* = p < 0.01

\* = p < 0.05

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Lima (150 m)** | **Cusco (3400 m)** | **Juliaca (3800 m)** | **La Rinconada (5100m -5300m)** |
| **Age (mois) (n = 271)** | **21** ; 11.5/35  N = 51 | **27** ; 17/37.75  N = 114 | **24** ; 15.75/35  N= 48 | **29.5** ; 21/37 **\***  N = 58 |
| **Poids (kg) (n = 257)** | **11** ; 9.4/13.25  N = 47 | **11.525** ; 9.98/14.1  N = 110 | **11.65** ; 10/13  N = 45 | **12**; 9.8/14  N = 55 |
| **Taille (cm) (n = 257)** | **80** ; 72/92.5  N = 47 | **85.5** ; 78/93  N = 110 | **85** ; 77.5/91.7  N= 45 | **86.5** ; 80/93  N = 55 |
| **Sexe (Fille = F ou Garçon = G en %) (n = 269)** | **F = 51 %**  **G = 49 %**  N = 49 | **F = 43,9 %**  **G = 56,1 %**  N = 114 | **F = 35,4 %**  **G = 64,6 %**  N = 48 | **F = 46,6 %**  **G = 53,4 %**  N = 58 |
| **BMI (n = 257)** | **16.9** ; 15.8/17.8  N = 47 | **16.3** ; 15.1/17.3  N = 110 | **16.3** ; 15.5/17.2  N = 45 | **15.8** ; 15/16.7 **\***  N = 55 |
| **[HB] (g / dL) (n = 247)** | **12** ; 11.45/12.55  N = 47 | **13.75** ; 13.20/14.45 **\***  N = 98 | **14.4** ; 13.9/15 **\***  N = 45 | **15.35** ; 14.7/16.25 **\* # §**  N = 57 |
| **Hématocrite (%) (n = 234)** | **35.5** ± 2.7  N = 41 | **39.7** ± 3 **\***  N = 96 | **41** ± 2.7 **\* §**  N = 44 | **45.4** ± 3.2 **\* § #**  N = 53 |

**Données pour les enfants de 0 à 3 ans (n = 271)**

* **\*** = différence significative avec Lima (150 m)
* **§** = différence significative avec Cusco (3400 m)
* **#** = différence significative avec Juliaca (3800 m)

Si les données sont normalement distribuées, les résultats sont présentés de façon **moyenne ± écart type** pour les variables continues (âge, poids, taille) et en % de l’échantillon pour le sexe. Si les données ne sont pas normales les résultats prennent la forme suivante : **médiane** ; **25ème / 75ème percentile.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Lima (150 m)** | **Cusco (3400 m)** | **Juliaca (3800 m)** | **La Rinconada (5100m -5300m)** |
| **Age (mois) (n = 364)** | **122** ; 107.75/134  N = 84 | **118** ; 103/130  N = 59 | **126** ; 109/134  N= 125 | **126** ; 107.75/136  N = 96 |
| **Poids (kg) (n = 362)** | **38.5** ; 32.55/45.75  N = 83 | **30.95** ; 26/38.7 **\***  N = 59 | **35.5** ; 29/42.5  N = 125 | **30** ; 25.45/40 **\* #**  N = 95 |
| **Taille (cm) (n = 362)** | **140.8** ± 10,75  N = 83 | **136** ± 9,38 **\***  N = 59 | **137.8** ± 9,34  N= 125 | **136.7** ± 11,07 **\***  N = 95 |
| **Sexe (Fille = F ou Garçon = G en %) (n = 363)** | **F = 39,8 %**  **G = 60,2 %**  N = 83 | **F = 64,4 % \***  **G = 35,6 %**  N = 59 | **F = 48,8 %**  **G = 51,2 %**  N = 125 | **F = 43,8 %**  **G = 56,2 %**  N = 96 |
| **BMI (n = 362)** | **19.7** ; 17.1/22  N = 83 | **16.5** ; 15.4/20.2 **\***  N = 59 | **18.4** ; 16.3/21.1  N = 125 | **16.7** ; 15.1/18.7 **\* #**  N = 95 |
| **[HB] (g / dL) (n = 356)** | **13.4** ; 12.85/13.85  N = 83 | **15** ; 14.5/15.75 **\***  N = 55 | **15.8** ; 15.3/16.4 **\***  N = 125 | **17.9** ; 17.1/18.6 **\* # §**  N = 93 |
| **Hématocrite (%) (n = 352)** | **39** ; 37.2/40.9  N = 83 | **41.5** ; 40.4/43.8 **\***  N = 53 | **45.3** ; 43.9/47.2 **\* §**  N = 124 | **51.7** ; 49.3/53.5 **\* # §**  N = 92 |
| **Masse totale d’hémoglobine absolue (g) (n = 304)** | **309**; 265/357  N = 76 | **303**; 270/378  N = 44 | **348** ; 305/416 **\***  N = 109 | **410**; 343/514 **\* § #**  N = 75 |
| **Masse totale d’hémoglobine relative rapportée au poids (g / kg) (n = 304)** | **7.89** ; 7.2/8.9  N = 76 | **9.35**; 8.5/10 **\***  N = 44 | **9.6**; 8.4/10.9 **\***  N = 109 | **12.5**; 11.3/13.3 **\* § #**  N = 75 |

**Données pour les enfants de 8 à 12 ans (n = 364)**

* **\*** = différence significative avec Lima (150 m)
* **§** = différence significative avec Cusco (3400 m)
* **#** = différence significative avec Juliaca (3800 m)

Si les données sont normalement distribuées, les résultats sont présentés de façon **moyenne ± écart type** pour les variables continues (âge, poids, taille) et en % de l’échantillon pour le sexe. Si les données ne sont pas normales les résultats prennent la forme suivante : **médiane** ; **25ème / 75ème percentile.**

**Remarques :**

Il est intéressant de noter que pour les 0-3 ans, où l’âge est significativement différent (notamment entre Lima et La Rinconada), aucun effet significatif n’apparaît pour la taille et le poids. Au contraire chez les 8-12 ans, aucune différence significative d’âge n’existe. Dès lors, plusieurs différences significatives apparaissent entre les villes pour le poids et la taille. Ainsi, soit dès lors que l’effet de l’âge est inexistant (comme pour les 8-12 ans), l’effet spécifique de l’altitude sur le poids et la taille se révèle ; soit ces différences de poids et de taille sont spécifiques aux 8-12 ans et non altitude ; soit un peu des deux.

**Il est aussi intéressant de noter que pour l’[HB] et l’HCT %, la comparaison Cusco-Juliaca est la seule qui n’est pas significative pour les 0-3 ans et que lorsqu’elle l’est pour les 8-12 ans, la p value est proche du seuil limite au contraire des autres paires > effet altitude diminué par écart altitude pas si important ?**