

Interactions et modifications d'effet en Epidémiologie

CERPOP, INSERM, EQUITY Team

Last compiled on 31 mai, 2023

Contents

1	Présentation	7
2	Introduction	9
2.1	Quand étudier les interactions ?	9
2.2	Les points les plus importants	11
I	Synthèse de la littérature	13
3	Notations	15
3.1	Variables et probabilités	15
3.2	Mesures d'effets	16
4	Interaction vs modification d'effets	19
4.1	Modification d'effets	19
4.2	Interaction	20
4.3	Synthèse	22
5	La question des échelles	23
5.1	Mesures des interactions	23
5.2	Lien entre les deux échelles	24
5.3	Synthèse	27

6	Types de paramètres	29
6.1	Avec les différences de risques (DR)	29
6.2	Avec les risques relatifs (RR)	29
6.3	Avec les Odds Ratio (OR)	29
6.4	Excès de risque à partir des RR (RERI)	30
6.5	Autres	30
II	Estimations, Interprétations, Présentations	33
7	Présentation des résultats	35
7.1	Recommandations	35
7.2	Proposition	36
8	Simulations	39
9	A partir de modèles de régression	43
9.1	Régression logistique	43
9.2	Régression lineaire	45
10	Analyses confirmatoires	49
10.1	Estimation par G-computation	49
10.2	Estimation par Modèle Structurel Marginal	55
10.3	Estimation avec TMLE	57
11	Représentations graphiques	75
III	En pratique	77
12	Proposition d'étapes	79
13	Exemple 1 - Y binaire	81

<i>CONTENTS</i>	5
14 Exemple 2 - Y quantitatif	83
14.1 1. Formuler les objectifs	83
14.2 2. Stratégies et méthodes	83
14.3 3. Analyse descriptive	84
14.4 4. Analyse exploratoire	84
14.5 5. Analyse confirmatoire	86
15 Exemple 3 - Y multinomial	87
16 Exemple 4 - X quantitatif	89
IV Conclusion	91
17 Synthèse générale	93
18 Pour aller plus loin...	95
18.1 Ajouter de la complexité	95
18.2 Interaction avec confusion intermédiaire	95
18.3 Interaction et médiation	95
19 Références	97

Chapter 1

Présentation

Ce document a été rédigé en tant que document de synthèse du travail du groupe “Interaction” de l’équipe EQUITY, CERPOP. Ce travail a consisté en une revue de la littérature et en une application détaillée des méthodes sur des analyses illustratives, dans un but d’auto-formation et pédagogique.

Les participant.e.s du groupe de travail sont :

- Hélène COLINEAUX
- Léna BONIN
- Camille JOANNES
- Benoit LEPAGE
- Lola NEUFCOURT
- Ainhoa UGARTECHE



The online version of this book is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Chapter 2

Introduction

Comment telle prédisposition génétique et telle exposition environnementale *inter-agissent*-elles ? L'effet de tel traitement varie-t-il selon les circonstances ? Selon les caractéristiques du patient ? Telle intervention peut-elle être bénéfique pour un groupe social et délétère pour un autre ?

De nombreuses questions épidémiologiques impliquent des mécanismes d'interactions ou de modifications d'effet. Pourtant, étudier ces mécanismes restent encore complexe aujourd'hui sur le plan méthodologique : quelle démarche adopter ? sur quelle échelle mesurer cette interaction ? comment interpréter les coefficients ? et cetera.

Dans ce document, nous proposons une synthèse de la littérature et une démarche progressive et appliquée pour explorer ces questions.

2.1 Quand étudier les interactions ?

2.1.1 *Prediction* versus *causalité*

La science des données cherche à répondre à 3 types d'objectifs Hernán et al. [2019] :

Description	Prédiction	Inférence causale
Résumer, décrire, visualiser	Reconnaissance des schémas et prévision	Compréhension
<u>Axé sur les données</u> : calculs simples +/- apprentissage non supervisé	<u>Axé sur les données</u> : modélisation statistique +/- apprentissage supervisé	<u>Non uniquement axé sur les données</u> : implique la combinaison de connaissances externes avec la modélisation statistique +/- apprentissage supervisé
Objectif : synthétiser l'information	Objectif : Prédire la valeur de l'outcome	Objectif : Estimer un effet causal

Selon le type d'objectif, la démarche d'analyse et les enjeux méthodologiques ne vont pas être les mêmes. Si l'objectif est prédictif, la démarche va être centrée sur la *prédiction de l'outcome*, à partir de covariables sélectionnées afin d'optimiser les performances de la prédiction, tout en prenant en compte leur disponibilité en pratique et la parcimonie du modèle.

Dans une démarche explicative, ou *étiologique*, au contraire, la démarche va être centrée sur l'*estimation d'un effet causal*, en prenant en compte les covariables en fonction de leur rôle vis-à-vis de l'effet d'intérêt (facteurs de confusion, colliders, médiateurs...).

En épidémiologie, à l'exception des cas où l'on souhaite développer un test ou score diagnostique ou pronostique, les objectifs sont le plus souvent explicatifs. On cherche en effet, la plupart du temps, à identifier des liens de cause à effet, afin de pouvoir agir sur les causes pour modifier les effets.

Finalement, pour répondre à la question “quand doit-on prendre en compte les interactions ?”, il est d'abord nécessaire d'identifier dans quel type de démarche l'on s'inscrit :

- **Démarche prédictive** : on ajoutera alors les interactions dans le modèle de prédiction, pour le rendre plus *flexible*, si cela améliore les performances de la prédiction VanderWeele and Knol [2014].
- **Démarche explicative/étiologique** : on étudiera les interactions ou modifications d'effet, si cela répond directement à l'objectif. Par exemple :
 - Si l'objectif est du type “l'effet de X sur Y varie-t-il en fonction de V ?”, on prendra en compte l'interaction entre X et V .
 - Les objectifs qui nécessitent la prise en compte de l'interaction peuvent aussi être du type : “Quel est l'effet conjoint de X et V sur Y ?” ou “Quel part de l'effet de X sur Y disparaît quand V est modifié ?”, etc.
 - Par contre, si l'objectif est simplement d'estimer l'effet de X sur Y , ou l'effet médié par un médiateur M , la prise en compte des interactions entre X et des covariables (facteurs de confusion ou médiateurs)

n'est pas indispensable pour répondre à la question scientifique. Un effet "moyen" pourra être estimé. Des termes d'interactions peuvent cependant être ajoutés (mais non interprétés), si cela améliore la précision de l'estimation (enjeu d'optimisation du modèle).

2.1.2 Types d'objectifs

Dans ce document, nous nous intéresserons principalement aux interactions et modifications d'effet dans une démarche étiologique/ explicative.

Les objectifs pouvant nécessiter l'étude de l'interaction/modification d'effet sont VanderWeele and Knol [2014] :

- **Cibler des sous-groupes.** Par exemple, identifier des sous-groupes pour lesquels l'intervention aura le plus d'effet afin de pouvoir cibler l'intervention en cas de ressources limitées, ou s'assurer que l'intervention est bénéfique pour tous les groupes et pas délétères pour certains groupes.
- **Explorer les mécanismes d'un effet.** Par exemple, en cas d'intervention qui n'a d'effet qu'en présence ou absence d'une caractéristique particulière (définition mécanistique de l'interaction) ou seulement conjointement à une autre intervention.
- **Etudier l'effet d'une intervention pour éliminer une partie de l'effet d'une exposition non modifiable.** Par exemple, quelle part de l'effet du niveau d'éducation des parents sur la mortalité disparaîtrait si on intervenait sur le tabagisme à l'adolescence ? Ce type d'objectif est proche d'un objectif ciblant la *médiation* d'un effet, par exemple la médiation de l'effet du niveau d'éducation des parents *par* le tabagisme, mais les mécanismes envisagés et explorés ne sont pas exactement les mêmes. Explorer ces deux types de mécanismes peut nécessiter des approches spécifiques (voir chapitre 18)

2.2 Les points les plus importants

La première étape importante consiste donc à **définir précisément l'objectif** :

- L'objectif est-il de type descriptif, prédictif ou explicatif ?
- Si l'on est dans une démarche explicative, d'inférence causale, est-ce que la mesure d'un effet d'interaction est nécessaire pour y répondre ? (identifier précisément l'effet que l'on cherche à estimer, ou *estimand*).

Ensuite, de **nombreuses questions** se posent pour réaliser une analyse d'interaction, auxquelles nous tentons de répondre dans ce document :

- S'agit-il d'une interaction ou une modification d'effet ? (Chapitre 4)
- Sur quelle échelle la mesure-t-on ? Un effet d'interaction peut en effet être défini sur une échelle multiplicative ou additive, et les résultats entre ces échelles peuvent être contradictoires. (Chapitre 5)
- Quels paramètres présenter et comment les interpréter ? (Chapitre 6)
- Comment estimer ces paramètres ? (Chapitre 9 et Chapitre 10)
- Comment représenter cette interaction graphiquement ? (Chapitre 11)

Part I

Synthèse de la littérature

Chapter 3

Notations

3.1 Variables et probabilités

On note :

- un outcome : Y ,
- deux expositions : X et V

La probabilité de l'outcome Y dans chaque strate définie par les 2 expositions est notée :

- $p_{xv} = P(Y = 1|X = x, V = v)$

Exemple

On a deux expositions X , le tabagisme actif à 20 ans, et V , le fait d'avoir vécu un évènement traumatique pendant l'enfance. L'outcome Y est binaire et représente le fait d'avoir au moins une pathologie chronique à 60 ans ($Y = 1$) ou aucune ($Y = 0$).

On décrit (données complètement fictives) :

$X \setminus V$	$V = 0$	$V = 1$
$X = 0$	$P_{00} = 0,1$	$P_{10} = 0,2$
$X = 1$	$P_{01} = 0,4$	$P_{11} = 0,9$

Interprétation : La probabilité d'avoir au moins une pathologie chronique à 60 ans quand on n'a pas vécu d'évènement traumatique pendant l'enfance et pas fumé à 20 ans est de 10%, tandis qu'elle est de 90% quand on a vécu un évènement traumatique et fumé.

3.2 Mesures d'effets

L'effet d'une variable X sur Y peut être mesuré sur deux échelles : additive (différence de risque/probabilité) ou multiplicative (rapport de risque/probabilité).

Concernant les différences de risques (DR, effets additifs)

On va noter $P(Y = 1|do(X = 1))$ la probabilité d'observer $Y = 1$ sous une intervention contrefactuelle où la totalité de la population étudiée est exposée à $X = 1$ (notée $do(X = 1)$).

De même, on va noter $P(Y = 1|do(X = 1, V = 1))$ la probabilité d'observer $Y = 1$ sous une intervention contrefactuelle conjointe à la fois sur X et sur V où la totalité de la population étudiée est exposée à $X = 1$ et $V = 1$ (notée $do(X = 1, V = 1)$).

- L'effet d'une exposition X binaire sur Y est : $DR(X) = P(Y = 1|do(X = 1)) - P(Y = 1|do(X = 0))$
 - qu'on peut estimer, si les conditions d'identifiabilité sont réunies,
 - par $P(Y = 1|X = 1) - P(Y = 1|X = 0) = p_1 - p_0$
- L'effet conjoint de X et V est : $DR(X, V) = p_{11} - p_{00}$
- L'effet de X sur Y pour chaque valeur fixée de V est : $DR(X, V = 0) = p_{10} - p_{00}$ et $DR(X, V = 1) = p_{11} - p_{01}$

Exemple

Différences de risques pour l'exemple 1

$X \setminus V$	$V = 0$	$V = 1$
$X = 0$	$p_{00} = 0,1$	$p_{10} = 0,2$
$X = 1$	$p_{01} = 0,4$	$p_{11} = 0,9$

- $DR(X, V) = p_{11} - p_{00} = 0,90 - 0,10 = +0,80$
- $DR(X, V = 0) = p_{10} - p_{00} = 0,40 - 0,10 = +0,30$
- $DR(X, V = 1) = p_{11} - p_{01} = 0,90 - 0,20 = +0,70$

Le fait d'être doublement exposé par rapport à pas du tout augmente le risque de +80%. Dans une population n'ayant pas vécu d'événement traumatique, le fait de fumer à 20 ans augmente le risque de +30%, alors que dans une population ayant vécu un événement traumatique, il est augmenté de +70%.

Concernant, les rapports de risque (effets multiplicatifs)

on peut notamment utiliser les **risques relatifs** (RR). On donc :

- L'effet d'une exposition X binaire sur Y est : $RR(X) = P(Y = 1|do(X = 1))/P(Y = 1|do(X = 0))$
 - qu'on peut estimer, si les conditions d'identifiabilité sont réunies,
 - par $P(Y = 1|do(X = 1))/P(Y = 1|do(X = 0)) = p_1/p_0$
- L'effet conjoint de X et V est : $RR(X, V) = p_{11}/p_{00}$
- L'effet de X sur Y pour chaque valeur fixée de V est : $RR(X, V = 0) = p_{10}/p_{00}$ et $RR(X, V = 1) = p_{11}/p_{01}$

Exemple

Risques relatifs pour l'exemple 1

$X \setminus V$	$V = 0$	$V = 1$
$X = 0$	$p_{00} = 0,1$	$p_{10} = 0,2$
$X = 1$	$p_{01} = 0,4$	$p_{11} = 0,9$

- $RR(X, V) = 0,9/0,1 = \times 9$
- $RR(X, V = 0) = 0,4/0,1 = \times 4$
- $RR(X, V = 1) = 0,9/0,2 = \times 4,5$

Le risque quand on est doublement exposé par rapport à pas du tout est multiplié par 9. Dans une population n'ayant pas vécu d'événement traumatique, le fait de fumer à 20 ans multiplie le risque par 4, alors que dans une population ayant vécu un événement traumatique, il est multiplié par 4,5.

Chapter 4

Interaction vs modification d'effets

Dans le champ des analyses d'interaction, deux termes peuvent être rencontrés : “interaction” et “modification d'effet”. Quel est la différence entre ces deux termes ?

4.1 Modification d'effets

La question de la modification d'effet consiste à identifier si un scénario contrefactuel modifiant le traitement ou l'exposition X donne un résultat différent dans différents groupes V de patients (estimer l'effet d'une exposition séparément en fonction d'une autre variable) Corraini et al. [2017].

Si l'on compare avec un essai d'intervention, c'est comme s'il y avait une seule intervention X et que l'analyse était stratifiée sur V . On analyse donc l'effet du scénario $do(X)$ dans chaque groupe de V .

En observationnel, l'effet causal qui nous intéresse est donc celui de X mais pas celui de V . On ajustera sur les facteurs de confusion de la relation $X \rightarrow Y$.

On ne fait pas d'hypothèse sur les mécanismes de la modification d'effet, qui peut être causale (de façon directe ou indirecte), ou non-causale (présence d'une modification d'effet par proxy ou cause commune, sans qu'il existe d'effet direct ou indirect du modificateur d'effet vers le critère de jugement, comme dans la figure en bas de page) VanderWeele and Robins [2007].

Exemples d'objectifs : identifier des groupes pour lesquels le traitement ne serait pas utile, ou explorer si l'effet du traitement est homogène/hétérogène en fonction de l'âge, du sexe, etc.

On a une modification de l'effet de X par V si l'effet de X est différent dans deux strates définies par V :

- en additif : $DR(X|V = 0) \neq DR(X|V = 1)$
 - soit $p_{10} - p_{00} \neq p_{11} - p_{01}$
- en multiplicatif : $RR(X|V = 0) \neq RR(X|V = 1)$
 - soit $p_{10}/p_{00} \neq p_{11}/p_{01}$

Exemple

Modification d'effet dans l'exemple 1

En additif :

- effet dans le groupe $V = 0$: $DR(X|V = 0) = 0,40 - 0,10 = +0,30$
- effet dans le groupe $V = 1$: $DR(X|V = 1) = 0,90 - 0,20 = +0,70$
- donc $DR(X|V = 0) \neq DR(X|V = 1)$

En multiplicatif :

- effet dans le groupe $V = 0$: $RR(X|V = 0) = 0,40/0,10 = \times 4,0$
- effet dans le groupe $V = 1$: $RR(X|V = 1) = 0,90/0,20 = \times 4,5$
- donc $RR(X|V = 0) \neq RR(X|V = 1)$

Ici l'effet du tabagisme est différent selon que les personnes ont vécu un événement traumatique ou non, sur l'échelle additive et multiplicative (données fictives). On peut donc dire que le fait d'avoir vécu un événement traumatique modifie l'effet du tabac. Attention, dans cet exemple, on fait l'hypothèse de l'absence de facteurs de confusion entre le tabagisme et l'outcome, ce qui est en réalité peu probable.

4.2 Interaction

Quand on s'intéresse à l'interaction, on s'intéresse plutôt à l'effet conjoint de 2 expositions (ou plus) sur un outcome. Il y a une interaction synergique si l'effet conjoint est supérieur à l'effet de la somme des individuels. Il y a une interaction antagoniste lorsque l'effet conjoint est inférieur à la somme des effets individuels Corraini et al. [2017].

Si l'on compare avec un essai d'intervention, c'est comme s'il y a plusieurs interventions selon le nombre de combinaison. On analyse donc l'effet du scénario $do(X, V)$. Ici l'effet causal d'intérêt est vraiment l'effet conjoint des deux variables.

Dans un schéma observationnel, l'effet causal qui nous intéresse est donc celui de l'interaction $X * V$. On ajustera sur les facteurs de confusion des deux relations $X \rightarrow Y$ et $V \rightarrow Y$. On fait l'hypothèse que les mécanismes de l'effet conjoint de X et V sont causaux.

Par définition, on a une interaction si l'effet conjoint de X et V sur Y ($DR(X, V)$) est différent de la somme (ou du produit sur l'échelle multiplicative) :

- de l'effet isolé de X sur Y (où V est constant, fixé à $V = 0$), noté $DR(X, V = 0)$ (ou $RR(X, V = 0)$)
- et de l'effet isolé de V sur Y (où X est constant, fixé à $X = 0$), noté $DR(V, X = 0)$ (ou $RR(V, X = 0)$)

On a ainsi,

- en additif : $DR(X, V) \neq DR(X, V = 0) + DR(V, X = 0)$
 - $p_{11} - p_{00} \neq (p_{10} - p_{00}) + (p_{01} - p_{00})$
 - $p_{11} \neq p_{10} + p_{01} - p_{00}$
 - $p_{11} - p_{10} - p_{01} + p_{00} \neq 0$
- en multiplicatif $RR(X, V) \neq RR(X, V = 0) \times RR(V, X = 0)$
 - $p_{11}/p_{00} \neq (p_{10}/p_{00}) \times (p_{01}/p_{00})$
 - $p_{11} \neq (p_{10} \times p_{01})/p_{00}$
 - $(p_{00} \times p_{11})/(p_{10} \times p_{01}) \neq 1$

Exemple

Interaction dans l'exemple 1

En additif :

- effet joint : $DR(X, V) = 0,90 - 0,10 = +0,80$
- somme des effets individuels : $DR(X, V = 0) + DR(V, X = 0) = +0,30 + 0,10 = +0,40$
- donc $DR(X, V) \neq DR(X, V = 0) + DR(V, X = 0)$

En multiplicatif :

- effet joint : $RR(X, V) = 0,9/0,1 = \times 9$
- produit des effets individuels : $RR(X, V = 0) \times RR(V, X = 0) = 4 \times 2 = \times 8$

- donc $DR(X, V) \neq DR(X, V = 0) \times DR(V, X = 0)$

Ici l'effet joint des 2 expositions est supérieur à la somme ou au produit des effets individuels, il y a donc une interaction synergique entre les deux expositions.

4.3 Synthèse

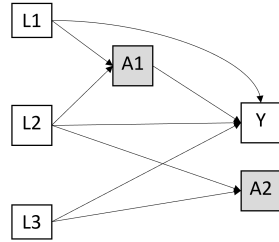
Mathématiquement, les formulations sont équivalentes :

- échelle additive: $p_{10} - p_{00} \neq p_{11} - p_{01} \iff p_{11} \neq (p_{10} + p_{01}) - p_{00}$
- échelle multiplicative : $p_{10}/p_{00} \neq p_{11}/p_{01} \iff p_{11} \neq (p_{10} \times p_{01})/p_{00}$

La différence se joue plutôt sur :

- la façon dont la question est posée (effet de X selon V , *versus* effet conjoint de X et V),
- sur les hypothèses causales formulées (scénarii $do(X)|V$ *versus* $do(X, V)$)
- et donc sur les sets de facteurs de confusion à considérer (seulement sur la relation $X \rightarrow Y$ *versus* les deux relations $X \rightarrow Y$ et $V \rightarrow Y$).

Il existe des cas où l'identification d'une interaction ou d'une modification d'effet ne conduira pas à la même démarche et donc au même résultat VanderWeele [2009]. Prenons le DAG suivant :



Dans ce cas, il n'y a pas d'interaction entre $A1$ et $A2$, car il n'y a pas d'effet direct ni indirect de $A2 \rightarrow Y$. L'effet de $A1 \rightarrow Y$ restera le même quelle que soit la valeur que l'on pourrait attribuer à $A2$:

$$P[Y = 1|do(A1 = 1, A2 = 0)] - P[Y = 1|do(A1 = 0, A2 = 0)] = P[Y = 1|do(A1 = 1, A2 = 1)] - P[Y = 1|do(A1 = 0, A2 = 1)]$$

Par contre, il peut y avoir une modification de l'effet de $A1$ par $A2$, en particulier s'il existe une interaction $A1 * L2 \rightarrow Y$ ou $A1 * L3 \rightarrow Y$, on s'attend à ce que les contrastes suivants soient différents :

$$P[Y = 1|do(A1 = 1), A2 = 1] - P[Y = 1|do(A1 = 2), A2 = 1] \neq P[Y = 1|do(A1 = 1), A2 = 0] - P[Y = 1|do(A1 = 2), A2 = 0].$$

Chapter 5

La question des échelles

5.1 Mesures des interactions

Echelle additive

Une façon simple de mesurer l'interaction est de mesurer à quel point l'effet conjoint de deux facteurs est différents de la somme de leurs effets individuels VanderWeele and Knol [2014] :






- $AI = DR(X, V) - [DR(X|V = 0) + DR(V|X = 0)]$
- $AI = (p_{11} - p_{00}) - [(p_{10} - p_{00}) + (p_{01} - p_{00})]$
- soit $AI = p_{11} - p_{10} - p_{01} + p_{00}$

Exemple

Mesure de l'interaction dans l'exemple 1

- $DR(X, V) - [DR(X|V = 0) + DR(V|X = 0)] = 0.8 - (0.3 + 0.1) = +0.4$
- soit $p_{11} - p_{10} - p_{01} + p_{00} = 0.9 - 0.4 - 0.2 + 0.1 = +0.4$
- ou $(p_{11} - p_{01}) - (p_{10} - p_{00}) = (0.9 - 0.2) - (0.4 - 0.1) = 0.7 - 0.3 = +0.4$
- ou $(p_{11} - p_{10}) - (p_{01} - p_{00}) = (0.9 - 0.4) - (0.2 - 0.1) = 0.5 - 0.1 = +0.4$

soit :

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	P00 = +0,1	P01 = +0,2	+0,1 
$X = 1$	P10 = +0,4	P11 = +0,9 	+0,5 
Effet X	+0,3 	+0,7 	+0,4 (+0,5)

+0,4

Echelle multiplicative

En cas d'outcome binaire, c'est souvent le RR ou l'OR qui est utilisé pour mesurer les effets. La mesure de l'interaction sur une échelle multiplicative serait donc VanderWeele and Knol [2014] :






- $MI = \frac{RR_{11}}{RR_{10} \times RR_{01}}$
- soit $MI = \frac{p_{11}/p_{00}}{(p_{10}/p_{00}) \times (p_{01}/p_{00})}$
- soit $MI = \frac{p_{11} \times p_{00}}{p_{10} \times p_{01}}$

Exemple

Mesure de l'interaction dans l'exemple 1

- $\frac{RR(X,V)}{RR(X|V=0) \times RR(V|X=0)} = 9/(4 \times 2) = \times 1,1$
- soit $\frac{p_{11}/p_{00}}{(p_{10}+p_{01})/p_{00}} = \frac{0,9/0,1}{(0,4 \times 0,2)/0,1} = \times 1,1$
- ou $\frac{p_{11}/p_{01}}{p_{10}/p_{00}} = \frac{0,9/0,2}{0,4/0,1} = \times 4,5 / \times 4 = \times 1,1$
- ou $\frac{p_{11}/p_{10}}{p_{01}/p_{00}} = \frac{0,9/0,4}{0,2/0,1} = \times 2,25 - \times 2 = \times 1,1$

ou :

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	P00 = +0,1	P01 = +0,2	x2 
$X = 1$	P10 = +0,4	P11 = +0,9 	x2,25 
Effet X	x4 	x4,5 	+1,1 (x8)

+1,1

5.2 Lien entre les deux échelles

Un apparent paradoxe

Mesurer l'interaction sur une seule échelle peut être trompeur Mathur and VanderWeele [2018]. On peut régulièrement observer une interaction positive dans

une échelle (par exemple $p_{11} - p_{10} - p_{01} + p_{00} > 0$) et négative dans l'autre (par exemple $(p_{11} \times p_{00}) / (p_{10} \times p_{01}) < 1$).

Exemple

Dans cet exemple (on a juste modifié la probabilité p_{11} , on observe une interaction additive positive (l'effet de X augmente de +20% quand $V = 1$ par rapport à $V = 0$) mais une interaction multiplicative négative (l'effet de X est multiplié par 0,9 - donc diminue - quand $V = 1$ par rapport à $V = 0$).

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P_{00} = +0,1$	$P_{01} = +0,2$	+0,1
$X = 1$	$P_{10} = +0,4$	$P_{11} = +0,7$	+0,3
Effet X	+0,3	+0,5	(+0,5)

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P_{00} = +0,1$	$P_{01} = +0,2$	$\times 2$
$X = 1$	$P_{10} = +0,4$	$P_{11} = +0,7$	$\times 1,75$
Effet X	$\times 4$	$\times 3,5$	$\times 0,9$ ($\times 8$)

Il a même été démontré que si on n'observe pas d'interaction sur une échelle, alors on en observera obligatoirement sur l'autre échelle... VanderWeele and Knol [2014].

Exemple

Dans cet exemple, il n'y a pas d'interaction multiplicative (effet de X identique quelque soit V), mais sur l'échelle additive, on observe une interaction positive.

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P_{00} = +0,1$	$P_{01} = +0,2$	+0,1
$X = 1$	$P_{10} = +0,4$	$P_{11} = +0,8$	+0,4
Effet X	+0,3	+0,6	(+0,5)

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P_{00} = +0,1$	$P_{01} = +0,2$	$\times 2$
$X = 1$	$P_{10} = +0,4$	$P_{11} = +0,8$	$\times 2$
Effet X	$\times 4$	$\times 4$	$\times 1$ ($\times 8$)

et dans cet autre exemple, il n'y a pas d'interaction additive (effet de X identique quelque soit V), mais sur l'échelle multiplicative, on observe une interaction négative.

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P_{00} = +0,1$	$P_{01} = +0,2$	+0,1
$X = 1$	$P_{10} = +0,4$	$P_{11} = +0,5$	+0,1
Effet X	+0,3	+0,3	(+0,5)

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P_{00} = +0,1$	$P_{01} = +0,2$	$\times 2$
$X = 1$	$P_{10} = +0,4$	$P_{11} = +0,5$	$\times 1,25$
Effet X	$\times 4$	$\times 2,5$	$\times 0,6$ ($\times 8$)

Le continuum

Dans un article de 2019 VanderWeele [2019], Vanderweele décrit le continuum existant entre les 2 échelles.

Par exemple, avec deux expositions ayant un effet positif (qui augmentent le risque) sur l'outcome en l'absence de l'autre exposition, lorsque l'effet joint est

très important, l'interaction est positive sur les 2 échelles. Mais lorsque la taille de l'effet joint diminue, l'interaction multiplicative devient négative alors que l'interaction additive reste positive. Puis, lorsque la taille de l'effet joint diminue encore, l'interaction devient négative sur les deux échelles.

X1 \ X2	X2 = 0	X2 = 1	Effet X2
X1 = 0	P00 = +0,1	P01 = +0,2	+0,1
X1 = 1	P10 = +0,4	P11 = + Δ	E(X2 X1=1)
Effet X1	+0,3	E(X1 X2=1)	(+0,5)

X1 \ X2	X2 = 0	X2 = 1	Effet X2
X1 = 0	P00 = +0,1	P01 = +0,2	x2
X1 = 1	P10 = +0,4	P11 = + Δ	RR(X2 X1=1)
Effet X1	x4	RR(X1 X2=1)	(+0,8)

	P11	MI	AI			RR(X1 X2=1) E(X1 X2=1)	RR(X2 X1=1) E(X2 X1=1)
1	0.9	1.1	+0.4	M+	positive-multiplicative	4.5	2.3
				A+	positive-additive	+0.7	+0.5
2	0.8	1.0	+0.3	M ₀	no-multiplicative	4	2
				A+	positive-additive	+0.6	+0.4
3	0.7	0.9	+0.2	M-	negative-multiplicative	3.5	1.75
				A+	positive-additive	+0.5	+0.3
4	0.5	0.6	+0.0	M-	negative-multiplicative	2.5	1.25
				A ₀	zero-additive	+0.3	+0.1
5	0.45	0.56	-0.05	M-	negative-multiplicative	2.3	1.1
				A-	negative-additive	+0.25	+0.05
6	0.4	0.5	-0.1	M-	single	2	1
				A-	pure interaction	+0.2	0.0
7	0.3	0.4	-0.2	M-	single	1.5	0.75
				A-	qualitative interaction	+0.1	-0.1
8	0.2	0.3	-0.3	M-	single-qualitative	1	0.5
				A-	single-pure interaction	0.0	-0.2
9	0.15	0.2	-0.35	M-	double	0.8	0.4
				A-	qualitative interaction	-0.05	-0.25
10	0.1	0.1	-0.4	M-	perfect antagonism	0.5	0.3
				A-		-0.1	-0.3
11	0.05	0.06	-0.45	M-	inverted interaction	0.3	0.1
				A-		-0.15	-0.35

Interaction pure et qualitative

Dans ce continuum, deux cas particuliers d'interaction peuvent être retrouvées :

- **Interaction pure** de X en fonction de V , si X n'a un effet que dans une seule strate de V . Par exemple, $p_{10} = p_{00}$ et $p_{11} \neq p_{01}$.

Par exemple ici, V a un effet si $X=0$ mais pas si $X=1$:

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	+0,1
X = 1	P10 = +0,4	P11 = +0,4	+0,0
Effet X	+0,3	+0,2	(+0,5)

X \ V	V = 0	V = 1	Effet V
X = 0	P00 = +0,1	P01 = +0,2	x2
X = 1	P10 = +0,4	P11 = +0,4	x1
Effet X	x4	x2	(x8)

- **Interaction qualitative** de X en fonction de V , si l'effet de X dans une strate de V va dans la direction opposée de l'autre strate de V .

Par exemple ici, V a un effet positif si $X=0$ mais négatif si $X=1$:

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P00 = +0,1$	$P01 = +0,2$	+0,1
$X = 1$	$P10 = +0,4$	$P11 = +0,3$	-0,1
Effet X	+0,3	+0,1	(+0,5)

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P00 = +0,1$	$P01 = +0,2$	x2
$X = 1$	$P10 = +0,4$	$P11 = +0,3$	x0,75
Effet X	x4	x1,5	(x8)

5.3 Synthèse

Quelle échelle choisir pour mesurer un effet d'interaction ?

Même si en pratique l'échelle multiplicative est plus utilisée, car les outcomes sont souvent binaires en épidémiologie et donc les modèles logistiques sont souvent utilisés Knol and VanderWeele [2012], il semble y avoir un consensus pour privilégier plutôt l'échelle additive, plus appropriée pour évaluer l'utilité en santé publique VanderWeele and Knol [2014] Knol and VanderWeele [2012].

Si on reprend l'exemple ci dessous :

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P00 = +0,1$	$P01 = +0,2$	+0,1
$X = 1$	$P10 = +0,4$	$P11 = +0,7$	+0,3 +0,2
Effet X	+0,3 +0,2	+0,5 +0,2	(+0,5)

$X \setminus V$	$V = 0$	$V = 1$	Effet V
$X = 0$	$P00 = +0,1$	$P01 = +0,2$	x2
$X = 1$	$P10 = +0,4$	$P11 = +0,7$	x1,75 x0,9
Effet X	x4 x0,9	x3,5 x0,9	(x8)

X représente un traitement dont on ne dispose que de 100 doses et Y un outcome de santé favorable (guérison). Il faut choisir si on donne 100 doses au groupe $V = 0$ ou au groupe $V = 1$.

Si on donne 100 doses au groupe $V = 0$, 30 personnes seront guéries grace au traitement (30 personnes de plus que l'évolution naturelle, $X = 0$) contre 50 personnes si on les donne au groupe $V = 1$. Donc il est préférable d'allouer les doses au groupe $V = 1$.

Pourtant si on avait réfléchi à partir de l'échelle multiplicative, on aurait choisi le groupe $V = 0$ car l'effet du traitement est de $RR=4$ dans le groupe $V = 0$ et $RR=3,5$ dans le groupe $V = 1$...

On peut donc conclure à un effet multiplicatif plus fort d'un traitement dans un groupe alors qu'en terme d'utilité (nombre de personnes favorablement impactées), l'échelle additive nous conduirait à choisir l'autre groupe...

Idéalement, les interactions devraient cependant être reportées sur les 2 échelles Knol and VanderWeele [2012] VanderWeele and Knol [2014].

Chapter 6

Types de paramètres

Plusieurs paramètres peuvent être utilisés pour décrire une interaction, sur l'échelle additive ou multiplicative.

6.1 Avec les différences de risques (DR)

On a déjà défini un paramètre d'interaction sur l'échelle additive (AI) à partir des différences d'effets VanderWeele and Knol [2014] :

- $AI = DR(X, V) - [DR(X|V = 0) + DR(V|X = 0)]$
- $AI = (p_{11} - p_{00}) - [(p_{10} - p_{00}) + (p_{01} - p_{00})]$
- soit $AI = p_{11} - p_{10} - p_{01} + p_{00}$

6.2 Avec les risques relatifs (RR)

On a aussi défini un paramètre d'interaction sur l'échelle multiplicative (MI) à partir des risques relatifs VanderWeele and Knol [2014] :

- $MI = \frac{RR_{11}}{RR_{10} \times RR_{01}}$
- soit $MI = \frac{p_{11}/p_{00}}{(p_{10}/p_{00}) \times (p_{01}/p_{00})}$
- soit $MI = \frac{p_{11} \times p_{00}}{p_{10} \times p_{01}}$

6.3 Avec les Odds Ratio (OR)

Souvent en épidémiologie, lorsque l'outcome Y est binaire, les effets sont mesurés par des odds ratios estimés à partir de modèles de régression logistique.

Un paramètre d'interaction sur l'échelle multiplicative (MI_{OR}) peut être estimé à partir de ces OR VanderWeele and Knol [2014] :

- $MI_{OR} = \frac{OR_{11}}{OR_{10} \times OR_{01}}$

En général, la mesure MI_{OR} et MI_{RR} seront proches si l'outcome est rare VanderWeele and Knol [2014].

6.4 Excès de risque à partir des RR (RERI)

Lorsque seulement les risques relatifs sont donnés mais que l'on souhaite évaluer l'interaction sur l'échelle additive, "l'excès de risque du à l'interaction" (RERI) ou "interaction contrast ratio" (ICR), peut être estimé à partir des risques relatifs VanderWeele and Knol [2014] :

$$RERI = \frac{AI}{p_{00}} = \frac{p_{11} - p_{10} - p_{01} + p_{00}}{p_{00}}$$

$$RERI = RR_{11} - RR_{10} - RR_{01} + 1$$

On voit que le RERI correspond à l'interaction mesurée sur l'échelle additive, rapportée au risque de base p_{00} .

Il faut noter que, bien que le RERI donne la direction (positive, négative ou nulle) de l'interaction additive, nous ne pouvons pas utiliser le RERI pour évaluer l'ampleur de l'interaction additive, à moins de connaître au moins p_{00} .

Si l'on a seulement l'OR et que l'outcome est rare, les OR peuvent approximer les RR, on a donc :

- $RERI_{OR} = OR_{11} - OR_{10} - OR_{01} + 1 \approx RERI_{RR}$

6.5 Autres

D'autres paramètres ont aussi été proposés VanderWeele and Knol [2014], tels que :

Le "Synergie index" (SI)

Il s'agit d'un paramètre explorant l'interaction additive :

- $S = \frac{RR_{11}-1}{(RR_{10}-1)+(RR_{01}-1)}.$

Il mesure l'augmentation relative du risque liée à l'exposition jointe, rapportée à la somme des augmentations relatives du risque liées à la 1ère et la 2ème exposition.

Pour rappel, l'augmentation relative du risque liée à l'exposition jointe correspond à l'augmentation absolue du risque (différence de risque), exprimée en pourcentage par rapport au risque de base p_{00} .

$$ARR(X, V) = \frac{DR(X, V)}{p_{00}} = \frac{p_{11}-p_{00}}{p_{00}} = RR_{11} - 1$$

L'augmentation relative du risque liée à l'exposition X ou V , exprimées en pourcentage par rapport au risque de base p_{00} sont respectivement :

- $ARR(X|V=0) = \frac{p_{10}-p_{00}}{p_{00}} = RR_{10} - 1$
- et $ARR(V|X=0) = \frac{p_{01}-p_{00}}{p_{00}} = RR_{01} - 1$

Une autre façon d'interpréter le "Synergie index" est d'indiquer que la différence liée à l'effet joint $DR(X, V)$ est égale à S fois la somme des différences liées aux effets individuels $DR(X|V=0) + DR(V|X=0)$:

$$S = \frac{p_{11}-p_{00}}{(p_{10}-p_{00})+(p_{01}-p_{00})}$$

Si le dénominateur est positif:

- si $S > 1$, alors $AI > 0$ et $RERI_{RR} > 0$
- si $S < 1$, alors $AI < 0$ et $RERI_{RR} < 0$

L'interprétation de l'indice de synergie devient difficile dans les cas où l'effet de l'une des expositions a un effet négatif et que le dénominateur de S est inférieur à 1.

6.5.1 Proportion attribuable (AP)

Il s'agit aussi d'un paramètre explorant l'interaction additive :

- $AP = \frac{RR_{11}-RR_{10}-RR_{01}+1}{RR_{11}}.$

Ce paramètre mesure la proportion du risque dans le groupe doublement exposé qui est due à l'interaction.

L'AP est en lien avec le $RERI_{RR}$:

- $AP > 0$ si et seulement si $RERI_{RR} > 0$
- $AP < 0$ si et seulement si $RERI_{RR} < 0$.

En fait $AP = \frac{RERI_{RR}}{RR_{11}-1}.$

Part II

Estimations, Interprétations, Présentations

Chapter 7

Présentation des résultats

7.1 Recommendations

Knol et VanderWeele ont émis des recommandations concernant la présentation des résultats d'une analyse d'interactions Knol and VanderWeele [2012]. Ces recommandations sont :

Pour une analyse d'une modification d'effet de A1 sur Y par A2

- Présenter les effectifs dans chaque catégorie
 - avec et sans l'outcome ($N_{a1,a2}(Y = 1)$ et $N_{a1,a2}(Y = 0)$)
- Présenter les risques relatifs (RR), les OR ou les différences de risque (RD)
 - avec les intervalles de confiance (IC)
 - pour chaque strate de A1 et de A2 avec une seule catégorie de référence
 - (éventuellement prise comme la strate A1.A2 présentant le plus faible risque de Y).
- Présenter les RR, OR ou RD avec les IC
 - de l'effet de A1 sur Y dans les strates de A2
- Présenter les mesures de la modification de l'effet avec les IC, sur des échelles
 - additives (par exemple, RERI)
 - et multiplicatives.
- Énumérez les facteurs de confusion pour lesquels la relation entre A1 et Y a été ajustée.

Interaction A1*A2 sur Y

- Présenter les effectifs dans chaque catégorie
 - avec et sans l’outcome ($N_{a1,a2}(Y = 1)$ et $N_{a1,a2}(Y = 0)$)
- Présenter les risques relatifs (RR), les OR ou les différences de risque (RD)
 - avec les intervalles de confiance (IC)
 - pour chaque strate de A1 et de A2 avec une seule catégorie de référence
 - (éventuellement prise comme la strate A1.A2 présentant le plus faible risque de Y).
- Présenter les RR, OR ou RD avec les IC
 - de l’effet de A1 sur Y dans les strates de A2
 - **et de A2 sur Y dans les strates de A1.**
- Présenter les mesures d’e la modification de l’effet’interaction avec les IC sur des échelles
 - additives (par exemple, RERI)
 - et multiplicatives.
- Énumérez les facteurs de confusion pour lesquels la relation entre A1 et Y **et la relation entre A2 et Y** ont été ajustées.

7.2 Proposition

Ces recommandations sont très utiles lorsque les interactions ont été évaluées à partir de modèles de régression (logistiques, log-linéaires ou linéaires) permettant d’estimer directement des OR, des RR ou des DR, conditionnellement aux facteurs de confusion.

En inférence causale, des associations marginales plutôt que conditionnelles sont souvent estimées (que ce soit en termes de différence de risque, de risques relatifs ou d’odds ratio). Dans la suite de ce document, nous proposons une variante des recommandations de Knol et VanderWeele, adaptée à des estimations marginales. Nous proposons en effet :

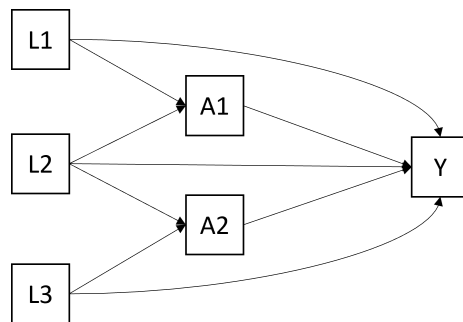
- De présenter les effets marginaux ou proportions prédites de Y dans chaque strate A1.A2,
 - plutôt les effectifs avec et sans l’outcome
- Ne pas forcément présenter une différence de risque ou un rapport de risque
 - pour chaque strate de A1 et de A2 avec une seule catégorie de référence

- Mais présenter les effets
 - de A1 dans chaque strate de A2
 - et de A2 dans chaque strate de A1 (si analyse d'interaction)
 - dans une échelle multiplicative **et** additive.

Chapter 8

Simulations

Pour la description des différents types d'estimation, on a simulé des données selon le DAG suivant (toutes les variables sont binaires):



Le code ayant permis de simuler les données est le suivant :

```
rm(list=ls())

param.causal.model <- function(p_L1 = 0.50,
                                p_L2 = 0.20,
                                p_L3 = 0.70,      # baseline confounders
                                b_A1 = 0.10,
                                b_L1_A1 = 0.15,
                                b_L2_A1 = 0.25, # modèle de A1
                                b_A2 = 0.15,
                                b_L1_A2 = 0.20,
                                b_L3_A2 = 0.20, # modèle de A2
                                b_Y = 0.10,      # modèle de Y
                                b_L1_Y = 0.02,
```

```

        b_L2_Y = 0.02,
        b_L3_Y = -0.02,
        b_A1_Y = 0.3,
        b_A2_Y = 0.1,
        b_A1A2_Y = 0.4 ) { # <- effet d'interaction Delta)

# coefficients pour simuler l'exposition
# exposition A1 # vérif
try(if(b_A1 + b_L1_A1 + b_L1_A1 > 1)
    stop("la somme des coefficient du modèle A1 dépasse 100%"))

# exposition A2 # vérif
try(if(b_A2 + b_L1_A2 + b_L3_A2 > 1)
    stop("la somme des coefficients du modèle A2 dépasse 100%"))

# coefficients pour simuler l'outcome, vérif
try(if(b_Y + b_L1_Y + b_L2_Y + b_L3_Y + b_A1_Y + b_A2_Y + b_A1A2_Y > 1)
    stop("la somme des coefficients du modèle Y dépasse 100%"))
try(if(b_Y + b_L1_Y + b_L2_Y + b_L3_Y + b_A1_Y + b_A2_Y + b_A1A2_Y < 0)
    stop("la somme des coefficients du modèle Y est inférieure à 0%"))

coef <- list(c(p_L1 = p_L1, p_L2 = p_L2, p_L3 = p_L3),
             c(b_A1 = b_A1, b_L1_A1 = b_L1_A1, b_L2_A1 = b_L2_A1),
             c(b_A2 = b_A2, b_L1_A2 = b_L1_A2, b_L3_A2 = b_L3_A2),
             c(b_Y = b_Y, b_L1_Y = b_L1_Y, b_L2_Y = b_L2_Y, b_L3_Y = b_L3_Y,
               b_A1_Y = b_A1_Y, b_A2_Y = b_A2_Y, b_A1A2_Y = b_A1A2_Y))
return(coef)
}

generate.data <- function(N, b = param.causal.model()) {

  L1 <- rbinom(N, size = 1, prob = b[[1]]["p_L1"])
  L2 <- rbinom(N, size = 1, prob = b[[1]]["p_L2"])
  L3 <- rbinom(N, size = 1, prob = b[[1]]["p_L3"])
  A1 <- rbinom(N, size = 1, prob = b[[2]]["b_A1"] +
    (b[[2]]["b_L1_A1"] * L1) + (b[[2]]["b_L2_A1"] * L2))
  A2 <- rbinom(N, size = 1, prob = b[[3]]["b_A2"] +
    (b[[3]]["b_L1_A2"] * L1) + (b[[3]]["b_L3_A2"] * L3))
  Y <- rbinom(N, size = 1, prob = (b[[4]]["b_Y"] +
    (b[[4]]["b_L1_Y"] * L1) +
    (b[[4]]["b_L2_Y"] * L2) +
    (b[[4]]["b_L3_Y"] * L3) +
    (b[[4]]["b_A1_Y"] * A1) +
    (b[[4]]["b_A2_Y"] * A2) +
    (b[[4]]["b_A1A2_Y"] * A1 * A2)) )

```



```

data.sim <- data.frame(L1, L2, L3, A1, A2, Y)
return(data.sim)
}

#### On simule une base de données
set.seed(12345)
# b = param.causal.model(b_A1A2_Y = -0.45)
b = param.causal.model()
df <- generate.data(N = 10000, b = b)
summary(df)
prop.table(table(df$Y, df$A1, df$A2, deparse.level = 2))

```

Au final, les probabilités de l'outcome $P(Y=1)$, dans chaque catégorie sont :

A2	label	levels	value
0	A1	0	0.10 (0.30)
0		1	0.41 (0.49)
1	A1	0	0.20 (0.40)
1		1	0.90 (0.30)

Chapter 9

A partir de modèles de régression

Dans une première étape exploratoire, on peut simplement utiliser les modèles de régression habituels : les modèles de régression logistique et linéaire.

9.1 Régression logistique

Lorsque l'on étudie un outcome binaire, on utilise souvent les modèles de régression logistique.

```
##
## Call:
## glm(formula = Y ~ as.factor(A1) + as.factor(A2) + as.factor(A1) *
##      as.factor(A2) + as.factor(L1) + as.factor(L2) + as.factor(L3),
##      family = binomial, data = df_f)
##
## Coefficients:
##                                Estimate Std. Error z value Pr(>|z|)
## (Intercept)                   -2.16540    0.06708 -32.281 < 2e-16 ***
## as.factor(A1)1                  1.75607    0.07604  23.093 < 2e-16 ***
## as.factor(A2)1                  0.75332    0.06831  11.028 < 2e-16 ***
## as.factor(L1)1                 0.15753    0.05702   2.763 0.00573 **
## as.factor(L2)1                 0.14128    0.06878   2.054 0.03996 *
## as.factor(L3)1                -0.14926    0.06141  -2.431 0.01507 *
## as.factor(A1)1:as.factor(A2)1  1.78587    0.14131  12.638 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 11037.7  on 9999  degrees of freedom
## Residual deviance:  8460.4  on 9993  degrees of freedom
## AIC: 8474.4
##
## Number of Fisher Scoring iterations: 4
```

A partir de cette sortie, on peut extraire :

- **A1|A2=0**
 - à partir du coefficient `as.factor(A1)1`
 - qui correspond à l'effet de A1 dans la catégorie de référence de A2,
 - soit $OR_{A1|A2=0} = \exp(1.756) = 5.789$.
- **A1|A2=1**
 - à partir du coefficient `as.factor(A1)1:as.factor(A2)1`,
 - qui correspond à la différence d'effet de A1 quand on passe dans l'autre catégorie de A2.
 - L'effet de A1 dans la catégorie A2=1 est donc
 - $OR_{A1|A2=1} = \exp(1.756 + 1.786) = 34.536$.
- **L'interaction multiplicative (IM)**
 - peut être estimée à partir du coefficient `as.factor(A1)1:as.factor(A2)1`
 - par $IM = \exp(1.786) = 5.966$,
 - qu'on peut retrouver en faisant $OR_{A1|A2=1}/OR_{A1|A2=0}$.
 - Ici l'interaction est significative (p-value >0.05).
- **A2|A1=0 et A2|A1=1**
 - On aurait aussi pu décrire l'interaction à partir de l'effet d'A2 dans chaque strate de A1
 - à partir de `as.factor(A2)1` et `as.factor(A1)1:as.factor(A2)1`,
 - avec : $OR_{A2|A1=0} = \exp(0.753) = 2.123$
 - et $OR_{A2|A1=1} = \exp(0.753 + 1.786) = 12.667$
- **L'interaction additive**
 - On peut explorer l'interaction sur l'échelle additive en estimant le RERI par
 - $RERI \approx OR_{11} - OR_{10} - OR_{01} + 1 =$
 - $OR_{A1,A2} - OR_{A1|A2=0} - OR_{A2|A1=0} + 1 =$
 - $\exp(1.786 + 0.753 + 1.786) - \exp(1.786) - \exp(0.753) + 1 = 68.477$.

En résumé, (le package `finalfit` permet de sortir quelques résultats proprement) :

names	OR
A1 A2=0	5.79 (4.99-6.72, p<0.001)
A2 A1=0	2.12 (1.86-2.43, p<0.001)
Interaction	5.96 (4.54-7.90, p<0.001)

```

explanatory = c("as.factor(A1)",
               "as.factor(A2)",
               "as.factor(A1)*as.factor(A2)",
               "as.factor(L1)",
               "as.factor(L2)",
               "as.factor(L3)")
dependent = "Y"

df_f %>%
  finalfit(dependent, explanatory)-> t

# le tableau t entier peut être imprimé, mais ici je sélectionne seulement les effets d'intérêt
# pour éviter la table 2 fallacy (les coefficient des facteurs de confusion L ne sont pas interprétés)

cbind(names = c("A1|A2=0", "A2|A1=0", "Interaction"),
      OR = t[c(12,14,13),6]) %>%
  as.data.frame %>%
  kbl() %>%
  kable_classic()

```

Attention, les modèles de régressions logistiques sont ici biaisés car les données sont générées à partir de modèles additifs.

9.2 Régression lineaire

Même si l'outcome binaire, on peut en théorie utiliser un modèle de régression linéaire et explorer les effets sur une échelle additive. Si l'outcome est quantitatif, on utilise aussi, en général, les modèles de régression linéaire.

```

##
## Call:
## lm(formula = Y ~ as.factor(A1) + as.factor(A2) + as.factor(A1) *
##      as.factor(A2) + as.factor(L1) + as.factor(L2) + as.factor(L3),
##      data = df)
##
## Residuals:
##      Min       1Q   Median       3Q      Max

```

```
## -0.93110 -0.19602 -0.10494 -0.08426  0.91574
##
## Coefficients:
##                                     Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      0.103835    0.008146  12.746 < 2e-16 ***
## as.factor(A1)1                   0.300796    0.011592  25.948 < 2e-16 ***
## as.factor(A2)1                   0.092280    0.008671  10.642 < 2e-16 ***
## as.factor(L1)1                   0.020677    0.007495   2.759 0.00581 **
## as.factor(L2)1                   0.019476    0.009410   2.070 0.03851 *
## as.factor(L3)1                  -0.019574    0.008085  -2.421 0.01549 *
## as.factor(A1)1:as.factor(A2)1    0.394034    0.017854  22.070 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.3615 on 9993 degrees of freedom
## Multiple R-squared:  0.2856, Adjusted R-squared:  0.2852
## F-statistic: 665.8 on 6 and 9993 DF,  p-value: < 2.2e-16
```

A partir de cette sortie, on peut extraire :

- **A1|A2=0**

- à partir du coefficient `as.factor(A1)1`
- qui correspond à l'effet de A1 dans la catégorie de référence de A2,
- soit $DR = +30,08\%$.

- **A1|A2=1**

- à partir du coefficient `as.factor(A1)1:as.factor(A2)1`,
- qui correspond à la différence d'effet de A1 quand on passe dans l'autre catégorie de A2.
- L'effet de A1 dans la catégorie A2=1 est donc
- $DR = 30.08 + 39.40 = 69.48 \%$.

- **L'interaction additive**

- à partir du coefficient `as.factor(A1)1:as.factor(A2)1`
- avec $AI = +39.40\%$,
- qu'on peut retrouver en faisant $DR(A1|A2 = 1) - DR(A1|A2 = 0)$.
- Ici l'interaction est significative (p-value >0.05).

- **A2|A1=0 et A2|A1=1**

- On aurait aussi pu décrire cette interaction à partir de l'effet d'A2 dans chaque strate de A1
- à partir de `as.factor(A2)1` et `as.factor(A1)1:as.factor(A2)1`,
- avec : $DR_{A1|A2=0} = +9.23\%$
- et $DR_{A1|A2=1} = 9.23 + 39.40 = 48.63\%$.

names	DR
A1 A2=0	0.30 (0.28 to 0.32, p<0.001)
A2 A1=0	0.09 (0.08 to 0.11, p<0.001)
Interaction	0.39 (0.36 to 0.43, p<0.001)

En résumé, (le package `finalfit` permet de sortir quelques résultats proprement) :

```

explanatory = c("as.factor(A1)",
                "as.factor(A2)",
                "as.factor(A1)*as.factor(A2)",
                "as.factor(L1)",
                "as.factor(L2)",
                "as.factor(L3)")
dependent = "Y"
df %>%
  finalfit(dependent, explanatory)-> t

cbind(names = c("A1|A2=0", "A2|A1=0", "Interaction"), DR = t[c(12,14,13),6]) %>%
  as.data.frame %>%
  kbl() %>%
  kable_classic()

```


Chapter 10

Analyses confirmatoires

10.1 Estimation par G-computation

Il s'agit d'une “G-methods” aussi appelée “standardisation” par Hernán Hernán and Robins [2020]. Le principe est le suivant :

```
## 1.a) on crée 4 tables correspondant aux 4 interventions contrefactuelles
df.A1_0.A2_0 <- df.A1_1.A2_0 <- df.A1_0.A2_1 <- df.A1_1.A2_1 <- df

df.A1_0.A2_0$A1 <- df.A1_0.A2_0$A2 <- rep(0, nrow(df))
df.A1_1.A2_0$A1 <- rep(1, nrow(df))
df.A1_1.A2_0$A2 <- rep(0, nrow(df))

df.A1_0.A2_1$A1 <- rep(0, nrow(df))
df.A1_0.A2_1$A2 <- rep(1, nrow(df))

df.A1_1.A2_1$A1 <- df.A1_1.A2_1$A2 <- rep(1, nrow(df))

## 1.b) on modélise le critère de jugement
# model.Y <- glm(Y ~ L1 + L2 + L3 + A1 + A2 + A1:A2, data = df, family = "binomial")
# modèle logistique biaisé (il y a des interactions avec les baseline)
model.Y <- glm(Y ~ L1 + L2 + L3 + A1 + A2 + A1:A2, data = df,
               family = "gaussian") # modèle non biaisé
# en pratique la régression logistique n'est pas tellement biaisée,
# mais peut être car il n'y a pas la place de mettre beaucoup de confusion
# par rapport aux effets importants de A1 et A2 ? (10 fois plus grands)

## 1.c) on prédit le critère de jugement sous les interventions contrefactuelles
Y.A1_0.A2_0 <- predict(model.Y, newdata = df.A1_0.A2_0, type = "response")
Y.A1_1.A2_0 <- predict(model.Y, newdata = df.A1_1.A2_0, type = "response")
```

```

Y.A1_0.A2_1 <- predict(model.Y, newdata = df.A1_0.A2_1, type = "response")
Y.A1_1.A2_1 <- predict(model.Y, newdata = df.A1_1.A2_1, type = "response")

## 1.d) on va enregistrer l'ensemble des résultats pertinents dans une table de longueurs
int.r <- matrix(NA,
               ncol = 26,
               nrow = nlevels(as.factor(df$A1)) * nlevels(as.factor(df$A2)))
int.r <- as.data.frame(int.r)
names(int.r) <- c("A1", "A2", "p", "p.lo", "p.up",
                 "RD.A1", "RD.A1.lo", "RD.A1.up", "RD.A2", "RD.A2.lo", "RD.A2.up",
                 "RR.A1", "RR.A1.lo", "RR.A1.up", "RR.A2", "RR.A2.lo", "RR.A2.up",
                 "a.INT", "a.INT.lo", "a.INT.up", "RERI", "RERI.lo", "RERI.up",
                 "m.INT", "m.INT.lo", "m.INT.up" )
int.r[,c("A1", "A2")] <- expand.grid(c(0,1), c(0,1))

# marginal effects (Y moyen dans chaque scénario) in the k1 x k2 table
# A1 = 0 et A2 = 0
int.r$p[int.r$A1 == 0 & int.r$A2 == 0] <- mean(Y.A1_0.A2_0)
# A1 = 1 et A2 = 0
int.r$p[int.r$A1 == 1 & int.r$A2 == 0] <- mean(Y.A1_1.A2_0)
# A1 = 0 et A2 = 1
int.r$p[int.r$A1 == 0 & int.r$A2 == 1] <- mean(Y.A1_0.A2_1)
# A1 = 1 et A2 = 1
int.r$p[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Y.A1_1.A2_1)

# risk difference (contrastes entre Y contrefactuels)
# RD.A1.A2is0
int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] <- mean(Y.A1_1.A2_0) - mean(Y.A1_0.A2_0)
# RD.A1.A2is1
int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Y.A1_1.A2_1) - mean(Y.A1_0.A2_1)
# RD.A2.A1is0
int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] <- mean(Y.A1_0.A2_1) - mean(Y.A1_0.A2_0)
# RD.A2.A1is1
int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Y.A1_1.A2_1) - mean(Y.A1_1.A2_0)

# relative risk (rapports entre Y contrefactuels)
# RR.A1.A2is0
int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0] <- mean(Y.A1_1.A2_0) / mean(Y.A1_0.A2_0)
# RR.A1.A2is1
int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Y.A1_1.A2_1) / mean(Y.A1_0.A2_1)
# RR.A2.A1is0
int.r$RR.A2[int.r$A1 == 0 & int.r$A2 == 1] <- mean(Y.A1_0.A2_1) / mean(Y.A1_0.A2_0)
# RR.A2.A1is1
int.r$RR.A2[int.r$A1 == 1 & int.r$A2 == 1] <- mean(Y.A1_1.A2_1) / mean(Y.A1_1.A2_0)

```

```

# additive interaction
int.r$a.INT[int.r$a1 == 1 & int.r$a2 == 1] <- mean(Y.A1_1.A2_1) -
  mean(Y.A1_1.A2_0) -
  mean(Y.A1_0.A2_1) +
  mean(Y.A1_0.A2_0)

# RERI
int.r$RERI[int.r$a1 == 1 & int.r$a2 == 1] <- (mean(Y.A1_1.A2_1) -
  mean(Y.A1_1.A2_0) -
  mean(Y.A1_0.A2_1) +
  mean(Y.A1_0.A2_0)) /
  mean(Y.A1_0.A2_0)

# multiplicative interaction
int.r$m.INT[int.r$a1 == 1 & int.r$a2 == 1] <- (mean(Y.A1_1.A2_1) *
  mean(Y.A1_0.A2_0)) /
  (mean(Y.A1_1.A2_0) *
  mean(Y.A1_0.A2_1))

## 1.e) Intervalles de confiance par bootstrap
set.seed(5678)
B <- 1000
bootstrap.est <- data.frame(matrix(NA, nrow = B, ncol = 15))
colnames(bootstrap.est) <- c("p.A1is0.A2is0", "p.A1is1.A2is0", "p.A1is0.A2is1", "p.A1is1.A2is1",
  "RD.A1.A2is0", "RD.A1.A2is1", "RD.A2.A1is0", "RD.A2.A1is1",
  "lnRR.A1.A2is0", "lnRR.A1.A2is1", "lnRR.A2.A1is0", "lnRR.A2.A1is1",
  "INT.a", "lnRERI", "lnINT.m")

for (b in 1:B){
  # sample the indices 1 to n with replacement
  bootIndices <- sample(1:nrow(df), replace=T)
  bootData <- df[bootIndices,]

  if ( round(b/100, 0) == b/100 ) print(paste0("bootstrap number ",b))

  # model (unbiased in this case)
  model.Y <- glm(Y ~ L1 + L2 + L3 + A1 + A2 + A1:A2,
    data = bootData, # use BootData here +++
    family = "gaussian")

  # counterfactual data sets
  boot.A1_0.A2_0 <- boot.A1_1.A2_0 <- boot.A1_0.A2_1 <- boot.A1_1.A2_1 <- bootData
  boot.A1_0.A2_0$A1 <- boot.A1_0.A2_0$A2 <- rep(0, nrow(df))
  boot.A1_1.A2_0$A1 <- rep(1, nrow(df))
  boot.A1_1.A2_0$A2 <- rep(0, nrow(df))
  boot.A1_0.A2_1$A1 <- rep(0, nrow(df))

```

```

boot.A1_0.A2_1$A2 <- rep(1, nrow(df))
boot.A1_1.A2_1$A1 <- boot.A1_1.A2_1$A2 <- rep(1, nrow(df))

# predict potential outcomes under counterfactual scenarios
Y.A1_0.A2_0 <- predict(model.Y, newdata = boot.A1_0.A2_0, type = "response")
Y.A1_1.A2_0 <- predict(model.Y, newdata = boot.A1_1.A2_0, type = "response")
Y.A1_0.A2_1 <- predict(model.Y, newdata = boot.A1_0.A2_1, type = "response")
Y.A1_1.A2_1 <- predict(model.Y, newdata = boot.A1_1.A2_1, type = "response")

# save results in the bootstrap table
bootstrap.est[b,"p.A1is0.A2is0"] <- mean(Y.A1_0.A2_0)
bootstrap.est[b,"p.A1is1.A2is0"] <- mean(Y.A1_1.A2_0)
bootstrap.est[b,"p.A1is0.A2is1"] <- mean(Y.A1_0.A2_1)
bootstrap.est[b,"p.A1is1.A2is1"] <- mean(Y.A1_1.A2_1)

bootstrap.est[b,"RD.A1.A2is0"] <- mean(Y.A1_1.A2_0) - mean(Y.A1_0.A2_0)
bootstrap.est[b,"RD.A1.A2is1"] <- mean(Y.A1_1.A2_1) - mean(Y.A1_0.A2_1)
bootstrap.est[b,"RD.A2.A1is0"] <- mean(Y.A1_0.A2_1) - mean(Y.A1_0.A2_0)
bootstrap.est[b,"RD.A2.A1is1"] <- mean(Y.A1_1.A2_1) - mean(Y.A1_1.A2_0)

bootstrap.est[b,"lnRR.A1.A2is0"] <- log(mean(Y.A1_1.A2_0) / mean(Y.A1_0.A2_0))
bootstrap.est[b,"lnRR.A1.A2is1"] <- log(mean(Y.A1_1.A2_1) / mean(Y.A1_0.A2_1))
bootstrap.est[b,"lnRR.A2.A1is0"] <- log(mean(Y.A1_0.A2_1) / mean(Y.A1_0.A2_0))
bootstrap.est[b,"lnRR.A2.A1is1"] <- log(mean(Y.A1_1.A2_1) / mean(Y.A1_1.A2_0))

bootstrap.est[b,"INT.a"] <- mean(Y.A1_1.A2_1) -
  mean(Y.A1_1.A2_0) - mean(Y.A1_0.A2_1) + mean(Y.A1_0.A2_0)
bootstrap.est[b,"lnRERI"] <- log((mean(Y.A1_1.A2_1) -
  mean(Y.A1_1.A2_0) - mean(Y.A1_0.A2_1) + mean(Y.A1_0.A2_0)) / mean(Y.A1_0.A2_0))
bootstrap.est[b,"lnINT.m"] <- log( (mean(Y.A1_1.A2_1) *
  mean(Y.A1_0.A2_0)) / (mean(Y.A1_1.A2_0) * mean(Y.A1_0.A2_1)))
}

# head(bootstrap.est)
# summary(bootstrap.est)
# par(mfrow = c(4,4))
# for(c in 1:ncol(bootstrap.est)) {
#   hist(bootstrap.est[,c], freq = FALSE, main = names(bootstrap.est)[c])
#   lines(density(bootstrap.est[,c]), col = 2, lwd = 3)
#   curve(1/sqrt(var(bootstrap.est[,c]) * 2 * pi) *
#     exp(-1/2 * ((x-mean(bootstrap.est[,c])) / sd(bootstrap.est[,c]))^2),
#     col = 1, lwd = 2, lty = 2, add = TRUE)
# par(mfrow = c(1,1))
# ok, on a des belles lois normales dans les distributions bootstrap, tout va bien
# pour les IC95%, je peux utiliser la déviation standard des distributions

```

```

# pour des distributions plus asymétriques, on utiliserait plutôt les percentiles 2.5% et 97.5%
# }

# marginal effects in the k1 x k2 table
# A1 = 0 et A2 = 0
int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0] -
  qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is0)
int.r$p.up[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0] +
  qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is0)
# A1 = 1 et A2 = 0
int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0] -
  qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is0)
int.r$p.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0] +
  qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is0)
# A1 = 0 et A2 = 1
int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] -
  qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is1)
int.r$p.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$p.A1is0.A2is1)
# A1 = 1 et A2 = 1
int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is1)
int.r$p.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$p.A1is1.A2is1)

# risk difference
# RD.A1.A2is0
int.r$RD.A1.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] -
  qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is0)
int.r$RD.A1.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] +
  qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is0)
# RD.A1.A2is1
int.r$RD.A1.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is1)
int.r$RD.A1.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$RD.A1.A2is1)
# RD.A2.A1is0
int.r$RD.A2.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] -
  qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is0)
int.r$RD.A2.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is0)
# RD.A2.A1is1
int.r$RD.A2.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is1)
int.r$RD.A2.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is1)

```

```

      qnorm(0.975) * sd(bootstrap.est$RD.A2.A1is1)

# relative risk
# RR.A1.A2is0
int.r$RR.A1.lo[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1 &
  int.r$A2 == 0]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A1.A2is0.lo))
int.r$RR.A1.up[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1 &
  int.r$A2 == 0]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A1.A2is0.up))

# RR.A1.A2is1
int.r$RR.A1.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A1.A2is1.lo))
int.r$RR.A1.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A1.A2is1.up))

# RR.A2.A1is0
int.r$RR.A2.lo[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A2.A1is0.lo))
int.r$RR.A2.up[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A2.A1is0.up))

# RR.A2.A1is1
int.r$RR.A2.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A2.A1is1.lo))
int.r$RR.A2.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RR.A2.A1is1.up))

# additive interaction
int.r$a.INT.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$a.INT.a)
int.r$a.INT.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * sd(bootstrap.est$a.INT.a)

# RERI
int.r$RERI.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RERI.lo))
int.r$RERI.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$RERI.up))

# multiplicative interaction
int.r$m.INT.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$m.INT.a))
int.r$m.INT.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 &
  int.r$A2 == 1]) +
  qnorm(0.975) * sd(bootstrap.est$m.INT.a))

```

Au final, on a :

	A2=0	A2=1	RD.A2 A1	RR.A2 A1
A1=0	$\$p_{\{00\}}\$=0.104$ [0.095,0.114]	$\$p_{\{01\}}\$=0.197$ [0.182,0.211]	0.092 [0.075,0.109]	1.89 [1.68,2.11]
A1=1	$\$p_{\{10\}}\$=0.405$ [0.379,0.431]	$\$p_{\{11\}}\$=0.891$ [0.87,0.913]	0.486 [0.453,0.52]	2.2 [2.05,2.36]
RD.A1 A2	0.301 [0.272,0.329]	0.695 [0.669,0.721]		
RR.A1 A2	3.89 [3.47,4.35]	4.54 [4.19,4.91]		

Note:

additive Interaction = 0.394 [0.357;0.431]

RERI = 3.78 [3.36;4.25]

multiplicative Interaction = 1.17 [1.02;1.33]

10.2 Estimation par Modèle Structurel Marginal

```
# On récupère les Y prédit précédents, que l'on fusionne
Y <- c(Y.A1_0.A2_0, Y.A1_1.A2_0, Y.A1_0.A2_1, Y.A1_1.A2_1)
length(Y)
# on aura une base de données de 40000 lignes

# On récupère les valeurs d'exposition qui ont servi dans les scénarios contrefactuels
# (garder le même ordre que pour les Y.A1.A2)
X <- rbind(subset(df.A1_0.A2_0, select = c("A1", "A2")),
           subset(df.A1_1.A2_0, select = c("A1", "A2")),
           subset(df.A1_0.A2_1, select = c("A1", "A2")),
           subset(df.A1_1.A2_1, select = c("A1", "A2")))
# dim(X)

## Modèle structurel marginal
msm.RD <- glm(Y ~ A1 + A2 + A1:A2,
              data = data.frame(Y,X),
              family = "gaussian") # ne pas ajuster sur les facteurs de confusion
msm.RD

## tableau des effets marginaux
results.MSM <- matrix(NA, ncol = 4, nrow = 4)
colnames(results.MSM) <- c("A2 = 0", "A2 = 1",
                           "RD within strata of A1",
                           "RR within strata of A1")
rownames(results.MSM) <- c("A1 = 0", "A1 = 1",
                           "RD within strata of A2",
                           "RR within strata of A2")

# 4 risques marginaux
results.MSM["A1 = 0", "A2 = 0"] <- msm.RD$coefficients["(Intercept)"]
results.MSM["A1 = 0", "A2 = 1"] <- msm.RD$coefficients["(Intercept)"] +
```

	A2 = 0	A2 = 1	RD within strata of A1	RR within strata of A1
A1 = 0	0.107	0.198	0.091	1.851
A1 = 1	0.411	0.889	0.478	2.164
RD within strata of A2	0.303	0.690	NA	NA
RR within strata of A2	3.834	4.483	NA	NA

Note:

additive Interaction = 0.387

multiplicative Interaction = 1.17

```

msm.RD$coefficients["A2"]
results.MSM["A1 = 1","A2 = 0"] <- msm.RD$coefficients["(Intercept)"] +
  msm.RD$coefficients["A1"]
results.MSM["A1 = 1","A2 = 1"] <- msm.RD$coefficients["(Intercept)"] +
  msm.RD$coefficients["A2"] + msm.RD$coefficients["A1"] + msm.RD$coefficients["A1:A2"]

# within strata of A2
results.MSM["RR within strata of A2", "A2 = 0"] <- results.MSM["A1 = 1","A2 = 0"] /
  results.MSM["A1 = 0","A2 = 0"]
results.MSM["RD within strata of A2", "A2 = 0"] <- results.MSM["A1 = 1","A2 = 0"] -
  results.MSM["A1 = 0","A2 = 0"]
results.MSM["RR within strata of A2", "A2 = 1"] <- results.MSM["A1 = 1","A2 = 1"] /
  results.MSM["A1 = 0","A2 = 1"]
results.MSM["RD within strata of A2", "A2 = 1"] <- results.MSM["A1 = 1","A2 = 1"] -
  results.MSM["A1 = 0","A2 = 1"]

# within strata of A1
results.MSM["A1 = 0", "RR within strata of A1"] <- results.MSM["A1 = 0","A2 = 1"] /
  results.MSM["A1 = 0","A2 = 0"]
results.MSM["A1 = 0", "RD within strata of A1"] <- results.MSM["A1 = 0","A2 = 1"] -
  results.MSM["A1 = 0","A2 = 0"]
results.MSM["A1 = 1", "RR within strata of A1"] <- results.MSM["A1 = 1","A2 = 1"] /
  results.MSM["A1 = 1","A2 = 0"]
results.MSM["A1 = 1", "RD within strata of A1"] <- results.MSM["A1 = 1","A2 = 1"] -
  results.MSM["A1 = 1","A2 = 0"]

results.MSM <- round(results.MSM,3)
RD.interaction <- msm.RD$coefficients["A1:A2"]
RR.interaction <- (results.MSM["A1 = 1","A2 = 1"] *
  results.MSM["A1 = 0","A2 = 0"]) /
  ( results.MSM["A1 = 0","A2 = 1"] *
    results.MSM["A1 = 1","A2 = 0"] )

```

Au final, on a (sans les IC):

10.3 Estimation avec TMLE

```
## 3- int.ltmleMSM()           pour estimer les différentes quantités d'intérêt,
##                             par gcomputation, IPTW ou tmle

int.ltmleMSM <- function(data = data,
                        Q_formulas = Q_formulas,
                        g_formulas = g_formulas,
                        Anodes = Anodes,
                        Lnodes = Lnodes,
                        Ynodes = Ynodes,
                        final.Ynodes = final.Ynodes,
                        SL.library = list(Q="SL.glm",
                                         g="SL.glm"),
                        gcomp = gcomp,
                        iptw.only = iptw.only,
                        survivalOutcome = FALSE,
                        variance.method = "ic",
                        B = 2000,
                        boot.seed = 12345) {

  # regime=
  # binary array: n x numAnodes x numRegimes of counterfactual treatment or a list of 'rule' functions
  regimes.MSM <- array(NA, dim = c(nrow(data), 2, 4)) # 2 variables d'exposition (A1, A2), 4 régimes
  regimes.MSM[, , 1] <- matrix(c(0,0), ncol = 2, nrow = nrow(data), byrow = TRUE) # exposé ni à A1, ni à A2
  regimes.MSM[, , 2] <- matrix(c(1,0), ncol = 2, nrow = nrow(data), byrow = TRUE) # exposé à A1 uniquement
  regimes.MSM[, , 3] <- matrix(c(0,1), ncol = 2, nrow = nrow(data), byrow = TRUE) # exposé à A2 uniquement
  regimes.MSM[, , 4] <- matrix(c(1,1), ncol = 2, nrow = nrow(data), byrow = TRUE) # exposé à A1 et A2

  # summary.measures = valeurs des coefficients du MSM associés à chaque régime
  # array: num.regimes x num.summary.measures x num.final.Ynodes -
  # measures summarizing the regimes that will be used on the right hand side of working.msm
  # (baseline covariates may also be used in the right hand side of working.msm and do not need to be included)
  summary.measures.reg <- array(NA, dim = c(4, 3, 1))
  summary.measures.reg[, , 1] <- matrix(c(0, 0, 0, # aucun effet ni de A1, ni de A2
                                          1, 0, 0, # effet de A1 isolé
                                          0, 1, 0, # effet de A2 isolé
                                          1, 1, 1), # effet de A1 + A2 + A1:A2
                                     ncol = 3, nrow = 4, byrow = TRUE)

  colnames(summary.measures.reg) <- c("A1", "A2", "A1:A2")

  if(gcomp == TRUE) {
    # test length SL.library$Q
    SL.library$Q <- ifelse(length(SL.library$Q) > 1, "SL.glm", SL.library$Q)
  }
}
```

```

# simplify SL.library$g because g functions are useless with g-computation
SL.library$g <- "SL.mean"

iptw.only <- FALSE
}

ltmle_MSM <- ltmleMSM(data = data,
                      Anodes = Anodes,
                      Lnodes = Lnodes,
                      Ynodes = Ynodes,
                      Qform = Q_formulas,
                      gform = g_formulas,
                      #deterministic.g.function = det.g,
                      regimes = regimes.MSM, # à la place de abar
                      working.msm= "Y ~ A1 + A2 + A1:A2",
                      summary.measures = summary.measures.reg,
                      final.Ynodes = final.Ynodes,
                      msm.weights = NULL,
                      SL.library = SL.library,
                      gcomp = gcomp,
                      iptw.only = iptw.only,
                      survivalOutcome = survivalOutcome,
                      estimate.time = FALSE,
                      variance.method = variance.method)

bootstrap.res <- data.frame("beta.Intercept" = rep(NA, B),
                           "beta.A1" = rep(NA, B),
                           "beta.A2" = rep(NA, B),
                           "beta.A1A2" = rep(NA, B))

if(gcomp == TRUE) {
  set.seed <- boot.seed

  for (b in 1:B){
    # sample the indices 1 to n with replacement
    bootIndices <- sample(1:nrow(data), replace=T)
    bootData <- data[bootIndices,]

    if ( round(b/100, 0) == b/100 ) print(paste0("bootstrap number ",b))

    boot_ltmle_MSM <- ltmleMSM(data = bootData,
                              Anodes = Anodes,
                              Lnodes = Lnodes,
                              Ynodes = Ynodes,

```

```

    Qform = Q_formulas,
    gform = g_formulas,
    #deterministic.g.function = det.g,
    regimes = regimes.MSM, # à la place de abar
    working.msm= "Y ~ A1 + A2 + A1:A2",
    summary.measures = summary.measures.reg,
    final.Ynodes = final.Ynodes,
    msm.weights = NULL,
    SL.library = SL.library,
    gcomp = gcomp,
    iptw.only = iptw.only,
    survivalOutcome = survivalOutcome,
    estimate.time = FALSE,
    variance.method = variance.method)

    bootstrap.res$beta.Intercept[b] <- boot_ltmle_MSM$beta["(Intercept)"]
    bootstrap.res$beta.A1[b] <- boot_ltmle_MSM$beta["A1"]
    bootstrap.res$beta.A2[b] <- boot_ltmle_MSM$beta["A2"]
    bootstrap.res$beta.A1A2[b] <- boot_ltmle_MSM$beta["A1:A2"]
  }
}

return(list(ltmle_MSM = ltmle_MSM,
           bootstrap.res = bootstrap.res))
}

### 4- summary.int()    pour enregistrer l'ensemble des estimations

summary.int <- function(data = data,
                        ltmle_MSM = ltmle_MSM,
                        estimator = c("gcomp", "iptw", "tmle")) {

  if(estimator == "gcomp") {
    try(if(ltmle_MSM$ltmle_MSM$gcomp == FALSE)
        stop("The ltmle function did not use the gcomp estimator, but the iptw +/- tmle estimator"))

    beta <- ltmle_MSM$ltmle_MSM$beta
  }

  if(estimator == "iptw") {
    try(if(ltmle_MSM$ltmle_MSM$gcomp == TRUE)
        stop("The ltmle function used the gcomp estimator, iptw is not available"))

    beta <- ltmle_MSM$ltmle_MSM$beta.iptw
  }
}

```

```

    IC <- ltmle_MSM$ltmle_MSM$IC.ipw
  }

  if(estimator == "tmle") {
    try(if(ltmle_MSM$ltmle_MSM$gcomp == TRUE) stop("The ltmle function used the gcomp"))

    beta <- ltmle_MSM$ltmle_MSM$beta
    IC <- ltmle_MSM$ltmle_MSM$IC
  }

  # on va enregistrer l'ensemble des résultats pertinent dans une table de longueur k1 :
  int.r <- matrix(NA,
                 ncol = 34,
                 nrow = nlevels(as.factor(data$A1)) * nlevels(as.factor(data$A2)))
  int.r <- as.data.frame(int.r)
  names(int.r) <- c("A1", "A2", "p", "sd.p", "p.lo", "p.up",
                  "RD.A1", "sd.RD.A1", "RD.A1.lo", "RD.A1.up",
                  "RD.A2", "sd.RD.A2", "RD.A2.lo", "RD.A2.up",
                  "RR.A1", "sd.lnRR.A1", "RR.A1.lo", "RR.A1.up",
                  "RR.A2", "sd.lnRR.A2", "RR.A2.lo", "RR.A2.up",
                  "a.INT", "sd.a.INT", "a.INT.lo", "a.INT.up", "RERI", "sd.lnRERI", "RE",
                  "m.INT", "sd.ln.m.INT", "m.INT.lo", "m.INT.up" )
  int.r[,c("A1", "A2")] <- expand.grid(c(0,1), c(0,1))

  # on peut retrouver les IC95% par delta method
  # A1 = 0 et A2 = 0
  int.r$p[int.r$A1 == 0 & int.r$A2 == 0] <- plogis(beta["(Intercept)"])

  # A1 = 1 et A2 = 0
  int.r$p[int.r$A1 == 1 & int.r$A2 == 0] <- plogis(beta["(Intercept)"] +
                                                  beta["A1"])

  # A1 = 0 et A2 = 1
  int.r$p[int.r$A1 == 0 & int.r$A2 == 1] <- plogis(beta["(Intercept)"] +
                                                  beta["A2"])

  # A1 = 1 et A2 = 1
  int.r$p[int.r$A1 == 1 & int.r$A2 == 1] <- plogis(beta["(Intercept)"] +
                                                  beta["A1"] +
                                                  beta["A2"] +
                                                  beta["A1:A2"])

  # RD.A1.A2 is 0
  int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0]

```

```

# RD.A1.A2is1
int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[

# RD.A2.A1is0
int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] - int.r$p[

# RD.A2.A1is1
int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[

# RR.A1.A2is0
int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$p[int.r$A1 == 1 & int.r$A2 == 0]) -

# RR.A1.A2is1
int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$p[int.r$A1 == 1 & int.r$A2 == 1]) -

# RR.A2.A1is0
int.r$RR.A2[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$p[int.r$A1 == 0 & int.r$A2 == 1]) -

# RR.A2.A1is1
int.r$RR.A2[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$p[int.r$A1 == 1 & int.r$A2 == 1]) -

# additive interaction
int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[
  int.r$p[int.r$A1 == 0 & int.r$A2 == 1] + int.r$p[int.r$A1 == 0 & int.r$A2 == 0]

# RERI
int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - i
  int.r$p[int.r$A1 == 0 & int.r$A2 == 1] +
  log(int.r$p[int.r$A1 == 0 & int.r$A2 == 0]))

# multiplicative interaction
int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$p[int.r$A1 == 1 & int.r$A2 == 1]) -
  log(int.r$p[int.r$A1 == 0 & int.r$A2 == 1]))

## IC95%
if(estimator == "iptw" | estimator == "tmle") {
  # A1 = 0 et A2 = 0
  grad <- c(int.r$p[int.r$A1 == 0 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
  v <- t(grad) %*% var(IC) %*% grad
  int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 0] <- sqrt(v / nrow(data))

  int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0] -
    qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 0]
  int.r$p.up[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0] +
    qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 0]

```

```

# A1 = 1 et A2 = 0
grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]) +
          int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]))
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 0] <- sqrt(v / nrow(data))

int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 0]
int.r$p.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 0]

# A1 = 0 et A2 = 1
grad <- c(int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]) +
          int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]))
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 1]
int.r$p.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 1]

# A1 = 1 et A2 = 1
grad <- rep(int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]) +
            int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]))
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 1]
int.r$p.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 1]

# RD.A1.A2is0
grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]) +
          int.r$p[int.r$A1 == 0 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 0]) +
          int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]))
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 0] <- sqrt(v / nrow(data))

int.r$RD.A1.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] -
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 0]
int.r$RD.A1.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] +
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 0]

# RD.A1.A2is1

```

```

grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RD.A1.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 1]
int.r$RD.A1.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 1]

# RD.A2.A1is0
grad <- c(int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.RD.A2[int.r$A1 == 0 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RD.A2.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 0 & int.r$A2 == 1]
int.r$RD.A2.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 0 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 0 & int.r$A2 == 1]

# RD.A2.A1is1
grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.RD.A2[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RD.A2.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 1 & int.r$A2 == 1]
int.r$RD.A2.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 1 & int.r$A2 == 1]

# RR.A1.A2is0
grad <- c(int.r$p[int.r$A1 == 0 & int.r$A2 == 0] - int.r$p[int.r$A1 == 1 & int.r$A2 == 0],
        1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0], 0, 0)
v <- t(grad) %*% var(IC) %*% grad

```

```

int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 0] <- sqrt(v / nrow(data))

int.r$RR.A1.lo[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0] -
                                                         qnorm(0.975) * int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 0]))
int.r$RR.A1.up[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 0] +
                                                         qnorm(0.975) * int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 0]))

# RR.A1.A2is1
grad <- c(int.r$p[int.r$A1 == 0 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 1],
          1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1],
          int.r$p[int.r$A1 == 0 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 1],
          1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1] )
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RR.A1.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 1] -
                                                         qnorm(0.975) * int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 1]))
int.r$RR.A1.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1 & int.r$A2 == 1] +
                                                         qnorm(0.975) * int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 1]))

# RR.A2.A1is0
grad <- c(int.r$p[int.r$A1 == 0 & int.r$A2 == 0] - int.r$p[int.r$A1 == 0 & int.r$A2 == 1],
          1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1], 0 )
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.lnRR.A2[int.r$A1 == 0 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RR.A2.lo[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0 & int.r$A2 == 1] -
                                                         qnorm(0.975) * int.r$sd.lnRR.A2[int.r$A1 == 0 & int.r$A2 == 1]))
int.r$RR.A2.up[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0 & int.r$A2 == 1] +
                                                         qnorm(0.975) * int.r$sd.lnRR.A2[int.r$A1 == 0 & int.r$A2 == 1]))

# RR.A2.A1is1
grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 0] - int.r$p[int.r$A1 == 1 & int.r$A2 == 1],
          int.r$p[int.r$A1 == 1 & int.r$A2 == 0] - int.r$p[int.r$A1 == 1 & int.r$A2 == 1],
          1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1],
          1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1])
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.lnRR.A2[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RR.A2.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1 & int.r$A2 == 1] -
                                                         qnorm(0.975) * int.r$sd.lnRR.A2[int.r$A1 == 1 & int.r$A2 == 1]))
int.r$RR.A2.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1 & int.r$A2 == 1] +
                                                         qnorm(0.975) * int.r$sd.lnRR.A2[int.r$A1 == 1 & int.r$A2 == 1]))

# additive interaction

```



```

grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.a.INT[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$a.INT.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.a.INT[int.r$A1 == 1 & int.r$A2 == 1]
int.r$a.INT.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.a.INT[int.r$A1 == 1 & int.r$A2 == 1]

# RERI
grad <- c((int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] + int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
        (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 0])),
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 1 & int.r$A2 == 0] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] + int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 0 & int.r$A2 == 1]
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] + int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] * (1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1]
        (int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 0]
        int.r$p[int.r$A1 == 0 & int.r$A2 == 1] + int.r$p[int.r$A1 == 0 & int.r$A2 == 0]
v <- t(grad) %*% var(IC) %*% grad
int.r$sd.lnRERI[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$RERI.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1]
        qnorm(0.975) * int.r$sd.lnRERI[int.r$A1 == 1 & int.r$A2 == 1]
int.r$RERI.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1]
        qnorm(0.975) * int.r$sd.lnRERI[int.r$A1 == 1 & int.r$A2 == 1]

# multiplicative interaction

```

```

grad <- c(int.r$p[int.r$A1 == 1 & int.r$A2 == 0] + int.r$p[int.r$A1 == 0 & int.r$A2 == 1] -
          int.r$p[int.r$A1 == 1 & int.r$A2 == 1] - int.r$p[int.r$A1 == 0 & int.r$A2 == 0] -
          int.r$p[int.r$A1 == 1 & int.r$A2 == 0] - int.r$p[int.r$A1 == 1 & int.r$A2 == 1] -
          int.r$p[int.r$A1 == 0 & int.r$A2 == 1] - int.r$p[int.r$A1 == 1 & int.r$A2 == 0] -
          1 - int.r$p[int.r$A1 == 1 & int.r$A2 == 1])
v <- t(grad) %% var(IC) %% grad
int.r$sd.ln.m.INT[int.r$A1 == 1 & int.r$A2 == 1] <- sqrt(v / nrow(data))

int.r$m.INT.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1] *
                                                         qnorm(0.975) * int.r$sd.ln.m.INT[int.r$A1 == 1 & int.r$A2 == 1])
int.r$m.INT.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1] *
                                                         qnorm(0.975) * int.r$sd.ln.m.INT[int.r$A1 == 1 & int.r$A2 == 1])

bootstrap.res <- ltmle_MSM$bootstrap.res
}

if(estimator == "gcomp") {
  ltmle_MSM$bootstrap.res$p.A1_0.A2_0 <- plogis(ltmle_MSM$bootstrap.res$beta.Intercept +
                                                  ltmle_MSM$bootstrap.res$beta.A1)
  ltmle_MSM$bootstrap.res$p.A1_1.A2_0 <- plogis(ltmle_MSM$bootstrap.res$beta.Intercept +
                                                  ltmle_MSM$bootstrap.res$beta.A1)
  ltmle_MSM$bootstrap.res$p.A1_0.A2_1 <- plogis(ltmle_MSM$bootstrap.res$beta.Intercept +
                                                  ltmle_MSM$bootstrap.res$beta.A2)
  ltmle_MSM$bootstrap.res$p.A1_1.A2_1 <- plogis(ltmle_MSM$bootstrap.res$beta.Intercept +
                                                  ltmle_MSM$bootstrap.res$beta.A1 +
                                                  ltmle_MSM$bootstrap.res$beta.A2 +
                                                  ltmle_MSM$bootstrap.res$beta.A1A2)

  ltmle_MSM$bootstrap.res$RD.A1.A2_0 <- ltmle_MSM$bootstrap.res$p.A1_1.A2_0 - ltmle_MSM$bootstrap.res$p.A1_0.A2_0
  ltmle_MSM$bootstrap.res$RD.A1.A2_1 <- ltmle_MSM$bootstrap.res$p.A1_1.A2_1 - ltmle_MSM$bootstrap.res$p.A1_0.A2_1
  ltmle_MSM$bootstrap.res$RD.A2.A1_0 <- ltmle_MSM$bootstrap.res$p.A1_0.A2_1 - ltmle_MSM$bootstrap.res$p.A1_0.A2_0
  ltmle_MSM$bootstrap.res$RD.A2.A1_1 <- ltmle_MSM$bootstrap.res$p.A1_1.A2_1 - ltmle_MSM$bootstrap.res$p.A1_0.A2_1

  ltmle_MSM$bootstrap.res$lnRR.A1.A2_0 <- log(ltmle_MSM$bootstrap.res$p.A1_1.A2_0 / ltmle_MSM$bootstrap.res$p.A1_0.A2_0)
  ltmle_MSM$bootstrap.res$lnRR.A1.A2_1 <- log(ltmle_MSM$bootstrap.res$p.A1_1.A2_1 / ltmle_MSM$bootstrap.res$p.A1_0.A2_1)
  ltmle_MSM$bootstrap.res$lnRR.A2.A1_0 <- log(ltmle_MSM$bootstrap.res$p.A1_0.A2_1 / ltmle_MSM$bootstrap.res$p.A1_0.A2_0)
  ltmle_MSM$bootstrap.res$lnRR.A2.A1_1 <- log(ltmle_MSM$bootstrap.res$p.A1_1.A2_1 / ltmle_MSM$bootstrap.res$p.A1_0.A2_1)

  ltmle_MSM$bootstrap.res$a.INT <- ltmle_MSM$bootstrap.res$p.A1_1.A2_1 -
                                   ltmle_MSM$bootstrap.res$p.A1_1.A2_0 -
                                   ltmle_MSM$bootstrap.res$p.A1_0.A2_1 +
                                   ltmle_MSM$bootstrap.res$p.A1_0.A2_0

  ltmle_MSM$bootstrap.res$lnRERI <- log((ltmle_MSM$bootstrap.res$p.A1_1.A2_1 -
                                         ltmle_MSM$bootstrap.res$p.A1_1.A2_0 -
                                         ltmle_MSM$bootstrap.res$p.A1_0.A2_1 +
                                         ltmle_MSM$bootstrap.res$p.A1_0.A2_0) /
                                         (ltmle_MSM$bootstrap.res$p.A1_1.A2_1 -
                                         ltmle_MSM$bootstrap.res$p.A1_1.A2_0 -
                                         ltmle_MSM$bootstrap.res$p.A1_0.A2_1 +
                                         ltmle_MSM$bootstrap.res$p.A1_0.A2_0))
}

```

```

ltml_MSM$bootstrap.res$p.A1_0.A2_0) / ltml_MSM$boot

ltml_MSM$bootstrap.res$ln.m.INT <- log((ltml_MSM$bootstrap.res$p.A1_1.A2_1 * ltml_MSM$boot
(ltml_MSM$bootstrap.res$p.A1_1.A2_0 * ltml_MSM$bo

# A1 = 0 et A2 = 0
int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 0] <- sd(ltml_MSM$bootstrap.res$p.A1_0.A2_0)
int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 0]
int.r$p.up[int.r$A1 == 0 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 0] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 0]

# A1 = 1 et A2 = 0
int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 0] <- sd(ltml_MSM$bootstrap.res$p.A1_1.A2_0)
int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 0]
int.r$p.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 0] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 0]

# A1 = 0 et A2 = 1
int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 1] <- sd(ltml_MSM$bootstrap.res$p.A1_0.A2_1)
int.r$p.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 1]
int.r$p.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 0 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 0 & int.r$A2 == 1]

# A1 = 1 et A2 = 1
int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltml_MSM$bootstrap.res$p.A1_1.A2_1)
int.r$p.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 1]
int.r$p.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$p[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.p[int.r$A1 == 1 & int.r$A2 == 1]

# RD.A1.A2is0
int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 0] <- sd(ltml_MSM$bootstrap.res$RD.A1.A2_0)
int.r$RD.A1.lo[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] -
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 0]
int.r$RD.A1.up[int.r$A1 == 1 & int.r$A2 == 0] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 0] +
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 0]

# RD.A1.A2is1
int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltml_MSM$bootstrap.res$RD.A1.A2_1)
int.r$RD.A1.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 1]
int.r$RD.A1.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A1[int.r$A1 == 1 & int.r$A2 == 1] +

```

```

qnorm(0.975) * int.r$sd.RD.A1[int.r$A1 == 1 & int.r$A2 == 1]

# RD.A2.A1is0
int.r$sd.RD.A2[int.r$A1 == 0 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$RD.A2.
int.r$RD.A2.lo[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 0 & int.r
qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 0 & int.r$A2 == 1]
int.r$RD.A2.up[int.r$A1 == 0 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 0 & int.r
qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 0 & int.r$A2 == 1]

# RD.A2.A1is1
int.r$sd.RD.A2[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$RD.A2.
int.r$RD.A2.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 1 & int.r
qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 1 & int.r$A2 == 1]
int.r$RD.A2.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$RD.A2[int.r$A1 == 1 & int.r
qnorm(0.975) * int.r$sd.RD.A2[int.r$A1 == 1 & int.r$A2 == 1]

# RR.A1.A2is0
int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 0] <- sd(ltmle_MSM$bootstrap.res$lnRR
int.r$RR.A1.lo[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1
qnorm(0.975) * int.r$sd.lnRR
int.r$RR.A1.up[int.r$A1 == 1 & int.r$A2 == 0] <- exp(log(int.r$RR.A1[int.r$A1 == 1
qnorm(0.975) * int.r$sd.lnRR

# RR.A1.A2is1
int.r$sd.lnRR.A1[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$lnRR
int.r$RR.A1.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1
qnorm(0.975) * int.r$sd
int.r$RR.A1.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A1[int.r$A1 == 1
qnorm(0.975) * int.r$sd

# RR.A2.A1is0
int.r$sd.lnRR.A2[int.r$A1 == 0 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$lnRR
int.r$RR.A2.lo[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0
qnorm(0.975) * int.r$sd.lnRR
int.r$RR.A2.up[int.r$A1 == 0 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 0
qnorm(0.975) * int.r$sd.lnRR

# RR.A2.A1is1
int.r$sd.lnRR.A2[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$lnRR
int.r$RR.A2.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1
qnorm(0.975) * int.r$sd.lnRR
int.r$RR.A2.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RR.A2[int.r$A1 == 1
qnorm(0.975) * int.r$sd.lnRR

# additive interaction

```

```

int.r$sd.a.INT[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$a.INT)
int.r$a.INT.lo[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.a.INT[int.r$A1 == 1 & int.r$A2 == 1]
int.r$a.INT.up[int.r$A1 == 1 & int.r$A2 == 1] <- int.r$a.INT[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.a.INT[int.r$A1 == 1 & int.r$A2 == 1]

# RERI
int.r$sd.lnRERI[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$lnRERI)
int.r$RERI.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.lnRERI[int.r$A1 == 1 & int.r$A2 == 1])
int.r$RERI.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$RERI[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.lnRERI[int.r$A1 == 1 & int.r$A2 == 1])

# multiplicative interaction
int.r$sd.ln.m.INT[int.r$A1 == 1 & int.r$A2 == 1] <- sd(ltmle_MSM$bootstrap.res$ln.m.INT)
int.r$m.INT.lo[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1] -
  qnorm(0.975) * int.r$sd.ln.m.INT[int.r$A1 == 1 & int.r$A2 == 1])
int.r$m.INT.up[int.r$A1 == 1 & int.r$A2 == 1] <- exp(log(int.r$m.INT[int.r$A1 == 1 & int.r$A2 == 1] +
  qnorm(0.975) * int.r$sd.ln.m.INT[int.r$A1 == 1 & int.r$A2 == 1])

bootstrap.res <- ltmle_MSM$bootstrap.res
}

return(list(int.r = int.r,
            bootstrap.res = bootstrap.res))
}

### Obtention du MSM par la fonction ltmle, estimation par gcomp, iptw ou tmle
# avec la fonction int.ltmleMSM()

# on définit les arguments de la fonction ltmleMSM du package ltmle
library(ltmle)
library(SuperLearner)

## arguments à renseigner
Q_formulas = c(Y="Q.kplus1 ~ L1 + L2 + L3 + A1 * A2") # useful to add A1 * A2 interaction here
g_formulas = c("A1 ~ L1 + L2",
               "A2 ~ L1 + L3")

SL.library = list(Q=list("SL.glm", c("SL.glm", "screen.corP"),
                        "SL.xgboost", "SL.rpartPrune", #"SL.randomForest",
                        "SL.step.interaction", c("SL.step.interaction", "screen.corP"),
                        "SL.glmnet", "SL.stepAIC",
                        "SL.mean"),
                  g=list("SL.glm", c("SL.glm", "screen.corP"),
                        "SL.xgboost", "SL.rpartPrune", #"SL.randomForest",

```



```

g_formulas = g_formulas,
Anodes = c("A1", "A2"),
Lnodes = c("L1", "L2", "L3"),
Ynodes = c("Y"),
final.Ynodes = "Y",
# SL.library = SL.library,
SL.library = list(Q="SL.glm.gaussian", #
                  g="SL.mean"),

gcomp = TRUE,          # si FALSE, fait tmle + IPTW
iptw.only = FALSE,
# si (gcomp = FALSE et iptw.only = TRUE), fait uniquement iptw
survivalOutcome = FALSE,
variance.method = "ic",
B = 1000, # nombre d'échantillons bootstrap
boot.seed = 54321) # seed pour l'échantillonnage bootstrap

### 3) Calcul des paramètres utiles pour l'analyse de l'interaction
# avec la fonction summary.int()

### récupération des résultats tmle
summary.tmle <- summary.int(data = df,
                           ltmle_MSM = interaction.ltmle,
                           estimator = c("tmle"))

# summary.tmle$int.r

### récupération des résultats iptw
summary.iptw <- summary.int(data = df,
                           ltmle_MSM = interaction.ltmle,
                           estimator = c("iptw"))

# summary.iptw$int.r

### récupération des résultats gcomputation
summary.gcomp <- summary.int(data = df,
                             ltmle_MSM = interaction.gcomp,
                             estimator = c("gcomp"))

# summary.gcomp$int.r
# head(summary.gcomp$bootstrap.res)
# # vérifier la normalité des estimations bootstrap
#   bootstrap.est <- subset(summary.gcomp$bootstrap.res,
#                             select =
#                               c("p.A1_0.A2_0",
#                                 "p.A1_1.A2_0",

```

```

# "p.A1_0.A2_1",
# "p.A1_1.A2_1",
# "RD.A1.A2_0",
# "RD.A1.A2_1",
# "RD.A2.A1_0",
# "RD.A2.A1_1",
# "lnRR.A1.A2_0",
# "lnRR.A1.A2_1",
# "lnRR.A2.A1_0",
# "lnRR.A2.A1_1",
# "a.INT",
# "lnRERI",
# "ln.m.INT"))
# par(mfrow = c(4,4))
# for(c in 1:ncol(bootstrap.est)) {
#   hist(bootstrap.est[,c], freq = FALSE, main = names(bootstrap.est)[c])
#   lines(density(bootstrap.est[,c]), col = 2, lwd = 3)
#   curve(1/sqrt(var(bootstrap.est[,c]) * 2 * pi) * exp(-1/2*((x-mean(bootstrap.e
#     col = 1, lwd = 2, lty = 2, add = TRUE)
# par(mfrow = c(1,1))
# }

```

Au final, on a (présentation selon recommandation Knol et al. Knol and VanderWeele [2012]):

TMLE

```

## $out.table
##
##                                     A2=0                                     A2=1
## A1=0      $p_{00}$=0.104 [0.095,0.113]  $p_{01}$=0.195 [0.18,0.21]
## A1=1      $p_{10}$=0.408 [0.378,0.439]  $p_{11}$=0.903 [0.88,0.927]
## RD.A1|A2          0.304 [0.272,0.336]          0.708 [0.68,0.737]
## RR.A1|A2          3.93 [3.5,4.41]          4.63 [4.55,4.72]
##
##          RD.A2|A1          RR.A2|A1
## A1=0      0.091 [0.073,0.109]  1.88 [1.67,2.11]
## A1=1      0.495 [0.457,0.534]  2.21 [2.04,2.4]
## RD.A1|A2
## RR.A1|A2
##
## $interaction.effects
## [1] "additive Interaction = 0.404 [0.362;0.447]"
## [2] "RERI = 3.89 [3.45;4.4]"
## [3] "multiplicative Interaction = 1.18 [1.02;1.36]"

```


IPTW

```
## $out.table
##
##                               A2=0                               A2=1
## A1=0      $p_{00}$=0.104 [0.095,0.113]  $p_{01}$=0.195 [0.18,0.21]
## A1=1      $p_{10}$=0.408 [0.377,0.439]  $p_{11}$=0.904 [0.88,0.927]
## RD.A1|A2      0.304 [0.272,0.336]      0.709 [0.68,0.737]
## RR.A1|A2      3.93 [3.5,4.41]      4.63 [4.55,4.72]
##
##                               RD.A2|A1                               RR.A2|A1
## A1=0      0.091 [0.073,0.109]  1.88 [1.67,2.11]
## A1=1      0.496 [0.457,0.534]  2.22 [2.05,2.4]
## RD.A1|A2
## RR.A1|A2
##
## $interaction.effects
## [1] "additive Interaction = 0.405 [0.362;0.447]"
## [2] "RERI = 3.9 [3.45;4.4]"
## [3] "multiplicative Interaction = 1.18 [1.02;1.36]"
```

G-computation

```
## $out.table
##
##                               A2=0                               A2=1
## A1=0      $p_{00}$=0.104 [0.095,0.112]  $p_{01}$=0.197 [0.183,0.211]
## A1=1      $p_{10}$=0.4 [0.373,0.427]  $p_{11}$=0.893 [0.872,0.914]
## RD.A1|A2      0.296 [0.268,0.325]      0.697 [0.671,0.722]
## RR.A1|A2      3.86 [3.46,4.31]      4.54 [4.46,4.61]
##
##                               RD.A2|A1                               RR.A2|A1
## A1=0      0.093 [0.077,0.11]  1.9 [1.7,2.12]
## A1=1      0.494 [0.46,0.527]  2.23 [2.08,2.4]
## RD.A1|A2
## RR.A1|A2
##
## $interaction.effects
## [1] "additive Interaction = 0.4 [0.363;0.438]"
## [2] "RERI = 3.86 [3.46;4.32]"
## [3] "multiplicative Interaction = 1.18 [1.03;1.34]"
```


Chapter 11

Représentations graphiques

Part III

En pratique

Chapter 12

Proposition d'étapes

1. Formuler l'objectif

- Est-ce un objectif prédictif ou explicatif ?
- Si démarche explicative, s'agit-il plutôt d'une analyse d'interaction ou de modification d'effet?

2. Stratégies et méthodes

- Poser les hypothèses sur un DAG ou schéma conceptuel
- Identifier le ou les estimand(s), c'est-à-dire l'effet ou le paramètre que l'on va chercher à estimer pour répondre à l'objectif, par exemple :
 - effet conjoint de X et V sur Y, sur l'échelle multiplicative = $OR_{X,V}$
 - effet de X sur Y dans chaque strate de V, sur l'échelle additive = $DR_{X|V=0}$ et $DR_{X|V=1}$
 - effet d'interaction sur l'échelle additive et multiplicative AI et MI
- Elaborer l'estimateur, notamment :
 - quelles est(sont) l'exposition(s) d'intérêt ?
 - quels sont les facteurs de confusion +/- les médiateurs si besoin ?
 - quels types de modélisation va être utilisée ?

3. Analyses descriptives

- Description habituelle de la population
- Décrire, dans un tableau croisé,
 - le Y moyen ou la proportion de $Y = 1$
 - pour chaque catégorie de X et V

4. Analyses exploratoires

- Analyses stratifiées
 - pour une analyse de modification d'effet,
 - il est possible en exploratoire, d'estimer l'effet de X sur Y
 - de façon stratifiée sur V (on découpe la population)
 - les effets ne seront directement pas comparables
- Régressions avec terme d'interaction
 - un modèle dans la population totale peut être utilisé
 - avec un terme d'interaction entre X et V
 - les différents paramètres peuvent être déduits des résultats du modèle
 - voir Chapitre 9
- Marges/ Effets prédits
 - .

5. **Analyses confirmatoires**, voir Chapitre 10

- G-computation
- MSM
- TMLE

Chapter 13

Exemple 1 - Y binaire

Chapter 14

Exemple 2 - Y quantitatif

14.1 1. Formuler les objectifs

Dans cette étude Colineaux et al. [2023], on s'est intéressé à :

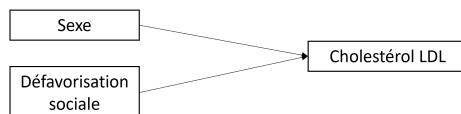
- comment l'effet du sexe sur le taux de cholestérol LDL vers 45 ans varie en fonction de la défavorisation sociale précoce,
- comment l'effet de la défavorisation sociale précoce sur le taux de cholestérol LDL varie en fonction du sexe.

La démarche ici est explicative : on cherche à comprendre des mécanismes causaux.

A partir de la formulation des objectifs, on pourrait dire qu'on s'intéresse ici plutôt à deux modifications d'effet. On va donc devoir à la fois agir sur le sexe $do(S)$ et sur la défavorisation sociale $do(D)$. Donc la démarche, en fait, sera plutôt une analyse d'interaction $do(S, D)$

14.2 2. Stratégies et méthodes

Le DAG (sans les médiateurs) était :



Les **estimands** étaient définis sur l'échelle additive par :

Sexe	Défavorisation	Mean(Chol LDL)
Male	Non	3.57
Male	Oui	3.60
Female	Non	3.24
Female	Oui	3.37

- La modification de l'effet du sexe en fonction par la défavorisation sociale précoce :

$$\begin{aligned}
& - (Y_{s=1|d=0} - Y_{s=0|d=0}) - (Y_{s=1|d=1} - Y_{s=0|d=1}) \\
& - \text{ou } (Y_{s=1,d=0} - Y_{s=0,d=0}) - (Y_{s=1,d=1} - Y_{s=0,d=1})
\end{aligned}$$

- La modification de l'effet de la défavorisation sociale précoce par la sexe

$$\begin{aligned}
& - (Y_{d=1|s=0} - Y_{d=0|s=0}) - (Y_{d=1|s=1} - Y_{d=0|s=1}) \\
& - \text{ou } (Y_{d=1,s=0} - Y_{d=0,s=0}) - (Y_{d=1,s=1} - Y_{d=0,s=1})
\end{aligned}$$

Les deux formulations sont ici équivalentes car il n'y pas de facteurs de confusion, donc, par exemple, $Y_{d=1|s=0} = Y_{s=0|d=1} = Y_{d=1,s=0}$

L'estimateur : Les effets ont été estimés par g-computation (*standardisation par régression*) Hernán and Robins [2020]. Des régressions linéaires ont été utilisées pour estimer les *potential outcomes* pour chaque scénario, désignées par $\bar{Q}(S, D) = E(Y|S, D)$. A partir des fonctions $\bar{Q}(S, D)$ estimées, nous avons prédit la valeur de l'outcome Y pour chaque individu i pour chaque scénario. Les valeurs moyennes de Y dans chaque scénario vont ensuite nous permettre d'estimer les estimands selon leurs définitions précisées ci-dessus. Ces modèles $\bar{Q}(S, D)$ vont comprendre 2 variables : le sexe et la défavorisation sociale précoce (il n'y a pas ici de facteurs de confusion).

14.3 3.Analyse descriptive

Dans cette population (N=17 272), il y avait 51,4% d'hommes et 60,5% de personnes ayant été précocement défavorisées.

On peut commencer par décrire les moyennes de cholestérol dans chaque catégorie de sexe et de défavorisation sociale :

14.4 4.Analyse exploratoire

La sortie d'une modèle linéaire simple serait :

```
# Call:
# lm(formula = t8_ldl ~ as.factor(sex) + as.factor(soc_group) +
#     as.factor(sex) * as.factor(soc_group), data = ba_1)
#
# Residuals:
#      Min       1Q   Median       3Q      Max
# -3.1996 -0.6427 -0.0688  0.5318  4.6312
#
# Coefficients:
#                                     Estimate Std. Error t value Pr(>|t|)
# (Intercept)                        3.24270     0.01594  203.475 < 2e-16 ***
# as.factor(sex)1                     0.32553     0.02227   14.616 < 2e-16 ***
# as.factor(soc_group)2.Défau         0.12614     0.02052    6.148 8.02e-10 ***
# as.factor(sex)1:as.factor(soc_group)2.Défau -0.09473     0.02863   -3.308 0.000941 ***
# ---
# Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
#
# Residual standard error: 0.9194 on 17268 degrees of freedom
# Multiple R-squared:  0.0231, Adjusted R-squared:  0.02293
# F-statistic: 136.1 on 3 and 17268 DF, p-value: < 2.2e-16
```

On peut en déduire (échelle additive) que :

- L'effet du sexe (d'être homme plutôt que femme) est :
 - Quand on est favorisé : $DR(S|D = 0) = +0.326$ mmol/L
 - Quand on est défavorisé : $DR(S|D = 1) = 0.326 - 0.095 = +0.231$ mmol/L
- L'effet de la défavorisation est :
 - Quand on est une femme : $DR(D|S = 0) = +0.126$ mmol/L
 - Quand on est un homme : $DR(D|S = 1) = 0.126 - 0.095 = +0.031$ mmol/L
- L'effet d'être un homme et défavorisé
 - plutôt que femme et favorisé est
 - $DR(D, S) = 0.326 + 0.126 - 0.095 = +0.357$ mmol/L
- **L'effet d'interaction/modification d'effet** est : $AI = -0.095$ mmol/L

On pourrait aussi déduire :

- $Y_{00} = 3.24$ mmol/L
- $Y_{10} = 3.243 + 0.326 = 3.57$ mmol/L
- $Y_{01} = 3.243 + 0.126 = 3.37$ mmol/L
- $Y_{11} = 3.243 + 0.326 + 0.126 - 0.095 = 3.6$ mmol/L

	A2=0	A2=1	RD.A2 A1
A1=0	$\$p_{\{00\}}\$=3.243$ [3.213,3.273]	$\$p_{\{01\}}\$=3.369$ [3.344,3.394]	0.126 [0.087,0.165]
A1=1	$\$p_{\{10\}}\$=3.568$ [3.538,3.598]	$\$p_{\{11\}}\$=3.6$ [3.574,3.625]	0.031 [-0.008,0.071]
RD.A1 A2	0.326 [0.283,0.368]	0.231 [0.195,0.267]	

^a additive Interaction = -0.095 [-0.15;-0.039]

14.5 5.Analyse confirmatoire

Si l'on utilise le package proposé par B Lepage pour réaliser cet analyse avec la TMLE (effets d'interaction calculés à partir des paramètres d'une modèle structurel marginal estimé à l'aide du package R ltmle), les résultats sont :

On retrouve des résultats qui peuvent être interprétés ainsi :

- l'effet d'être un homme (ou "la différence H-F) est moins fort de additive Interaction = -0.095 [-0.15;-0.039] mmol/L lorsqu'on est défavorisé précocement
- l'effet de la défavorisation est moins fort de additive Interaction = -0.095 [-0.15;-0.039] mmol/L chez les hommes

En réalité, on a réalisé cette analyse par g-computation (voir chapitre 10) sur des données imputées et bootstrappées (l'exemple ci-dessus a été réalisé sur une seule des bases bootstrappées, ce qui explique les différences), et les résultats, présentés selon les recommandations modifiées de Knol et VanderWeele, étaient:

	Né-e-s-Avantagé-e-s		Né-e-s désavantagé-e-s		ET du désavantage précoce	
Nés-Hommes (moyenne)	3,48	[3,44 à 3,52]	3,49	[3,45 à 3,52]	+0,01	[-0,04 à 0,05]
Nés-Femmes (moyenne)	3,24	[3,20 à 3,28]	3,33	[3,29 à 3,36]	+0,09	[0,04 à 0,13]
ET d'être né homme	+0,24	[0,19 à 0,29]	+0,16	[0,12 à 0,20]	(-0,07)	[-0,13 à -0,02]

Chapter 15

Exemple 3 - Y multinomial

Chapter 16

Exemple 4 - X quantitatif

Part IV

Conclusion

Chapter 17

Synthèse générale

La première étape importantes consiste à **définir précisément l’objectif**. Et, si l’on est dans une démarche explicative, d’inférence causale, il s’agit de définir si la mesure d’un effet d’interaction est nécessaire pour y répondre (identifier précisément l’effet que l’on cherche à estimer, ou *estimand*).

Le fait de choisir une **démarche d’analyse d’interaction ou de modification d’effet** repose sur :

- la façon dont la question est posée (effet de X selon V ou effet conjoint de X et V),
- sur les hypothèses causales formulées (scénarii $do(X)$ ou $do(X, V)$)
- et donc sur les sets de facteurs de confusion à considérer (seulement sur $X \rightarrow Y$ ou $X.V \rightarrow Y$).

Concernant le **choix de l’échelle**, idéalement, les interactions devraient être reportées sur les 2 échelles Knol and VanderWeele [2012] VanderWeele and Knol [2014]. Cependant, l’échelle additive est plus appropriée pour évaluer l’utilité en santé publique VanderWeele and Knol [2014] Knol and VanderWeele [2012].

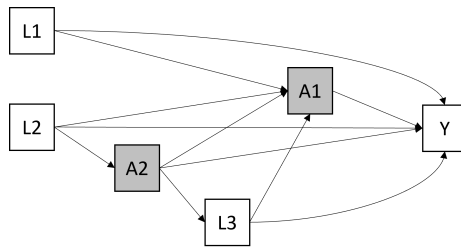
Concernant les paramètres,

Chapter 18

Pour aller plus loin...

18.1 Ajouter de la complexité

A1 et A2 sont rarement indépendants. Scénario plus probable :



18.2 Interaction avec confusion intermédiaire

18.3 Interaction et médiation

VanderWeele [2013]

VanderWeele [2014]

Chapter 19

Références

Bibliography

- Hélène Colineaux, Lola Neufcourt, Cyrille Delpierre, Michelle Kelly-Irving, and Benoit Lepage. Explaining biological differences between men and women by gendered mechanisms. *Emerging Themes in Epidemiology*, 20(1):2, 2023.
- Priscila Corraini, Morten Olsen, Lars Pedersen, Olaf M Dekkers, and Jan P Vandenbroucke. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. *Clinical Epidemiology*, 9:331–338, June 2017. ISSN 1179-1349. doi: 10.2147/CLEP.S129728.
- Miguel A Hernán and James M Robins. *Causal Inference: What If - PREPRINT*. Chapman & Hall/CRC Boca Raton, FL, 2020.
- Miguel A Hernán, John Hsu, and Brian Healy. A second chance to get causal inference right: a classification of data science tasks. *Chance*, 32(1):42–49, 2019.
- Mirjam J. Knol and Tyler J. VanderWeele. Recommendations for presenting analyses of effect modification and interaction. *International Journal of Epidemiology*, 41(2):514–520, April 2012. ISSN 1464-3685. doi: 10.1093/ije/dyr218.
- Maya B Mathur and Tyler J VanderWeele. R function for additive interaction measures. *Epidemiology (Cambridge, Mass.)*, 29(1):e5, 2018.
- Tyler J. VanderWeele. On the distinction between interaction and effect modification. *Epidemiology (Cambridge, Mass.)*, 20(6):863–871, November 2009. ISSN 1531-5487. doi: 10.1097/EDE.0b013e3181ba333c.
- Tyler J. VanderWeele. A Three-way Decomposition of a Total Effect into Direct, Indirect, and Interactive Effects. *Epidemiology*, 24(2):224–232, March 2013. ISSN 1044-3983. doi: 10.1097/EDE.0b013e318281a64e.
- Tyler J. VanderWeele. A unification of mediation and interaction: a 4-way decomposition. *Epidemiology (Cambridge, Mass.)*, 25(5):749–761, September 2014. ISSN 1531-5487. doi: 10.1097/EDE.0000000000000121.
- Tyler J. VanderWeele. The Interaction Continuum. *Epidemiology*, 30(5):648–658, September 2019. ISSN 1044-3983. doi: 10.1097/EDE.0000000000001054.

Tyler J. VanderWeele and Mirjam J. Knol. A Tutorial on Interaction. *Epidemiologic Methods*, 3(1):33–72, December 2014. ISSN 2161-962X. doi: 10.1515/em-2013-0005. Publisher: De Gruyter.

Tyler J. VanderWeele and James M. Robins. Four types of effect modification: a classification based on directed acyclic graphs. *Epidemiology (Cambridge, Mass.)*, 18(5):561–568, September 2007. ISSN 1044-3983. doi: 10.1097/EDE.0b013e318127181b.