

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/257700353>

# Principe et raisonnement en biostatistiques

**Article** in *Kinésithérapie la Revue* · June 2013

DOI: 10.1016/j.kine.2013.03.002

---

CITATION

1

---

READS

142

**1 author:**



**Mathieu Vergnault**  
University of Bordeaux

23 PUBLICATIONS 40 CITATIONS

SEE PROFILE



# Principe et raisonnement en biostatistiques

## *Principle and reasoning in biostatistics*

Groupe d'études et de recherches appliquées à la rééducation (GERAR), société française des professionnels en activités physiques adaptées (SFP-APA), 20, rue d'Aubervilliers, 75019 Paris, France

Mathieu Vergnault

Reçu le 23 mai 2012 ; accepté le 8 mars 2013

### RÉSUMÉ

Dans la structure d'un article scientifique, les biostatistiques possèdent une place à part entière. Au sein d'une pratique basée sur les preuves, les analyses statistiques objectivent la présence d'arguments en faveur ou non des hypothèses posées au départ par le chercheur. Mais également parce que la conclusion d'une étude repose sur les résultats des tests statistiques réalisés à partir des groupes comparés. L'inférence statistique nécessite une méthodologie et une rigueur appropriée afin de prendre en compte les indices nécessaires pour évaluer la qualité du traitement statistique. Le sérieux d'une analyse statistique, via l'utilisation raisonnée des tests statistiques, permet d'éviter certains biais afin d'interpréter correctement la dualité existante entre la significativité statistique et le cliniquement significatif.

Niveau de preuve. – Non adapté.

© 2013 Elsevier Masson SAS. Tous droits réservés.

### SUMMARY

*In a scientific article's framework, biostatistics have a full place. According to a practice based on evidence, statistical analysis focus on the presence of arguments in favor or not, of the assumptions made initially by the researcher. In a second hand, studies' conclusions are based on the results of statistical tests made from groups to compare. Statistical inference requires rigour and appropriate methodology to consider the clues needed to assess the quality of statistical processing. The seriousness of a statistical analysis avoids some biases, and the rational use of tests to interpret correctly the duality between statistical significance and clinically significant.*

Level of evidence. – Not applicable.

© 2013 Elsevier Masson SAS. All rights reserved.

Un travail de recherche vise presque toujours à comparer deux ou plusieurs groupes, en vérifiant si l'existence d'une différence est liée à divers signes diagnostiques ou à certains traitements thérapeutiques [1]. Le clinicien examine alors si la différence observée est également due aux fluctuations du hasard [1–4]. Pour revenir aux bases de la probabilité, la chance de trouver un résultat plutôt qu'un autre ne se retrouve pas que dans les lancers de dés [5].

Le raisonnement statistique veut qu'à la suite de diverses observations, une ou plusieurs hypothèses soient énoncées et que celles-ci soient vérifiées grâce à des tests statistiques [2]. Au lieu d'essayer de démontrer qu'un événement est probablement arrivé, les analyses statistiques tentent de démontrer que c'est improbable que cet événement ne se passe pas [4].

Toute analyse statistique fréquentiste se base sur le test d'hypothèses. Pour départager

### Mots clés

Biais  
Intervalle de confiance  
Puissance statistique  
Signification statistique  
Tests d'hypothèses

### Keywords

Bias  
Confidence intervals  
Statistical power  
Statistical significance  
Hypothesis testing

DOIs des articles originaux :

<http://dx.doi.org/10.1016/j.kine.2013.03.003>,  
<http://dx.doi.org/10.1016/j.kine.2013.03.001>,  
<http://dx.doi.org/10.1016/j.kine.2013.03.004>

Adresse e-mail :  
[mathieu.vergnault@hotmail.fr](mailto:mathieu.vergnault@hotmail.fr)

## Biostatistiques : la dimension collective de nos individualités

deux distributions, le chercheur formule une hypothèse principale, appelée hypothèse nulle ou  $H_0$ . Celle-ci assume qu'il n'existe pas de différence entre les échantillons testés, autrement dit que rien ne permet de les différencier. Rejeter l'hypothèse nulle, c'est donc accepter l'existence d'une différence entre les deux échantillons analysés. La deuxième hypothèse, appelée l'hypothèse alternative ou  $H_1$ , est alors considérée. Celle-ci consiste à dire qu'il existe une différence entre ces échantillons [2,3,6–12].

Le but de la démarche statistique est donc de permettre une conclusion via les tests d'hypothèses formulées a priori à partir d'un échantillon [7,11].

L'action de rejeter ou d'accepter l'hypothèse nulle, et de considérer qu'il existe une différence ou non, est soumise à deux types d'erreur [12] :

- l'erreur  $\alpha$  (alpha) ou risque de première espèce ou encore de type I, se retrouve lorsque le clinicien considère qu'une différence existe entre les échantillons alors que celle-ci est le simple fruit du hasard ;
- l'erreur  $\beta$  (bêta) ou risque de deuxième espèce ou encore risque de type II, est le fait de conclure à tort qu'il n'existe pas de différence alors qu'en réalité, cette différence existe.

La plupart des études fixe le risque  $\alpha$  à 5 %, libre à chacun de diminuer ce risque pour augmenter le degré de chance de ne pas commettre cette erreur [13]. Le risque  $\beta$  quant à lui, est estimé arbitrairement à 20 %.

Cependant, une erreur est commise dans de nombreux articles par les auteurs en affirmant que l'hypothèse nulle est vraie en cas de résultats non significatifs. L' $H_0$  est acceptée selon le contexte de l'étude, mais en aucun cas elle est vraie [14–17].

*Les biostatistiques quantifient l'incertitude qui se retrouve obligatoirement lors de comparaison d'échantillons, mais ne quantifie en aucun cas un traitement ou une thérapeutique.*

La quantification d'une incertitude ne lève pas l'incertitude des décisions que le chercheur prend lorsqu'il accepte ou non l'hypothèse nulle [18].

## LE RÉSULTAT STATISTIQUE

### La valeur de $p$ -value

Depuis les années 1960, les études scientifiques utilisent presque exclusivement la  $p$ -value, mais elle fait l'objet d'une méconnaissance et d'une mauvaise utilisation [7,8,19,20]. Les erreurs fréquentes de définition de la  $p$ -value ainsi que sa dépendance à la taille de l'échantillon fragilisent et remettent en question son utilisation au sein de la logique statistique fréquentiste [14,16,21].

*La  $p$ -value est la probabilité qui quantifie le désaccord observé entre la thérapeutique utilisée et l'hypothèse nulle, de ce fait le choix d'une hypothèse par rapport à une autre se fait grâce à la  $p$ -value.*

Il convient d'appeler une différence statistique significative toute différence qui a, au minimum, cinq chances sur 100 de se produire simplement par hasard [22,23]. Le seuil de la signification statistique est fixé historiquement à 5 %, celui-ci peut être plus restrictif en fonction des études et correspond au risque

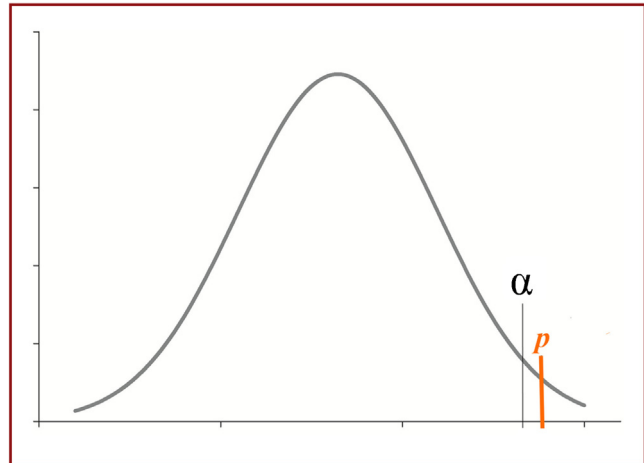


Figure 1. Courbe gaussienne avec une valeur du risque  $\alpha$  et une  $p$ -value lorsque la valeur de la  $p$ -value est inférieure à  $\alpha$ .

d'erreur  $\alpha$  [3,4,22,23]. De ce fait une  $p$ -value inférieure à 5 % rejette l' $H_0$  et l'hypothèse alternative est plus probable. Le risque  $\alpha$  est défini a priori et la  $p$ -value se calcule a posteriori à partir des observations recueillies (Fig. 1) [24].

*La  $p$ -value ne renseigne pas sur l'importance de la différence.*

En ce sens il n'y a pas de différence hautement significative ou de résultats significativement élevé, mais plutôt des petites  $p$ -values concluant à une différence significative ou des résultats significativement différents entre deux échantillons. C'est toute la différence entre des  $p$ -value de 0,04 et 0,003 [21,23,25]. Il est indispensable de connaître la valeur exacte de cette  $p$ -value afin de désacraliser ce seuil de  $p < 0,05$  [26–28]. Une  $p$ -value non significative ne prouve pas que l'étude soit sans résultat [29,30], sachant que la  $p$ -value est dépendante de la taille de l'échantillon, il sera difficile de trouver une différence statistiquement significative avec des groupes de petites tailles. Il faut toujours garder en tête que *the absence of evidence is not evidence of absence* [31].

### L'intervalle de confiance

Lorsqu'une  $p$ -value est égale à 0,048 ou à 0,052, quelle décision prendre quand les valeurs sont proches d'une signification statistique ?

*L'intervalle de confiance (IC) complète la  $p$ -value en donnant une information sur l'ensemble des valeurs compatibles avec une moyenne ou une proportion.*

L'IC<sub>95</sub> % est le plus cité dans les études, cet IC est alors une fourchette de valeurs où la véritable valeur possède 95 % de chances de s'y situer [28,32]. En fonction de la limite clinique et des valeurs statistiquement significatives qui ont été fixées, l'IC contient les valeurs qui sont cliniquement importantes. [26,33]. Concrètement, lorsque la différence entre deux thérapeutiques n'est pas statistiquement significative, l'amplitude de l'IC contient la valeur zéro [33–37]. Plus l'écart entre la valeur retrouvée et les bornes supérieures et inférieures de

## Biostatistiques : la dimension collective de nos individualités

L'IC<sub>95 %</sub> est faible, plus votre résultat est précis. Cette précision est dépendante du nombre de sujets présents dans l'échantillon, plus l'échantillon est important plus cette précision augmente [7].

*La p-value détermine donc la probabilité existante que l'effet observé soit seulement dû à la chance et l'IC indique l'écart des valeurs pour le réel effet dans une population.*

Cet IC est au cœur de la pratique basée sur des résultats précis en rapport avec la pratique et modère la portée des résultats de l'étude [4,27,37–41].

Au lieu d'annoncer que l'effet d'une prise en charge est statistiquement significatif, il est préférable d'annoncer les résultats en y intégrant l'IC : le renforcement musculaire en excentrique a augmenté la force de 41 N.m en moyenne, de 123 à 164 N.m (IC<sub>95 %</sub> = 27 à 55 N.m,  $p = 0,007$ ). Bien que de plus en plus recommandée par les différentes revues scientifiques, l'utilisation de l'IC peine à trouver sa place dans l'interprétation des résultats au sein des études cliniques [17].

### Taille de l'effet et puissance statistique

Énoncé en début d'article, des différences significatives sont retrouvées parce que la taille de l'échantillon est tout simplement grand. L'analyse de la puissance statistique (PS) et de la taille de l'effet (TE) permet de relativiser l'importance des tests d'hypothèses et donc de l'interprétation de la significativité statistique. Bien que ces termes commencent à être de plus en plus intégrés dans les articles scientifiques, il est encore rare de voir des discussions autour de la TE et de la PS [8,17,42].

La TE supplémente les tests d'hypothèses en donnant une indication sur la présence du phénomène dans une population via les valeurs de la distribution. De ce fait, plus la TE sera grande, plus l'écart avec l'H<sub>0</sub> sera important, et inversement. Il existe plusieurs TE en fonction des analyses statistiques, par exemple le «  $r$  » de Pearson peut constituer une estimation de la taille de l'effet. Dépendant de la variabilité des mesures et de l'hétérogénéité des distributions, une TE relative peut être calculée via des études cliniques précédentes [43,44].

L'oubli de traiter et d'interpréter les TE des différentes études engendre des interprétations contradictoires entre diverses études portant sur le même thème. Pour les échantillons de petites tailles, il faut des différences importantes pour retrouver une  $p$ -value inférieure à 0,05. Même si la TE donne une signification aux écarts retrouvés, celle-ci reste tout de même dépendante à la taille de l'échantillon [45].

La PS d'un test d'hypothèse est la probabilité qu'il rejette H<sub>0</sub> lorsque H<sub>0</sub> est fausse, elle est souvent annotée 1- $\beta$ . C'est donc sa capacité à déceler une différence significative quand il y en a réellement une [6,42]. La PS d'un test peut être comparée à celle d'une loupe : si un signe est perçu, son existence peut être affirmée. Dans le cas contraire, cela ne veut pas dire que ce signe n'existe pas, mais peut être serait-il perceptible avec une loupe plus puissante [22]. Celle-ci dépend principalement de trois paramètres : le seuil de signification, la taille de l'échantillon et la taille de l'effet. Ainsi, quand le seuil de signification est grand et que la taille de l'échantillon est important, la TE et la PS augmentent. Comme la TE, la PS est dépendante de l'analyse statistique [8,10,27,29,46].

De nombreuses études énoncent des différences via des conclusions hâtives en se basant sur des puissances

statistiques trop faibles car la taille des échantillons est trop petite, même si la  $p$ -value est inférieure à 0,05 [1,17]. La plupart du temps les logiciels récents apportent sur un plateau la puissance statistique ainsi que la taille de l'effet [17]. À la suite de ces constats, la notion du nombre nécessaire de sujets (*Number needed to treat* [NNT]) fait son apparition. Le NNT permet alors de se préserver d'études peu puissantes [22].

### Les biais

Dépendante du protocole mis en place, la relation causale entre un type de prise en charge et un effet observé est sujette à différents biais. Un biais est la distorsion entre l'estimation d'un résultat par rapport à la réalité, engendrant ainsi des erreurs dans les analyses statistiques de l'étude et des limites dans leur interprétation. Ces biais sont systématiquement présents dans la conception du protocole mais tendent à diminuer si l'étude est rigoureusement conçue et conduite [47–50]. Voici quelques exemples de biais facilement retrouvés dans la littérature scientifique.

#### Biais de sélection

Ils interviennent lors de la constitution des groupes. Les groupes testés sont-ils réellement comparables ? Parfois les sujets choisis ne sont pas si indépendants de l'effet testé. Afin de ne pas influencer sur la répartition des groupes, il est fortement conseillé d'homogénéiser les groupes comparés grâce à une randomisation : tirer au hasard la répartition des individus dans les groupes à comparer. Il est donc nécessaire de bien définir a priori les critères d'inclusion et d'exclusion [7,48,49,51].

#### Biais de confusion

Ils sont présents lorsque les protocoles de recherches ne possèdent pas de groupe témoin. Ce biais provoque une erreur d'appréciation entre les effets d'une prise en charge et les conséquences d'une pathologie ou d'un traitement médicamenteux [7,51].

#### Biais d'information

Ils apparaissent quand des données sont manquantes, et sont d'autant plus importants si ce manque de données apparaît plus dans un groupe que dans l'autre [7,50].

#### Biais de suivi ou de réalisation

Ils surviennent lorsqu'il existe des différences méthodologiques de prise en charge entre les groupes (fréquence des suivis, réalisation des tests, etc.). La mise en place d'une étude en double insu est un bon moyen pour éviter ce type de biais [7,48,51].

#### Biais d'évaluation

Particulièrement fréquent lorsque les critères de jugement s'appuient sur des mesures subjectives. Ce biais est constaté lorsque le critère de jugement n'est pas recherché de la même manière entre deux populations (encouragements lors d'un test de marche de six minutes, etc.) [7,48,51]. Les effets Hawthorne et Rosenthal sont également des biais qui, s'ils ne sont pas pris en compte, peuvent influencer l'interprétation des tests statistiques. L'effet Hawthorne apparaît

## Biostatistiques : la dimension collective de nos individualités

lorsque les sujets faisant partie de l'étude ne possèdent pas le même comportement, simplement parce qu'ils font partie d'une expérience [52]. L'effet Rosenthal quant à lui, intervient lorsque l'attitude de l'investigateur influence le sujet de façon non intentionnelle [53]. Et encore ici, les stratégies de simple ou de double insu peuvent être mise en place afin de diminuer les biais d'interprétation [54].

## CONCLUSION

Le raisonnement statistique peut paraître barbare à première lecture, c'est d'ailleurs pour cela qu'il existe de nombreuses interprétations erronées, mais une fois ingérée, sa logique est d'une rigueur inflexible. Comme dans un processus de recherche clinique, les biostatistiques répondent à des hypothèses posées en aval par une question. Ces réponses se retrouvent dans le résultat des analyses statistiques. Afin d'appréhender au mieux ce résultat la *p-value*, l'intervalle de confiance, la puissance statistique et la taille de l'effet sont des indices qui possèdent leur poids dans l'interprétation des résultats. Même si les analyses sont correctement réalisées et leurs interprétations justifiées, il reste néanmoins que celles-ci sont dépendantes de biais qu'il faut connaître afin qu'ils soient connus et si possible évités.

## Points à retenir

- L'H0 est acceptée selon le contexte de l'étude, mais en aucun cas elle est vraie.
- La *p-value* ne renseigne pas sur l'importance de la différence.
- L'absence de preuve d'efficacité n'est pas la preuve d'absence d'efficacité.

## Déclaration d'intérêts

L'auteur déclare ne pas avoir de conflits d'intérêts en relation avec cet article.

## Remerciements

Je remercie les membres du GERAR, Anne Ayala et Nathanaël Shahmaei, pour avoir pris le temps de lire, d'analyser, de vérifier et de critiquer ce travail.

## RÉFÉRENCES

- [1] Schwartz D, Lazar P, Papoz L. Statistique médicale et biologique, 5<sup>e</sup> éd., Flammarion Médecine-Sciences; 1993.
- [2] Ancelle T. Statistique épidémiologie, 3<sup>e</sup> éd., Maloine: Collection « sciences fondamentales »; 2007.
- [3] Falissard D. Comprendre et utiliser les statistiques dans les sciences de la vie, 3<sup>e</sup> éd., Masson; 2005.
- [4] Browner WS. Publishing and presenting clinical research, 2nd ed., Lippincott Williams & Wilkins; 2006.
- [5] Guyatt G, Jaeschke R, Heddle N, Cook D, Shannon H, Walter S. Basic statistics for clinicians: 1. Hypothesis testing. Can Med Assoc J 1995;152(1):27–32.
- [6] Glantz SA. Primer of biostatistics, 6th ed., McGraw-Hill Medical Publishing division; 2005.
- [7] Roy PM, Calvel L, Dubart AE, Jabre P, Ricard-Hibon A, Thys F, et al. Guide pratique de recherche clinique. S-Editions; 2008.
- [8] Bourque J, Blais JG, Larose F. L'interprétation des tests d'hypothèses : *p*, la taille de l'effet et la puissance. Rev Sci Educ 2009;35(1):211–26.
- [9] Meyer N, Vinzio S, Goichot B. La statistique bayésienne : une approche des statistiques adaptée à la clinique. Rev Med Interne 2009;39:242–9.
- [10] Pereira B, Thezenas S. Compréhension des tests statistiques. In: Kramar A, Mathoulin-Pélissier S, editors. Méthodes biostatistiques appliquées à la recherche clinique en cancérologie. John Libbey Eurotext; 2011.
- [11] Carrat F. Comparer deux groupes. Rev Med Interne 2000;21:118–21.
- [12] Huguier M, Flahault. Biostatistiques au quotidien. Elsevier; 2000.
- [13] Bewick V, Cheek L, Ball J. Statistics review 9: one-way analysis of variance. Crit Care 2004;8:130–6.
- [14] Lecoutre B, Poitevineau J. Aller au-delà des tests de signification traditionnels: vers de nouvelles normes de publication. Année Psychol 2000;100(4):683–713.
- [15] Lecoutre B, Poitevineau J, Lecoutre MP. A reason why not to ban null hypothesis significance tests. Rev MODULAD 2005;33:249–53.
- [16] Lecoutre B, Lecoutre MP, Poitevineau J. Uses, abuses and misuses of significance tests in the scientific community: won't the bayesian choice be unavoidable? Int Stat Rev;69(3):399–417.
- [17] Champely S, Verdot C. Que signifie la significativité statistique ? L'apport de la taille d'effet et de la puissance statistique. Staps 2007;77:49–61.
- [18] Duhaut P, Ducroix JP. Des chiffres et des lettres. Rev Med Interne 2005;26:970–2.
- [19] Goodman S. *P* value. In: Elston R, Olson J, Palmer, editors. Biostatistical genetics and genetic epidemiology. Wiley; 2002.
- [20] Sterne JAC, Davey Smith G. Sifting the evidence – what's wrong with significance tests? Br Med J 2001;322:226–31.
- [21] Grendár M. Is the *p*-value a good measure of evidence? Asymptotic consistency criteria. Stat Probability Lett 2012;82:1116–9.
- [22] Cucherat M. Interprétation des essais cliniques pour la pratique médicale. 2009. <http://www.spc.univ-lyon1.fr/polycop>.
- [23] Hill C, Laplanche A. Quand dit-on qu'une différence est statistiquement significative. In: Kramar A, Mathoulin-Pélissier S, editors. Méthodes biostatistiques appliquées à la recherche clinique en cancérologie. John Libbey Eurotext; 2011.
- [24] Riou B, Landais P. Principes des tests d'hypothèse en statistique :  $\alpha$ ,  $\beta$  et *P*. Ann Fr Anesth Reanim 1998;17:1168–80.
- [25] Gore SM. Assessing methods - Art of significance testing. Br Med J 1981;283:600–2.
- [26] Lang T. Twenty statistical errors even YOU can find in biomedical research articles. Croat Med J 2004;45:361–70.
- [27] Coutant C. Introduction à l'analyse critique statistique d'un article. Tribune des internes. Gynecol Obstet Fertil 2009;37:765–8.
- [28] Sim J, Reid N. Statistical inference by confidence intervals: issues of interpretation and utilization. Phys Ther 1999;79:186–95.
- [29] Maisonneuve H. Guide du thésard, 7<sup>e</sup> éd., Éditions Scientifiques L&C; 2010.
- [30] Perneger T, Perrier A. Analyses de variance (1<sup>re</sup> partie). Rev Mal Respir 2004;21:797–801.
- [31] Whitley E, Ball J. Statistics review 4: sample size calculations. Crit Care 2002;6:335–41.



## Biostatistiques : la dimension collective de nos individualités



- [32] Mélot C. Qu'est-ce qu'un intervalle de confiance ? *Rev Mal Respir* 2003;20:599–601.
- [33] Ada L. Laissez tomber le petit *p*. *Kinesither Rev* 2009;96:1.
- [34] Gore SM. Assessing methods – Confidence intervals. *Br Med J* 1981;283:660–2.
- [35] Gore SM. Assessing methods – Many variables. *Br Med J* 1981;283:901–5.
- [36] Poinso D. 2004. Statistiques pour statophobes. [en ligne : <http://perso.univ-rennes1.fr/denis.poinso>].
- [37] Montesi E. Petit lexique pour lecture critique. *Rev Med Gen* 2009;262:140–5.
- [38] Whithley E, Ball J. Statistics review 3: hypothesis testing and *P* values. *Crit Care* 2002;6:222–5.
- [39] Salmi LR. Lecture critique et rédaction médicale scientifique, comment lire, rédiger et publier une étude clinique ou épidémiologique. Elsevier; 1998.
- [40] Gardner MJ, Altman DG. Confidence intervals rather than *P* values: estimation rather than hypothesis testing. *Br Med J* 1986;292(6522):746–50.
- [41] Riffenburgh RH. *Statistics in medicine*, 2nd ed., Elsevier Academic Press; 2006.
- [42] Laurencelle L. Inventer ou estimer la puissance statistique ? Quelques considérations utiles pour le chercheur. *Tutorials Quant Methods Psychol* 2007;3(2):35–42.
- [43] Olejnik S, Algina J. Measures of effect size for comparative studies: applications, interpretations, and limitations. *Contemp Educ Psychol* 2000;25:241–86.
- [44] Durlak JA. How to select, calculate, and interpret effect sizes. *J Pediatr Psych* 2009;34(9):917–28.
- [45] Guéguen N. L'importance d'un effet : méthodologie simple de détermination et d'évaluation de « l'effect size ». *Eur J Sci Res* 2009;38(1):20–5.
- [46] Michel R, Ollivier-Gay L, Spiegel A, Boutin JP. Les tests statistiques : intérêt, principes et interprétations. *Med Trop* 2002;62:561–3.
- [47] Atkinson G, Nevill AM. Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Med* 1998;4:217–38.
- [48] Aptel F, Cucherat M, Blumen-Ohana E, Denis P. Interprétation des essais cliniques. *J Fr Ophtalmol* 2011;34:755–61.
- [49] Gail MH. Biais. In: Elston R, Olson J, Palmer, editors. *Biostatistical genetics and genetic epidemiology*. Wiley; 2002.
- [50] Shapiro S, Rosenberg L, Palmer LJ. Biais in case-control studies. In: Elston R, Olson J, Palmer, editors. *Biostatistical genetics and genetic epidemiology*. Wiley; 2002.
- [51] Riteau AS. Points-clés pour l'analyse d'un article scientifique : la recherche des biais dans un essai thérapeutique. *Tribune des internes. Gynecol Obstet Fertil* 2011;39:402–4.
- [52] Macefield R. Usability studies and the Hawthorne effect. *J Usabil Stud* 2007;2(3):145–54.
- [53] Triola MM, Triola MF. *Biostatistique pour les sciences de la vie et de la santé*. Pearson Educ 2009.
- [54] Gore SM. Assessing clinical trials – Double-blind trials. *Br Med J* 1981;283:122–4.