



**HAL**  
open science

## Méta-analyse basée sur la théorie des possibilités

Abdelhak Imoussaten, Meryeme Hassouna, Jacky Montmain, Gérard Dray

► **To cite this version:**

Abdelhak Imoussaten, Meryeme Hassouna, Jacky Montmain, Gérard Dray. Méta-analyse basée sur la théorie des possibilités. LFA 2023 - 32e rencontres francophones sur la logique floue et ses applications, Nov 2023, Bourges, France. hal-04202558

**HAL Id: hal-04202558**

**<https://hal.science/hal-04202558v1>**

Submitted on 11 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Méta-analyse basée sur la théorie des possibilités

## Possibility theory based meta-analysis

A. Imoussaten<sup>1</sup>

M. Hassouna<sup>1</sup>

J. Montmain<sup>1</sup>

G. Dray<sup>1</sup>

<sup>1</sup> EuroMov Digital Health in Motion, Univ Montpellier, IMT Mines Ales, Alès, France

### Résumé :

Les méta-analyses proposent des outils intéressants pour combiner les résultats d'études effectuées par plusieurs auteurs sur la force de l'effet observé d'une variable sur une autre, appelée aussi la taille d'effet. Les modèles proposés pour la méta-analyse le sont dans le cadre statistique ce qui permet de faire des hypothèses sur les lois de probabilités paramétriques rattachées à la taille d'effet. Cela permet aussi d'utiliser les outils classiques des statistiques que sont l'estimation, l'intervalle de confiance et les tests. Ce choix s'explique principalement parce que la théorie des probabilités est un cadre bien connu par les utilisateurs de la méta-analyse pour modéliser l'incertitude. Cependant, la théorie des probabilités montre ses limites pour combiner des informations provenant de sources hétérogènes. En l'absence d'hypothèse d'homogénéité, l'agrégation par une moyenne pondérée des tailles d'effet, adoptée dans les méta-analyses pour obtenir la taille d'effet finale, ne semble pas refléter l'information fournie par la plupart des sources. Dans cet article, nous proposons une méta-analyse qui tire profit des techniques de la théorie des possibilités pour combiner des informations incomplètes provenant de sources hétérogènes. Le résultat consiste à distinguer les valeurs plausibles et celles qui le sont moins pour la taille d'effet d'un traitement, par exemple, étant donnés les essais effectués par différents auteurs. Pour illustrer notre proposition, nous présentons une expérimentation sur un cas classique de la méta-analyse qui s'intéresse à l'effet de la prise de diurétiques pendant la grossesse.

**Mots-clés :** Méta-analyse, Combinaison des informations, Théorie des possibilités

### Abstract:

Meta-analyses offer interesting tools for combining the results of studies carried out by several authors on the strength of the observed effect of one variable on another, also called the effect size. The models proposed for the meta-analysis are within the statistical framework, which makes it possible to make hypotheses on the parametric probability laws related to the effect size variable. Then it is possible to use the classic statistical tools as estimation, confidence intervals and tests. This choice is mainly explained by the fact that probability theory is a well-known framework to model uncertainty for the meta-analysis users. However, probability theory shows its limits for combining information from heterogeneous sources. In the absence of the assumption of homogeneity, the aggregation by a weighted average of the effect sizes, adopted in the meta-analyses to obtain the fi-

nal effect size, does not seem to reflect the information provided by most sources. In this article, we propose a meta-analysis that takes advantage of possibility theory techniques to combine incomplete information from heterogeneous sources. The result is to distinguish between plausible and less plausible values for the effect size of a treatment, for example, given trials performed by different authors. To illustrate our proposal, we present an experiment on a classic case of meta-analysis which focuses on the effect of taking diuretics during pregnancy.

**Keywords:** Meta-analysis, Merging information, Possibility theory

## 1 Introduction

La méta analyse [9] est une approche statistique qui consiste à combiner les résultats empiriques d'études effectuées de manière indépendante et selon certains critères, i.e selon un protocole reproductible, sur la même question clinique pour en faire une synthèse statistique [16]. La synthèse consiste en la production d'une estimation agrégée la plus proche de la vérité commune inconnue. Elle a été appliquée le plus souvent pour l'estimation de l'effet d'un traitement suivant plusieurs essais cliniques randomisés [11] selon certaines hypothèses [20] [19]. En pratique, les méta-analyses sont conduites sur une liste d'articles et études qui représentent l'état de connaissances sur une question ou problème de recherche spécifique [16]. La force de la méta-analyse, et en même temps sa faiblesse, c'est qu'elle permet l'augmentation du nombre de cas étudiés, comparer aux études séparées, et d'en tirer une conclusion globale. L'objet d'étude est essentiellement la taille d'effet qui est une mesure de la force de l'effet observé d'une variable sur une autre [14].

Les tailles d'effet les plus répandues sont la différence standardisée entre deux moyennes et le rapport des cotes (odds ratio) qui compare la probabilité d'un événement dans deux groupes. Malgré les efforts fournis pour se rapprocher des hypothèses nécessaires sur les études qui composent une méta-analyse, malheureusement il subsiste encore divers problèmes méthodologiques. Si on laisse de côté les problèmes dus aux manipulations pour des raisons diverses des données, e.g., des études individuelles biaisées, qui peuvent avoir comme conséquence une estimation de l'effet global du traitement ne reflétant pas l'efficacité réelle d'un traitement [13], la méta-analyse peut être aussi critiquable pour deux raisons principales : 1) l'adoption de l'approche statistique basée sur la théorie des probabilités pour la représentation de l'incertitude et dont l'hypothèse d'homogénéité des essais est discutable; et 2) le choix de l'agrégation par moyenne pondérée des tailles d'effet d'études hétérogènes. L'agrégation est beaucoup plus adaptée lorsque l'on cherche un compromis entre les préférences d'un décideur sur plusieurs critères ou de plusieurs décideurs sur un même critère. Or, l'objectif d'une méta-analyse est de trouver une estimation globale, à partir des études, qui soit la plus proche possible de la véritable taille d'effet. Ainsi, la combinaison des informations provenant des études par des opérateurs logiques, nous semble plus appropriée pour la recherche de la vérité commune des études [7]. Dans ce sens, cet article, comme travail exploratoire, propose d'explorer les outils de représentations et de combinaisons logiques, tenant compte de la fiabilité des sources, issus de la théorie des possibilités. L'article est organisé de la façon suivante. Dans la Section 2 nous rappelons certains concepts élémentaires sur la méta-analyse et sur la théorie des possibilités. Puis, dans la Section 3, nous présentons l'approche proposée pour la méta-analyse basée sur la théorie des possibilités. Enfin, la Section 4 est dédiée à l'illustration et la comparaison des résultats de l'approche proposée sur des données classiques de la méta-analyse.

## 2 Rappels

### 2.1 La méta-analyse classique

Considérons  $K$  essais randomisés testant les effets d'un nouveau médicament. Dans chaque essai l'effet du traitement est estimé en termes de moyenne ou de différence de moyenne ou de rapport des cotes logarithmique si l'effet est étudié par rapport à un groupe témoins. Ce dernier cas est à l'origine de l'appellation "effect size", en référence à la force de l'écart entre la moyenne de l'effet chez les patients traités et celle chez les patients témoins. Dans une méta-analyse, le modèle à effets fixes considère l'hypothèse de l'homogénéité des effets du traitement dans toutes les  $K$  études. En d'autres termes, pour chaque étude  $k$ , l'effet estimé  $\hat{\theta}_k$  du traitement a une distribution de moyenne commune  $\theta$  et de variance individuelle  $v_k$  pour  $k = 1, \dots, K$ . La moyenne  $\theta$  est alors estimée, dans le cas du modèle fixe, comme une moyenne pondérée des estimations des effets dans chaque étude, c'est-à-dire,

$$\hat{\theta} = \frac{\sum_{k=1}^K w_k \cdot \hat{\theta}_k}{\sum_{k=1}^K w_k} \quad (1)$$

où  $w_k$  est le poids donné à l'étude  $k$ , généralement considéré comme l'inverse de la variance  $v_k$  [21]. Si de plus,  $\hat{\theta}$  peut être considéré comme suivant une loi normale, par exemple c'est le cas quand chaque  $\hat{\theta}_k$  suit approximativement une loi normale. Un intervalle de confiance pour  $\theta$  au niveau  $100 \cdot (1 - \alpha)\%$  peut alors être calculé comme suit [3] :

$$IC_{\alpha}(\theta) = \hat{\theta} \pm z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{\theta})}, \quad (2)$$

où  $\text{var}(\hat{\theta}) = 1 / \sum_{k=1}^K w_k$  et  $z_{1-\frac{\alpha}{2}}$  est le quantile d'ordre  $1 - \frac{\alpha}{2}$  qui provient des tables de la loi normale centrée réduite. L'estimation de la formule (1) et l'intervalle de confiance de la formule (2) tiennent aussi dans le cas

du modèle à effets aléatoires [3] où une hétérogénéité des effets du traitement entre les études est considérée. Ce dernier modèle fournit un moyen d'incorporer l'hétérogénéité dans l'estimation en incluant une composante inter-études de variance  $\sigma_B^2$ , où l'hétérogénéité est restreinte à une forme particulière. En effet, pour chaque étude  $k$  l'on considère  $\hat{\theta}_k \sim \mathcal{N}(\theta_k, v_k)$  où  $\theta_k \sim \mathcal{N}(\theta, \sigma_B^2)$ . C'est comme si l'on considère que chaque essai estime un effet légèrement différent, et que cet effet varie avec différents critères d'entrée, différents protocoles de traitement et différents facteurs environnementaux et génétiques etc. chez les patients de différents essais. L'estimation de l'effet du traitement de la formule (1) est à nouveau obtenue avec des poids ajustés pour incorporer  $\sigma_B^2$ , i.e.  $w_k^* = \frac{1}{(v_k + \sigma_B^2)}$  [3]. Un cas particulier de grandeur fréquemment utilisé en épidémiologie pour évaluer l'effet d'un traitement est celui du rapport des cotes. Il permet d'étudier l'effet observé d'une variable binaire sur une autre variable du même type. Plus précisément, cette grandeur se définit comme le rapport de la cote d'un événement, e.g. une maladie, arrivant à un **groupe traité** ou expérimentale, avec celle du même événement arrivant à un **groupe témoin**, i.e., groupe de participants qui ne sont pas soumis au traitement. Par exemple, supposons que deux groupes traités et témoins soient étudiés pour évaluer l'exposition à un facteur causal suspecté. Les données d'une telle étude sont représentées comme dans la table 1. Une esti-

Tableau 1 – Exemple d'expérimentation

	Nb. exposés / Nb. total	
	Traités	Témoins
Essai	$a/n_1$	$b/n_2$

mation approximative du risque relatif de la maladie associée à l'exposition au facteur peut être obtenue à partir de cette étude par l'estimation du rapport des cotes, noté  $OR$  [17] [1] :

$$OR = \frac{a}{n_1 - a} / \frac{b}{n_2 - b} = \frac{a \cdot (n_2 - b)}{(n_1 - a) \cdot c} \quad (3)$$

Un intervalle de confiance pour la grandeur  $OR$  peut être construit à l'aide de plusieurs méthodes qui varient dans leur précision. Parmi ces méthodes, la méthode *logit* [21] est largement recommandée comme approximation satisfaisante basée sur l'approximation par la loi normale de la distribution du logarithme du rapport des cotes  $\log(OR)$  pour laquelle l'écart-type est :

$$\sigma(\log(OR)) = \sqrt{\frac{1}{a} + \frac{1}{n_1 - a} + \frac{1}{b} + \frac{1}{n_2 - b}} \quad (4)$$

Un intervalle de confiance de  $100 \cdot (1 - \alpha)\%$  pour  $OR$  est trouvé à partir de celui de  $\log(OR)$  :

$$IC_\alpha(OR) = e^{\log(OR) \pm z_{1-\frac{\alpha}{2}} \cdot \sigma(\log(OR))} \quad (5)$$

Les formules (3) et (5) donnent les estimations pour un essai individuel. Par la suite, les formules (1) et (2) sont utilisées pour l'estimation et l'intervalle de confiance du rapport des cotes à partir d'un ensemble d'essais. Par ailleurs, les poids dans la moyenne pondérée de la formule (1) peuvent être vus comme des affaiblissements de chacune des sources d'information sur la vraie valeur de  $\theta$ . C'est l'information qu'un essai individuel contient sur sa propre moyenne. En quelque sorte, plus la variance est grande moins l'essai est considéré fiable. la présence d'hétérogénéité entre les essais nécessite d'affaiblir certaines sources dans le cadre des méta-analyses, plus particulièrement dans le contexte d'essais cliniques où elle est considérée comme une alternative à la solution qui consiste à éliminer des sources de moins bonne qualité [10]. Les auteurs de [10] suggèrent un affaiblissement basé sur l'expérience clinique plutôt que celui qui remplit certains critères statistiques.

## 2.2 Théorie des possibilités

**Représentation de l'information.** La théorie des possibilités est un cadre de représentation de l'incertitude et de l'imprécision pour permettre de distinguer les valeurs plausibles et celles qui sont peu plausibles pour un paramètre ou une variable d'intérêt [4] [8]. Plus précisément,

soit  $\Theta$  un intervalle réel fermé et borné contenant la vraie valeur inconnue  $\theta^*$  d'une quantité d'intérêt. Les valeurs de  $\Theta$  sont considérées mutuellement exclusives, i.e.,  $\theta^*$  ne peut avoir qu'une seule valeur (sa vraie valeur). Une distribution de possibilité définie sur  $\Theta$ ,  $\pi_{\theta^*} : \Theta \rightarrow [0, 1]$ , est une fonction réelle telle que :  $\pi_{\theta^*}$  associe à chaque élément  $\theta \in \Theta$  un degré de possibilité  $\pi_{\theta^*}(\theta)$  pour être la vraie valeur  $\theta^*$ . Ainsi,  $\pi_{\theta^*}$  quantifie à quel point une valeur de  $\Theta$  est plus ou moins plausible pour être  $\theta^*$ . On suppose qu'il existe au moins une valeur  $\theta \in \Theta$  telle que  $\pi_{\theta^*}(\theta) = 1$  car il est certain que la vraie valeur  $\theta^*$  se trouve dans  $\Theta$ . Lorsque pour deux différentes valeurs  $\theta_1$  et  $\theta_2$  de  $\Theta$ ,  $\pi_{\theta^*}(\theta_1) > \pi_{\theta^*}(\theta_2)$  cela signifie que  $\theta_1$  est considérée plus plausible que  $\theta_2$ . Il est certain que  $\theta^*$  ne prend pas une valeur  $\theta \in \Theta$  si  $\pi_{\theta^*}(\theta) = 0$ . A partir de  $\pi_{\theta^*}$ , la mesure de possibilité, noté  $\Pi$  et la mesure de nécessité, noté  $N$ , sont définies pour tout événement  $A \subseteq \Theta$ ,

$$\Pi(A) = \sup_{\theta \in A} \pi_{\theta^*}(\theta), \quad (6)$$

$$N(A) = 1 - \Pi(A^c). \quad (7)$$

où  $A^c$  représente le complémentaire de  $A$  dans  $\Theta$ . La distribution de possibilité  $\pi_{\theta^*}$  peut représentée une famille de sous-ensembles de confiance imbriqués et de nombre fini  $A_1 \subset A_2 \subset \dots \subset A_m$  où chaque  $A_i$  est attaché à un niveau de confiance positif  $\lambda_i$  tel que  $\lambda_i = N(A_i)$  [6]. La distribution représentant ces sous-ensembles, respectant le principe de l'engagement minimal, s'écrit [7],  $\forall \theta \in \Theta$ ,

$$\pi_{\theta^*}(\theta) = \min_{i=1,m} \max(\mu_{A_i}(\theta), 1 - \lambda_i), \quad (8)$$

où  $\mu_{A_i}$  est la fonction caractéristique de support  $A_i$ . Les auteurs de [18] suggèrent d'utiliser trois intervalles avec des niveaux de confiance prédéfinis :  $A_1$  avec  $\lambda_1 = 0.05$ ,  $A_2$  avec  $\lambda_2 = 0.5$ ,  $A_3$  avec  $\lambda_3 = 0.95$  en plus de  $A_4$  (avec  $\lambda_4 = 1$ ) qui est toujours  $\Theta$ . Par ailleurs, si le degré de certitude qu'une source donnée soit fiable est connu, disons  $w$ , alors il est possible de tenir compte de cette information en affaiblissant une distribution de possibilité  $\pi$  en

$$\pi' = \max(\pi, 1 - w). \quad (9)$$

Si  $w = 1$  (source fiable), alors  $\pi' = \pi$  et si  $w = 0$  (source non fiable), alors  $\pi' = 1$ .

#### Combinaison de distributions de possibilités [7].

L'approche possibiliste permet la combinaison d'informations imprécises provenant de sources hétérogènes. Ses principaux avantages par rapport aux théories classiques sont : la fidélité de la représentation des informations, l'absence de besoin de connaissances a priori, et une variété de méthodes de combinaison dont le choix dépend de la fiabilité des experts ou des sources et du niveau de conflits entre les opinions ou les informations fournies [7]. Les combinaisons logiques comme les combinaisons conjonctives notée  $\pi_{\wedge}$  ou  $\pi_*$ , qui s'appliquent lorsque toutes les sources sont fiables, et les combinaisons disjonctives, notées  $\pi_{\vee}$ , qui traitent le cas de sources non fiables cachées dans un groupe d'autres sources fiables, constituent les combinaisons principales en théorie des possibilités. En présence de conflits et/ou de sources non-fiables, des combinaisons logiques pondérées peuvent être envisagées. Les combinaisons conjonctives ont un sens si tous les  $\pi_k$  se chevauchent de manière significative, c'est-à-dire  $\exists \theta, \forall k, \pi_k(\theta) = 1$ , exprimant qu'il existe au moins une valeur de  $\Theta$  que toutes les sources considèrent comme tout à fait possible. Cependant, en cas de conflit, on peut normaliser  $\pi_{\wedge}$  tout en gardant trace d'un conflit partiel, la règle de combinaison conjonctive normalisée et affaiblie suivante a été proposée [7] [5] :  $\forall \theta \in \Theta$ ,

$$\pi_{AD}(\theta) = \max\left(\frac{\pi_{\wedge}(\theta)}{h_{\wedge}}, \min(\pi_{\vee}(\theta), 1 - h_{\wedge})\right) \quad (10)$$

où  $h_{\wedge} = \sup_{\theta \in \Theta} \min_{k \in 1,K} \pi_k(\theta)$  est un indice de consistance. La règle de combinaison dans l'équation (10) est une extension à plusieurs sources de celle définie dans [5] pour deux sources. Cette extension est suggérée avec précaution dans [5]. En effet, elle peut ne pas être efficace si les sources sont éparpillées car l'indice de consistance sera nul et c'est la combinaison disjonctive non-informative qui sera le résultat.

### 3 La proposition d'une approche possibiliste pour la méta-analyse

Notre approche se base sur le constat que dans le cas de la méta-analyse, où l'on cherche la vraie valeur de la taille de l'effet par combinaison des résultats de plusieurs essais, nous sommes plus proches de l'hypothèse de plusieurs capteurs différents qui sont utilisés pour mesurer la même quantité que de l'hypothèse de plusieurs mesures de cette même quantité par un même capteur, comme le résultat d'une expérimentation aléatoire, qui justifierait l'application d'outils statistiques. Ainsi, en dépit de tous les efforts pour les rendre homogènes, l'hétérogénéité existe entre les essais et elle ne peut pas être simplifiée par un modèle à effet aléatoire. Par conséquent, l'estimation par la moyenne, i.e. les données à fusionner peuvent être interprétées comme une statistique standard, est non seulement inadéquate dans certains cas mais elle peut, dans la plus part des cas, être éloignée des estimations apportées par chacune des sources sur la vraie valeur recherchée de la taille de l'effet. De façon générale, on peut considérer les informations apportées par un essai comme une expertise des auteurs de l'article dont elles sont issues et étant que telles, ces informations peuvent être contradictoires et/ou en conflits avec les autres informations apportées par d'autres sources. De plus, ces expertises peuvent être plus au moins fiables selon la date de l'étude, la taille de l'échantillon utilisé, et d'autres éléments subjectifs d'intérêts. Comme alternative à l'approche statistique pour traiter ce problème, la théorie des possibilités va chercher à déterminer les valeurs qui sont plausibles et celles qui le sont moins en fonction de valeurs sur lesquelles s'accordent la plus part des sources. De cette façon elle offre une réponse prudente qui correspond aux données disponibles. L'approche proposée est alors organisée de la façon suivante. Considérons  $K$  essais testant les effets d'un nouveau médicament. Dans chaque essai la taille de l'effet du traitement est estimée et des intervalles de confiances peuvent être cal-

culés.

- Chaque source  $k$ , e.g., fournit une famille de quatre intervalles de confiance imbriqués  $A_1^k$ ,  $A_2^k$ ,  $A_3^k$  et  $A_4^k$ . Dans le cas présent, les intervalles sont calculés selon la formule (2) ou la formule (5), en fonction de la taille d'effet concerné. Les quatre intervalles correspondent au niveau de confiance 0.05, 0.5, 0.95 et 1.
- Les informations apportées par la source  $k$  sont donc représentées par une distribution de possibilité  $\pi_k$  calculées à partir des intervalles de confiance selon la formule (8).
- Si l'information sur la fiabilité de chacune des sources est connue, chaque distribution  $\pi_k$  est affaiblie selon la formule (9).
- Finalement, les sources sont combinées selon la formule générale (10). En effet, il nous semble que cette combinaison est adéquate car il n'y a aucune raison de ne pas considérer que toutes les sources disent vrai ou du moins ne soient pas éparses.

## 4 Expérimentations

### 4.1 La méta-analyse classique

Dans cette section, nous présentons un exemple classique de la littérature [2] sur l'utilisation de la méta-analyse. Il s'agit d'une méta-analyse de neuf essais cliniques portant sur l'effet de la prise de diurétiques pendant la grossesse sur le risque de pré-éclampsie [2]. Les données sont résumées dans la Table 2. Le résultat de l'analyse avec le modèle à effet aléatoire de DerSimonian-Laird [3] est présentée sur la Figure 1. A noter que la taille de l'effet du traitement est mesurée en termes de rapport des cotes logarithmique  $\theta$ , et les résultats sont donc rapportés ici en termes de rapport des cotes  $e^\theta$ . Comme on peut le voir sur la figure de la méta-analyse, l'estimation fournie par le modèle à effet aléatoire, i.e 0.6, ne correspond à aucune des estimations fournies par les sources. Plus

Tableau 2 – Neuf essais sur les effets de la diurétiques sur la pré-éclampsie[2]

Essai (année de publication)	Cas de pré-éclampsie/Nombre de patients			
	Traités		Témoins	
Weseley & Douglas 1962	14/131	(10.7%)	14/136	(10.3%)
Flowers et al. 1962	21/385	(5.5%)	17/134	(12.7%)
Menzies 1964	14/57	(24.6%)	24/48	(50%)
Fallis et al. 1964	6/38	(15.8%)	18/40	(45%)
Cuadros & Tatum 1964	12/1011	(1.2%)	35/760	(4.6%)
Landesman et al. 1965	138/1370	(10.1%)	175/1336	(13.1%)
Kraus et al. 1966	15/506	(3%)	20/524	(3.8%)
Tervila & Vartiainen 1971	6/108	(5.6%)	2/103	(1.9%)
Campbell & MacGillivray 1975	65/153	(42.5%)	40/102	(39.2%)

marquant, cette valeur n'appartient pas à l'intervalle de confiance au niveau 95% de trois essais. Plusieurs autres modèles ont été proposées

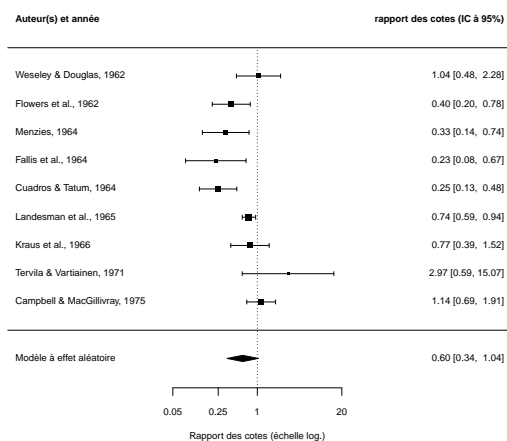


Figure 1 – La méta-analyse des données pré-éclampsie avec le package *metafor* de R

pour améliorer les résultat des modèles à effet fixe et aléatoire, le modèle Knapp-Hartung [15], le modèle de vraisemblance de profil [11] et le modèle bayésien hiérarchique [12]. Les figures 3 et 4 montrent les estimations et les intervalles de confiance fournis par ces méthodes qui sont très proches de ceux du modèle à effet aléatoire.

## 4.2 La méta-analyse basée sur la théorie des possibilités

Dans cette sous-section, nous allons appliquer les quatre étapes de l'approche possibiliste pro-

posée dans la Section 3. La Table 3, présente la première étape qui consiste à fournir les quatre intervalles de confiance pour chacune des sources. A partir de ces intervalles de confiance et sans affaiblissement, la Figure 2 montre les neuf distributions de possibilités résultats de la deuxième étape. Pour le be-

Tableau 3 – Intervalles de confiance

Essai	Rapport des cotes	Intervalle de confiance associés aux essais				Degré de Fiabilité
		$\lambda = 1$	$\lambda = 0.95$	$\lambda = 0.5$	$\lambda = 0.05$	
Weseley & Douglas 1962	1.04	[0.14, 7.69]	[0.48, 2.28]	[0.79, 1.36]	[1.01, 1.07]	0.86
Flowers et al. 1962	0.4	[0.07, 2.21]	[0.2, 0.78]	[0.34, 0.54]	[0.42, 0.44]	0.79
Menzies 1964	0.33	[0.04, 2.68]	[0.14, 0.74]	[0.38, 0.64]	[0.48, 0.50]	0.86
Fallis et al. 1964	0.23	[0.01, 3.53]	[0.08, 0.67]	[0.25, 0.50]	[0.34, 0.36]	0.86
Cuadros & Tatum 1964	0.25	[0.05, 1.35]	[0.13, 0.48]	[0.20, 0.32]	[0.25, 0.26]	0.86
Landesman et al. 1965	0.74	[0.41, 1.36]	[0.59, 0.94]	[0.71, 0.83]	[0.76, 0.78]	0.99
Kraus et al. 1966	0.77	[0.14, 4.37]	[0.39, 1.52]	[0.61, 0.98]	[0.76, 0.79]	1.0
Tervila & Vartiainen 1971	2.97	[0.05, 187.01]	[0.59, 15.07]	[1.63, 5.02]	[2.71, 3.02]	0.97
Campbell & MacGillivray 1975	1.14	[0.31, 4.21]	[0.69, 1.91]	[0.92, 1.27]	[1.07, 1.10]	0.9

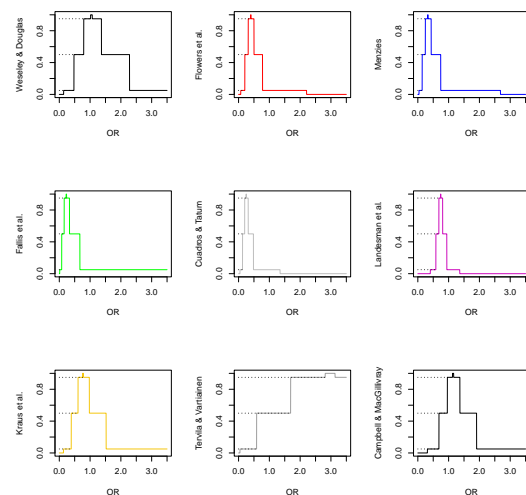


Figure 2 – Les distributions de possibilités des essais

soin de l'illustration et de façon subjective, nous avons calculé des degrés de fiabilité des informations fournies par les essais en fonction de plusieurs critères. Le premier critère est l'année de réalisation de l'essai, i.e., plus les années passent, plus les essais sont adaptés et ils deviennent de plus en plus homogènes et donc plus fiables. Le deuxième critère est la taille de l'échantillon et le troisième critère est l'équilibre entre les OR patients traités et les patients témoins. Les degrés de fiabilité sont présenté dans la Table 3 où l'on voit que les

sixième, septième et huitième essais obtiennent les plus grands degrés de fiabilité selon ces critères. La dernière étape consiste à calculer la combinaison des neuf distributions de possibilités en utilisant la formule (10). Dans cette illustration, nous considérons les opérateurs  $\wedge = \min$  et  $\vee = \max$ . La Figure 3 montre la distribution résultat de cette combinaison. Pour cette combinaison, on obtient un indice de

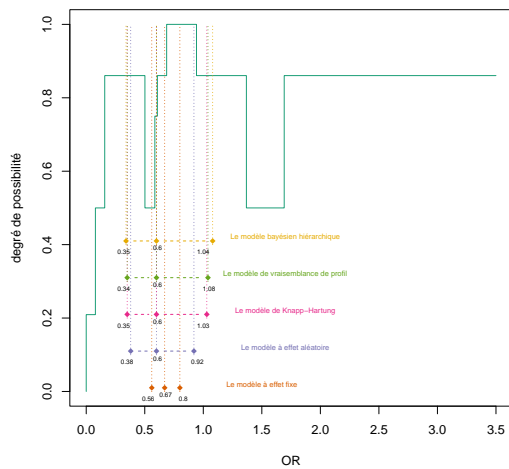


Figure 3 – Le résultat de la combinaison.

consistance de  $h_{\wedge} = 0.14$  qui n'est pas élevé ce qui traduit l'inconsistance entre les informations fournies par les sources. De plus, on peut constater qu'il y a trois modes ce qui signifie trois "avis" différents sur la vraie valeur de rapport des cotes. Le premier mode et le troisième mode résultent d'un fort conflit entre les sources dans les intervalles concernées ce qui se traduit par un degré de possibilité égal à  $1 - h_{\wedge} = 0.86$ . Par contre, le deuxième mode semble représenter les valeurs totalement plausibles sur lesquelles s'accordent la plus part des sources, i.e., l'intervalle  $[0.69, 0.94]$ . On remarque que toutes les estimations données par les autres modèles probabilistes ne tombent pas dans cet intervalle. En revanche, cet intervalle contient les parties supérieures des intervalles de confiance à 95% des autres modèles. A noter que la distribution de possibilité représentant l'essai "Tervila & Vartiainen", voir Figure 2

semble s'écarter des résultats des autres essais. En considérant cet essai non fiable, nous obtenons dans la Figure 4 une combinaison avec seulement un seul mode qui correspond au mode central de la Figure 3.

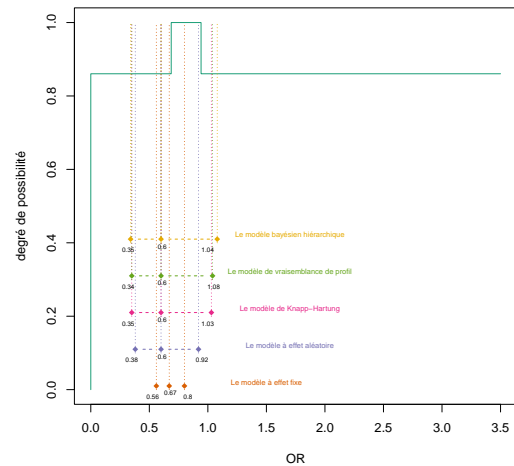


Figure 4 – Le résultat de la combinaison en considérant l'essai "Tervila & Vartiainen" comme non fiable.

## 5 Conclusion

La méta-analyse prend tout son sens récemment avec l'apparition du "covid" et le "covid long" pour lesquels plusieurs études ont été menées dans plusieurs laboratoires et dans plusieurs pays et cet outil semble adéquat pour résumer les connaissances sur ces deux "maladies". Ce premier travail exploratoire sur la méta-analyse possibiliste s'inscrit dans un projet qui vise à répondre à plusieurs questions sur le "covid long" en se basant sur tous les travaux et les données publiées récemment sur cette maladie. En effet, plusieurs symptômes post-covid apparaissent chez les atteints de covid et l'on voudrait les étudier pour distinguer ceux qui sont vraiment liés au covid-long, pour vérifier les dépendances entre eux et pour savoir pourquoi certains patients sont atteints par certains symptômes et pas d'autres.



## Références

- [1] Peter Armitage, Geoffrey Berry, and John Nigel Scott Matthews. *Statistical methods in medical research*. John Wiley & Sons, 2008.
- [2] Rory Collins, Salim Yusuf, and Richard Peto. Overview of randomised trials of diuretics in pregnancy. *Br Med J (Clin Res Ed)*, 290(6461) :17–23, 1985.
- [3] Rebecca DerSimonian and Nan Laird. Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3) :177–188, 1986.
- [4] Didier Dubois and Henri Prade. Representation and combination of uncertainty with belief functions and possibility measures. *Computational intelligence*, 4(3) :244–264, 1988.
- [5] Didier Dubois and Henri Prade. Combination of fuzzy information in the framework of possibility theory. *Data fusion in robotics and machine intelligence*, 12 :481–505, 1992.
- [6] Didier Dubois and Henri Prade. When upper probabilities are possibility measures. *FSS*, 49(1) :65–74, 1992.
- [7] Didier Dubois and Henri Prade. Possibility theory and data fusion in poorly informed environments. *Control Engineering Practice*, 2(5) :811–823, 1994.
- [8] Didier Dubois and Henri Prade. *Possibility theory : an approach to computerized processing of uncertainty*. Springer Science & Business Media, 2012.
- [9] Gene V Glass. Primary, secondary, and meta-analysis of research. *Educational researcher*, 5(10) :3–8, 1976.
- [10] Steven N Goodman. Meta-analysis and evidence. *Controlled Clinical Trials*, 10(2) :188–204, 1989.
- [11] Rebecca J Hardy and Simon G Thompson. A likelihood approach to meta-analysis with random effects. *Statistics in medicine*, 15(6) :619–629, 1996.
- [12] Julian PT Higgins, Simon G Thompson, and David J Spiegelhalter. A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society : Series A (Statistics in Society)*, 172(1) :137–159, 2009.
- [13] John E Hunter and Frank L Schmidt. *Methods of meta-analysis : Correcting error and bias in research findings*. Sage, 2004.
- [14] Ken Kelley and Kristopher J Preacher. On effect size. *Psychological methods*, 17(2) :137, 2012.
- [15] Guido Knapp and Joachim Hartung. Improved tests for a random effects meta-regression with a single covariate. *Statistics in medicine*, 22(17) :2693–2710, 2003.
- [16] Patrice Laroche. *La méta-analyse : Méthodes et applications en sciences sociales*. De Boeck, 2015.
- [17] Julie A Morris and Martin J Gardner. Statistics in medicine : Calculating confidence intervals for relative risks (odds ratios) and standardised ratios and rates. *British medical journal (Clinical research ed.)*, 296(6632) :1313, 1988.
- [18] Sandra A Sandri, Didier Dubois, and Henk W Kalfsbeek. Elicitation, assessment, and pooling of expert judgments using possibility theory. *IEEE transactions on fuzzy systems*, 3(3) :313–335, 1995.
- [19] Simon G Thompson. Controversies in meta-analysis : the case of the trials of serum cholesterol reduction. *Statistical methods in medical research*, 2(2) :173–192, 1993.
- [20] Anne Whitehead and John Whitehead. A general parametric approach to the meta-analysis of randomized clinical trials. *Statistics in medicine*, 10(11) :1665–1677, 1991.
- [21] Barnet Woolf et al. On estimating the relation between blood group and disease. *Ann hum genet*, 19(4) :251–253, 1955.