



**HAL**  
open science

# Analyse causale de mediation multiple et applications

Allan Jérolon

► **To cite this version:**

Allan Jérolon. Analyse causale de mediation multiple et applications. Mathématiques générales [math.GM]. Université Paris Cité, 2020. Français. NNT : 2020UNIP7047 . tel-03194558

**HAL Id: tel-03194558**

**<https://theses.hal.science/tel-03194558v1>**

Submitted on 9 Apr 2021

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

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



## UNIVERSITÉ DE PARIS

ÉCOLE DOCTORALE DE SCIENCES MATHÉMATIQUES DE PARIS CENTRE (ED386)  
LABORATOIRE DE MATHÉMATIQUES APPLIQUÉES À PARIS 5 (MAP5 - UMR CNRS  
8145)

---

# Analyse causale de médiation multiple et applications

---

*Par Allan JÉROLON*

THÈSE DE DOCTORAT DE MATHÉMATIQUES APPLIQUÉES,  
DIRIGÉE PAR Étienne BIRMELÉ,  
CO-ENCADRÉE PAR Flora ALARCON ET Vittorio PERDUCA,

PRÉSENTÉE ET SOUTENUE PUBLIQUEMENT LE 04 DÉCEMBRE 2020,

DEVANT UN JURY COMPOSÉ DE :

Anne PHILIPPE	Professeure à l'Université de Nantes,	Rapporteuse.
Nicolas SAVY	Maître de conférence à l'Université de Toulouse 1,	Rapporteur.
Antoine CHAMBAZ	Professeur à l'Université de Paris,	Examinateur.
Vivian VIALLO	Maître de conférence à l'Université Claude Bernard Lyon 1,	Examinateur.
Agathe GUILLOUX	Professeure à l'Université d'Évry Val d'Essonne,	Présidente du Jury.
Flora ALARCON	Maître de conférence à l'Université de Paris,	Co-encadrante.
Vittorio PERDUCA	Maître de conférence à l'Université de Paris,	Co-encadrant.
Étienne BIRMELÉ	Professeur à l'Université de Paris,	Directeur de thèse.

**Titre :** Analyse causale de médiation multiple et applications.

**Résumé :** L'analyse de médiation causale est une méthode statistique permettant d'étudier les mécanismes définissant les relations entre trois entités, la première étant une cause, la deuxième une variable de réponse et la troisième un ensemble de variables intermédiaires, aussi appelés médiateurs. Cette approche est devenue populaire dans différents domaines tels que les biostatistiques, l'épidémiologie et les sciences sociales. Son objectif est de décomposer et quantifier l'effet total d'une cause  $T$  sur une variable de réponse  $Y$  en un effet direct et un effet indirect à travers le(s) médiateur(s)  $M_i$ . Dans un diagramme causal ainsi défini peuvent intervenir des covariables mesurées ou non ayant des effets sur les trois variables  $T$ ,  $Y$  et  $M_i$ .

Quand plusieurs médiateurs sont impliqués dans le modèle de médiation, il y a trois cas possibles de relation entre ces médiateurs : 1) Les médiateurs sont indépendants conditionnellement au traitement et aux covariables mesurées. 2) Les médiateurs sont dépendants conditionnellement au traitement et aux covariables mesurées sans pour autant que l'un affecte l'autre. 3) Les médiateurs sont liés de façon causale, c'est-à-dire que l'un affecte l'autre.

Les contributions de cette thèse portent sur les deux premières configurations.

Plus particulièrement, l'état de l'art pour l'estimation des effets indirects individuels des médiateurs est de faire plusieurs analyses de médiation simple en parallèle, un médiateur à la fois. Cette approche conduit à une estimation biaisée de l'effet direct et n'est pas justifiée dans la situation 2) car l'hypothèse nécessaire d'Ignorabilité Séquentielle n'est alors pas vérifiée. Le premier chapitre de la thèse propose de résoudre ce problème par l'estimation de la loi multivariée des médiateurs conditionnellement au traitement et décrit les nouvelles hypothèses sous lesquelles une telle approche est valide. Une méthode d'estimation de l'effet direct et de tous les effets indirects individuels est développée, prenant en compte la corrélation entre les médiateurs. Le second chapitre est dédié à l'analyse de médiation multiple dans le cas où le nombre de médiateurs grandit. Deux stratégies sont explorées. La première repose sur une classification spectrale des médiateurs en blocs indépendants, chaque bloc pouvant être analysé avec l'approche précédente. La seconde, applicable dans des problèmes de grande dimension, repose sur la sélection de médiateurs potentiels à l'aide d'une approche de type Group-LASSO en amont de l'analyse de médiation. Le troisième chapitre présente le package R `multimediate` où les trois méthodes précédentes sont implémentées. On y présente l'algorithme d'estimation en fonction des différents types de médiateurs considérés (linéaires, binaires et catégoriels). Le dernier chapitre présente des applications sur des données réelles. D'une part une application en épidémiologie pour quantifier et décomposer l'effet de la thérapie hormonal sur le cancer du sein. D'autre part une application en psychiatrie, où nous analysons les effets des tempéraments affectifs sur l'anorexie mentale.

**Mots clés :** Analyse causale de médiation ; Ignorabilité séquentielle ; Médiateurs multiples ; Grande dimension ; Classification spectrale ; Sélection de variable.

**Title :** Causal mediation analysis with multiple mediators and applications.

**Abstract :** Causal mediation analysis is a statistical method used to study the mechanisms defining the relationships between three entities, the first being a cause, the second a response variable and the third a set of intermediate variables, also called mediators. This approach has become popular in various domains such as biostatistics, epidemiology and social sciences. The goal of such an analysis is to decompose and quantify the total effect of a cause  $T$  on a outcome  $Y$  into a direct effect and an indirect effect through the mediator(s)  $M_i$ . In a causal diagram defined in this way, measured or unmeasured covariates can intervene with effects on the three variables  $T$ ,  $Y$  and  $M_i$ .

When several mediators are involved in the mediation model, there are three possible situations : 1) The mediators are independent conditionally on the treatment and on the measured covariates. 2) Mediators are conditionally dependent on the treatment and on the measured covariates but this dependence is not of causal nature (i.e. mediators do not cause each other). 3) Mediators are causally related, that is, one affects the other.

The contributions of this thesis focus on the first two configurations.

More particularly, the state of the art for estimating the individual indirect effects of mediators is to perform several simple mediation analyzes in parallel, one mediator at a time. This approach leads to a biased estimate of the direct effect and is not justified in situation 2) because the necessary assumption of Sequential Ignorability is then not verified. The first chapter of the thesis proposes to solve this problem by estimating the multivariate distribution of mediators conditional on the treatment and describes the new hypotheses under which such an approach is valid. A method for estimating the direct effect and all individual indirect effects is developed, taking into account the correlation between the mediators. The second chapter is dedicated to the analysis of multiple mediation in the case where the number of mediators grows. Two strategies are explored. The first is based on a spectral classification of mediators into independent blocks, each block can be analyzed with the previous approach. The second strategy, applicable in large-scale problems, is based on the selection of potential mediators using a Group-LASSO type approach upstream of the mediation analysis. The third chapter presents the R `multimediate` package where the three previous methods are implemented. We present the estimation algorithm according to the different types of mediators considered (linear, binary and categorical). The last chapter presents applications on real data. On the one hand, I illustrate an application in epidemiology to quantify and disentangling the effect of hormonal therapy on breast cancer. On the other hand, I present an application in psychiatry, where I analyze the effects of affective temperaments on anorexia nervosa.

**Keywords :** Causal mediation analysis ; Sequential ignorability ; Multiple mediators ; High dimensional data ; Spectral clustering ; Variable selection.



# Remerciements

Je souhaite tout d'abord remercier mon directeur et mes co-encadrants de thèse Étienne Birmelé, Flora Alarcon et Vittorio Perduca qui m'ont permis de prendre part au monde de la recherche. Je tiens surtout à vous remercier pour vos conseils et votre encadrement qui m'ont permis d'améliorer mon travail tout au long de ma thèse.

Je remercie sincèrement Anne Philippe et Nicolas Savy pour avoir accepté de rapporter cette thèse, ainsi que Antoine Chambaz, Agathe Guilloux et Vivian Viallon, de constituer mon jury de thèse.

Mon travail de recherche a aussi été possible grâce aux financements accordés par la DIM-Math Innov, la FSMP et la Région Guadeloupe que je remercie sincèrement.

Le cadre de travail agréable que fournit le laboratoire du MAP5 sous la direction de Fabienne Comte puis d' Anne Estrade a énormément apporté. Le travail de Marie-Hélène Guabidi, Sandrine Fallu, Augustin Hangat et Isabelle Valéro ont grandement facilité les démarches administratives. Je remercie aussi Rémy Abergel, Max Paisley et Azedine Mani de m'avoir aidé dans l'utilisation des ressources informatiques.

Des applications concrètes de mon travail théorique étaient pour moi bien plus que nécessaires. Je remercie Laura Baglietto et Gianluca Severi pour les données épidémiologiques, ainsi que Enrica Marzola pour les données du domaine psychiatrique

Le domaine de la recherche permet de nombreux échanges enrichissant avec d'autres chercheurs. Je tiens à remercier pour leurs conseils, nos échanges et les réponses à certaines de mes interrogations, Céline Duval, Éric Lucon, Linda Valeri et Stijn Vansteelandt. De plus j'ai énormément apprécié les retraites scientifiques, organisées avec Flora Alarcon, Olivier Bouaziz, Vivien Goepf, Alexandra Lefebvre, Sabine Mercier, Grégory Nuel et Vittorio Perduca, qui se sont révélées très utiles pour ma recherche.

Parmi les échanges enrichissants j'ai beaucoup apprécié les discussions informelles, que j'ai eu avec Warith Harchaoui, qui ont souvent été productives, notamment concernant la classification spectrale. Je te remercie pour ton implication et ton sérieux ainsi que l'intérêt (scientifique) que tu portes aux autres. Je tiens aussi à remercier Noura Darwish pour son travail préliminaire sur l'analyse de médiation en grande dimension.

J'ai aussi pu comprendre l'importance d'adapter mon travail aux besoins du domaine applicatif grâce aux discussions organisées avec l'équipe médiation composée de Morgane Ballon, Jeanna-Eve Franck, Barbara Heude, Sandrine Lioret, Maria Melchior, Gwenn Menvielle, Aurélie Nakamura et Judith Van der Waerden.

Un très grand merci à Alkeos Michail, Fabien Vergnet et Vivien Geopp pour leur merveilleux accueil et pour avoir répondu à mes différentes interrogations en tant que nouveau doctorant. Nos différentes discussions m'ont souvent permis de décompresser et de trouver des solutions à certains problèmes mathématiques, administratifs et de programmation.

Je tiens aussi à remercier l'ensemble des doctorants, ATER et post-doctorants pour leur enthousiasme et notre groupe de travail des thésards et jeunes docteurs, où j'ai pu échanger avec Andrea Bondesan, Alessandro Cucchi, Marta Marulli, Pierre-Louis Antonsanti, Valentin De Bortoli, Anton François, Mario Gonzales, Rémi Laumont, Rémi Laumont, Claire Launay, Mariono Rodriguez, Alexandre Saint-Dizier, Vincent Vidal, Safa Ladgham, Frank Trujillo, Juliana Alves Pegoraro, Sinda Ammous, Emmanuel Caron, Florian Dussap, Vivien Goepp, Warith Harchaoui, Arthur Leroy, Quentin Marcou, Thanh Yen Nguyen, Ousmane Sacko.

J'ai aussi pu prendre part au monde de l'enseignement. Je tiens à remercier ceux avec qui j'ai pu travailler et qui m'ont permis d'apprécier cette partie de mon travail. Je pense aux membres de l'équipe d'enseignement du DUT Informatique de l'IUT de Paris Descartes dont Anne Estrade, Bérénice Grec, Sébastien Martin et Mélanie Gobert. Ainsi que l'équipe d'enseignement du département Mathématiques et Informatique de l'Université de Paris, dont Nicolas Loménié et Sylvain Lobry avec qui j'ai pu collaborer.

L'accomplissement de cette thèse a aussi été possible par le travail, l'encadrement et le soutien de personnes que j'ai rencontré bien avant le début de mon intérêt pour l'analyse de médiation causale. Un grand merci à Ken Dondin, à Marylène Troupé, Maître de conférence à l'Université des Antilles et à l'équipe d'enseignement de la classe préparatoire de BCPST du Lycée Général Technologique de Baimbridge, pour ne citer que ceux là.

Je remercie aussi mes proches et amis pour leur soutien et intérêt à mon égard, je pense notamment à Florian Anthoney, Néfertiti Bellony, Dany Castelneau, Sébastien Charon, Stanley Estilaire, Patience Longui, Emerick Mazaniello-Chezol, Daryl Naigre, Kara Ranguin-Luce, Jordan Sanchez, Maxime Szymanek et ma fiancée Cécilia Vergé-Dépré.

Et enfin mais pas des moindres je remercie mes parents, en particulier ma mère Marie-Aimée Jérolon, ma tante Clémence Jérolon et mes défunts grands-parents Justinien Jérolon et Gisèle Jérolon, pour leurs profondes implications dans mon éducation morale et scolaire et pour leurs rôles dans mes accomplissements.

La liste des personnes remerciées ici n'est point exhaustive. Il en manque beaucoup d'autres (famille, amis, collègues, enseignants, etc.) et je ne pourrai citer tout le monde. Toute personne (même morale), à qui de droit, y a bien évidemment sa place.

# Liste des symboles

## Lois de probabilités

$\mathcal{L}(\mu, s)$  loi logistique de moyenne  $\mu$  et de variance  $\frac{\pi^2}{3}s^2$

$\mathcal{N}(\mu, \sigma^2)$  loi normale de moyenne  $\mu$  et de variance  $\sigma^2$

$\Phi$  la fonction de distribution de la loi normale  $\mathcal{N}(0, 1)$

## Notations

$A^{\Gamma j}$  la transposée de la  $j$ -ième ligne d'une matrice  $A$

$A^{\Gamma}$  la transposée d'une matrice ou d'un vecteur  $A$

## Quantités causales

$\delta$  l'effet indirect naturel de  $T$  sur  $Y$ , médié par  $M$

$\delta^k$  l'effet indirect de  $T$  sur  $Y$ , médié par  $M^k$

$\delta^Z$  l'effet indirect de  $T$  sur  $Y$ , médié par l'ensemble des médiateurs du vecteur  $Z$

$\tau$  l'effet total

$\zeta$  l'effet direct naturel de  $T$  sur  $Y$

$\zeta^c$  l'effet direct contrôlé de  $T$  sur  $Y$

$PM^k$  la proportion médiée par  $M^k$

## Variables

$M^k$  le  $k$ -ième médiateur du vecteur  $Z$

$T$  le traitement ou l'exposition dans le modèle de causalité.

$W^k$  le complémentaire de  $M^k$  dans  $Z$

$X$  le vecteur des covariables de pré-traitement

$Y$  la variable de réponse du modèle de causalité.

$Z$  le vecteur de l'ensemble des médiateurs





# Table des matières

<b>Introduction</b>	<b>13</b>
0.1 Méthodes basée sur les modèles à équations linéaires structurelles (LSEM)	13
0.2 Notion des contrefactuels . . . . .	14
0.3 Analyse de médiation simple . . . . .	16
0.4 Problématiques . . . . .	21
0.5 Organisation de la thèse . . . . .	23
<b>1 Médiateurs liés de façon non-causale</b>	<b>27</b>
1.1 Introduction . . . . .	27
1.2 Définition des effets . . . . .	29
1.3 Hypothèses . . . . .	30
1.4 Identifiabilité . . . . .	31
1.5 Variable de réponse continue . . . . .	33
1.6 Variable de réponse binaire . . . . .	34
1.6.1 Identification des effets causaux . . . . .	34
1.6.2 Effets causaux à l'échelle des odds ratios . . . . .	36
1.7 Étude de simulation . . . . .	37
1.7.1 Méthode de simulation des données . . . . .	37
1.7.2 Limites des analyses simples indépendantes en présence d'une cause commune affectant les médiateurs. . . . .	39
1.7.3 Étude empirique des propriétés des estimateurs. . . . .	41
<b>2 Grande dimension</b>	<b>47</b>
2.1 Introduction . . . . .	47
2.2 Classification des médiateurs . . . . .	48
2.2.1 Classification spectrale . . . . .	49
2.2.2 Analyse de médiation par blocs . . . . .	51
2.2.3 Étude de simulation . . . . .	53
2.3 Sélection de médiateur en grande dimension. . . . .	55
Problème . . . . .	56

2.3.1	MAHI en présence d'un seul traitement . . . . .	57
2.3.2	Quelques méthodes existantes . . . . .	61
	Méthode CMF . . . . .	62
	Méthodes HIMA et HDMA . . . . .	63
2.3.3	Comparaison des méthodes . . . . .	64
	Métriques de comparaison . . . . .	64
	Paramétrage des méthodes . . . . .	66
	Résultats . . . . .	66
<b>3</b>	<b>Package R : multimediate</b>	<b>73</b>
3.1	Introduction . . . . .	73
3.2	Algorithmes d'estimation . . . . .	74
3.2.1	Algorithme d'Imai . . . . .	74
3.2.2	Estimation de la corrélation entre médiateurs . . . . .	75
3.2.3	Algorithme pour médiateurs multiples non causalement corrélés . . . . .	78
3.3	Présentation du package . . . . .	80
3.3.1	Analyse de médiation en présence de médiateur corrélés de façon non-causale . . . . .	80
	Présentation de la fonction <code>multimediate</code> . . . . .	80
	Exemple d'utilisation de la fonction <code>multimediate</code> . . . . .	81
3.3.2	Analyse de médiation par bloc . . . . .	83
	Présentation de la fonction <code>blockmediate</code> . . . . .	83
	Exemple d'utilisation de la fonction <code>blockmediate</code> . . . . .	85
3.3.3	Sélection de médiateur en grande dimension . . . . .	88
	Présentation de la fonction <code>mahi</code> . . . . .	88
	Exemple d'utilisation de la fonction <code>mahi</code> . . . . .	88
<b>4</b>	<b>Applications à des données réelles</b>	<b>91</b>
4.1	Effet de l'hormonothérapie substitutive sur le cancer du sein . . . . .	91
4.2	Effets des tempéraments sur l'anorexie mentale . . . . .	100
	<b>Conclusion</b>	<b>109</b>
	<b>Annexes</b>	<b>115</b>
A	Médiateurs liés de façon non-causale . . . . .	115
A.1	Lien entre $\delta^Z$ et $\sum_k \delta^k$ . . . . .	115
A.2	Hypothèses . . . . .	116
A.3	Preuve du Théorème 1.4.1 . . . . .	117
	Effet médié joint, effet direct et effet total . . . . .	117
	Effet médié par un médiateur d'intérêt . . . . .	119
A.4	Preuve du Corollaire 1.5.1 . . . . .	120

	Identifiabilité des coefficients . . . . .	120
	Effet médié par un médiateur d'intérêt . . . . .	121
	Effet direct . . . . .	123
	Effet médié joint . . . . .	123
A.5	Preuve du Corollaire 1.6.1 . . . . .	125
A.6	Modèles de simulation . . . . .	132
A.7	Résultats complémentaires . . . . .	135
B	Grande Dimension . . . . .	142
B.1	Propriétés du graphe laplacien non normalisé . . . . .	142
B.2	MAHI en présence de plusieurs traitements . . . . .	142
B.3	Méthode HIMA . . . . .	144
B.4	Méthode CMF . . . . .	147
C	Package R . . . . .	149
C.1	P-valeur . . . . .	149
C.2	Médiateurs catégoriels ordonnés . . . . .	150
C.3	Médiateurs linéaire et catégoriels ordonnés . . . . .	152
	<b>Bibliographie</b>	<b>155</b>
	<b>Table des figures</b>	<b>165</b>
	<b>Liste des tableaux</b>	<b>169</b>

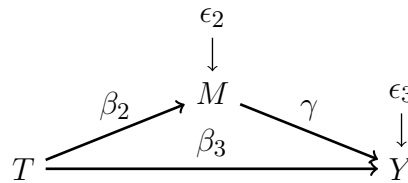


# Introduction

L'analyse causale de médiation est une méthode statistique permettant d'étudier les mécanismes définissant les relations entre trois entités, la première étant une cause, la deuxième une variable de réponse et la troisième un ensemble de variables intermédiaires. Cette approche est devenue populaire dans différents domaines tels que les biostatistiques, l'épidémiologie et les sciences sociales. Elle s'applique à des situations similaires à celles décrites sur le diagramme acyclique présenté dans la Figure 1, où une exposition (ou encore un traitement)  $T$  affecte une variable de réponse  $Y$  d'une part directement et d'autre part à travers une ou plusieurs variables appelées *médiateurs*. L'objectif de cette analyse est de quantifier et décomposer l'effet total de  $T$  sur  $Y$  en un effet *direct* et un effet *indirect* à travers le(s) médiateur(s).

## 0.1 Méthodes basée sur les modèles à équations linéaires structurelles (LSEM)

L'analyse de médiation a été initialement développée dans le cadre des modèles à équations linéaires structurelles (LSEM) (Baron and Kenny (1986), James et al. (1982), MacKinnon (2008)). Cette modélisation est une technique multivariée qui utilise un système d'équations liées entre elles pour capturer des relations complexes et dynamiques au sein d'un ensemble de variables observées et non observées. Une variable dépendante dans une équation de modèle peut devenir une variable indépendante dans d'autres composants du système. C'est précisément ce types de rôles réciproques qui permettent de déduire des relations causales. Ce système d'équation peut simplement être représenté par un diagramme causal. Dans une situation où toutes les variables sont continues et peuvent être décrites par des modèles linéaires, les équations structurelles associées à la Figure 1 sont les suivantes :

FIGURE 1 – Modèle de médiation simple avec un médiateur  $M$ .

$$\begin{aligned} Y &= \alpha_1 + \beta_1 T + \epsilon_1 \\ M &= \alpha_2 + \beta_2 T + \epsilon_2 \\ Y &= \alpha_3 + \beta_3 T + \gamma M + \epsilon_3, \end{aligned}$$

On suppose que les erreurs  $\epsilon_i$  sont indépendantes entre elles et suivent une loi normale  $\mathcal{N}(0, \sigma_i^2)$  pour  $i \in \{1, 2, 3\}$ .

L'effet direct est donné par  $\beta_3$  et l'effet indirect est donné par le produit des coefficients  $\beta_2 \times \gamma$ . L'effet total est donné par la somme de l'effet direct et indirect  $\beta_1 = \beta_3 + \beta_2 \times \gamma$ .

En pratique, pour estimer les effets direct, indirect et total, une première étape consiste à réaliser la régression du traitement sur la variable de réponse puis à vérifier la significativité de l'effet total  $\beta_1$ . Si le paramètre est significatif, une deuxième étape consiste à réaliser les deux autres régressions du modèle LSEM pour quantifier les effets direct et indirect. On contrôle leur significativité ; celle du produit des coefficients est obtenue par la delta méthode, décrite dans l'annexe B.4 (Sobel (1986)).

Suite aux travaux fondateurs de Robins and Greenland (1992) et Pearl (2001), un cadre formel basé sur des variables contrefactuelles s'est imposé comme l'approche standard de l'analyse de médiation, avec une littérature méthodologique croissante (voir par exemple Petersen et al. (2006), VanderWeele and Vansteelandt (2009, 2010), Lange et al. (2012) et le livre de VanderWeele (2015)). Dans un souci d'exhaustivité, nous passerons brièvement en revue les notions de base concernant les contrefactuels en utilisant les mêmes notations que Wasserman (2010).

## 0.2 Notion des contrefactuels

Est contrefactuel ce qui aurait pu être si . . . , mais qui n'est pas. La notion de contrefactuel est liée à la notion de cause à effet. Une variable contrefactuelle est "l'état" dans lequel serait une variable si sa cause était dans un "état" différent. On peut aussi utiliser les termes "variables potentielles" pour parler de "variables contrefactuelles". Pour illustrer

$$T \xrightarrow{\tau} Y$$

FIGURE 2 – Relation de cause à effet entre le traitement  $T$  et la variable de réponse  $Y$ , quantifié par l'effet causal moyen  $\tau$ .

ces propos, considérons un traitement binaire  $T$  et une variable de réponse binaire  $Y$ , comme c'est souvent le cas en épidémiologie :

$$T_i = \begin{cases} 1 & \text{si le patient } i \text{ prend le traitement } T \\ 0 & \text{sinon} \end{cases}$$

$$Y_i = \begin{cases} 1 & \text{si le patient } i \text{ est affecté par une maladie} \\ 0 & \text{sinon} \end{cases}$$

Dans ce cas très simple nous avons deux variables de réponse potentielles  $Y_i(1)$  et  $Y_i(0)$ .  $Y_i(1)$  est la variable de réponse potentielle qu'on aurait obtenue si le patient  $i$  prenait le traitement alors que  $Y_i(0)$  est la variable de réponse potentielle qu'on aurait obtenue si ce même patient  $i$  ne prenait pas le traitement. La relation entre la variable de réponse  $Y_i$  (réelle et observée) et les variables potentielles  $Y_i(1), Y_i(0)$  est la suivante :

$$Y_i = \begin{cases} Y_i(0) & \text{si } T_i = 0 \\ Y_i(1) & \text{si } T_i = 1 \end{cases} = Y_i(T_i) \quad (0.2.1)$$

Cette relation est appelée *relation de consistance*. Une seule des deux variables de réponse potentielles peut être observée comme illustré dans la Table 1, l'autre est appelée contrefactuelle.

$i$	$T_i$	$Y_i$	$Y_i(0)$	$Y_i(1)$
1	0	0	0	NA
2	0	0	0	NA
3	0	1	1	NA
4	1	1	NA	1
5	1	1	NA	1
6	1	0	NA	0

TABLE 1 – Exemple de données observables. On écrit NA lorsque la variable n'est pas observable.

Si les deux étaient observables on en apprendrait beaucoup. D'une part sur les types de patients comme présenté dans la Table 2 et d'autre part, si pour tout  $i$  on pouvait observer les deux variables contrefactuelles, on pourrait savoir pour chaque individu si le



Type du patient $i$	$Y_i(1)$	$Y_i(0)$
Survivant	0	0
Réceptif	0	1
Anti-Réceptif	1	0
Condamné	1	1

TABLE 2 – Type de patient en fonction des valeurs des deux variables de réponse potentielles.

traitement  $T$  a un effet causal sur la variable de réponse  $Y$  ou pas. L'*effet causal moyen*  $\tau$  du traitement sur la variable de réponse, est défini par :

$$\tau = \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \quad (0.2.2)$$

$\tau$  peut s'interpréter comme la moyenne de la variable de réponse si tous les patients étaient traités, à laquelle on soustrait la moyenne de la variable de réponse si tous les patients n'étaient pas traités. Ne pouvant observer les variables potentielles pour un même individu, on ne peut pas identifier  $\tau$ , c'est-à-dire qu'en pratique on ne peut pas accéder à la valeur de  $\tau$  en appliquant directement la formule (0.2.2). Cependant on peut calculer indirectement  $\tau$  sous l'hypothèse d'indépendance entre le traitement et les variables de réponse potentielles :

$$T \perp\!\!\!\perp \{Y(1), Y(0)\} \quad (0.2.3)$$

Cette hypothèse est respectée lorsque le traitement est attribué aléatoirement aux patients, c'est-à-dire sans chercher à privilégier certaines catégories de patients dans l'attribution du traitement, ce qui est le cas lors d'essais randomisés par exemple (c'est d'ailleurs ce qu'on appelle l'hypothèse de randomisation). Sous l'hypothèse de randomisation, la relation de consistance et sous l'hypothèse de positivité ( $0 < P(T = 1) < 1$ ), on observe :

$$\begin{aligned} \tau &= \mathbb{E}[Y(1)] - \mathbb{E}[Y(0)] \\ &= \mathbb{E}[Y(1)|T = 1] - \mathbb{E}[Y(0)|T = 0], \quad \text{d'après (0.2.3)} \\ &= \mathbb{E}[Y|T = 1] - \mathbb{E}[Y|T = 0], \quad \text{d'après (0.2.1)}. \end{aligned}$$

L'effet causal moyen est alors identifiable, car il est exprimé à partir de données observables (Table 1).

### 0.3 Analyse de médiation simple

Pour définir les effets causaux de médiation, nous devons étendre les notions précédentes. Considérons le diagramme causal de la Figure 3. L'objectif est d'expliquer l'effet

causal de  $T$  sur  $Y$  en le décomposant en deux parties, l'effet direct  $\zeta$  du traitement sur la variable de réponse et l'effet indirect du traitement à travers une variable intermédiaire  $M$ .

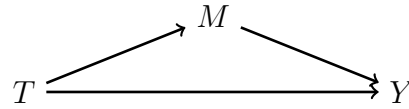


FIGURE 3 – Diagramme causal d'analyse de médiation simple avec un médiateur  $M$ .

$M$  étant une variable post-traitement (c'est-à-dire observée après l'administration du traitement), il existe deux médiateurs potentiels dans le cas d'un traitement binaire,  $M_i(0)$  et  $M_i(1)$  :

$$M_i = \begin{cases} M_i(0) & \text{si } T_i = 0 \\ M_i(1) & \text{si } T_i = 1 \end{cases} = M_i(T_i).$$

$Y$  étant une variable post-traitement et post-médiateur (c'est-à-dire observée après l'administration du médiateur), il existe alors quatre variables de réponse potentielles. Deux d'entre elles,  $Y(0, M(0))$  et  $Y(1, M(1))$  peuvent être observées alors que les deux autres,  $Y(0, M(1))$  et  $Y(1, M(0))$ , ne pourront pas l'être. L'interprétation de celles que l'on peut observer est analogue aux cas précédents :

$$Y_i = \begin{cases} Y_i(0, M_i(0)) & \text{si } T_i = 0 \\ Y_i(1, M_i(1)) & \text{si } T_i = 1 \end{cases} = Y_i(T_i, M_i) = Y_i(T_i).$$

L'interprétation des deux autres est plus complexe : pour  $t$  différent de  $t'$ ,  $Y_i(t, M_i(t'))$  est la valeur potentielle de la variable de réponse si le traitement avait été fixé à  $t$  et si le médiateur avait été fixé à sa valeur potentielle avec un traitement  $t'$ .

Comment, à l'aide de ces contrefactuels, définir les effets direct, indirect et total ? Prenons l'exemple de la discrimination à l'embauche, illustré dans la Figure 4.

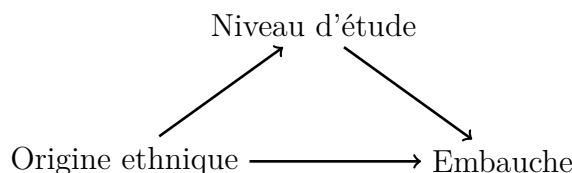


FIGURE 4 – Diagramme causal de discrimination à l'embauche.

Imaginons qu'une entreprise soit accusée de discrimination à l'embauche. Une défense possible pour cette entreprise est d'avancer qu'elle engage ses employés par rapport à leur

compétence ou niveau d'étude. Si on dispose des données listant les origines des personnes postulant dans cette entreprise, leur niveau d'étude et la décision de l'entreprise quant à l'embauche ou non, l'analyse de médiation permettrait de statuer sur la discrimination. En effet si l'effet direct de l'origine ethnique est de valeur non-nulle, on conclura qu'il y a discrimination. Inversement si tout l'effet de l'origine ethnique passe à travers le niveau d'étude on conclura qu'il n'y a pas de discrimination à l'embauche.

L'idée qui permet de comprendre comment sont définis les effets causaux de médiation ressort dans un rapport judiciaire d'une affaire opposant un salarié à son entreprise pour discrimination (Figure 4.9).

“The central question in any employment-discrimination case is whether the employer would have taken the same action had the employee been of a different race (age, sex, religion, national origin etc.) and everything else had been the same.” (Carson versus Bethlehem Steel Corp., 70 FEP Cases 921, 7th Cir. (1996), Quoted in Gastwirth 1997.)

FIGURE 5 – Extrait du rapport de l'affaire opposant Carson à Bethlehem Steel Corp.

L'origine ethnique (âge, sexe, religion) a-t-elle une influence direct sur l'embauche, toute chose étant égale par ailleurs ? L'origine ethnique a-t-elle une influence indirecte à travers leur niveau d'étude toute chose étant égale par ailleurs ?

C'est avec cette idée de "toute chose étant égale par ailleurs" que sont définis les effets direct  $\zeta$  et indirect  $\delta$  (Robins and Greenland (1992), Pearl (2001)). Ces effets sont aussi appelé effets naturels. Pour  $t \in \{0, 1\}$ , les effets sont définis de la manière suivante :

$$\delta(t) = \mathbb{E}[Y(t, M(1))] - \mathbb{E}[Y(t, M(0))] \quad (0.3.1)$$

$$\zeta(t) = \mathbb{E}[Y(1, M(t))] - \mathbb{E}[Y(0, M(t))] \quad (0.3.2)$$

On fait une différence de deux termes contrefactuels où, pour l'effet indirect  $\delta$ , le traitement affectant directement la variable de réponse est fixé à la même valeur,  $t$ , dans ces deux termes et où le traitement affectant le médiateur est fixé à 1 dans le premier terme et à 0 dans le deuxième, donnant ainsi accès à l'effet du traitement sur la variable de réponse à travers le médiateur.

De façon analogue pour l'effet direct  $\zeta$ , le traitement affectant le médiateur est fixé à la même valeur, et celui affectant directement la variable de réponse est fixé à 1 dans le premier terme et à 0 dans le deuxième, donnant ainsi accès à l'effet direct du traitement sur la variable de réponse.

L'effet direct définis dans l'équation 0.3.2 est appelé effet direct naturel, parce que au lieu de fixer une valeur du médiateur pour tous les individus, on fixe une certaine valeur du traitement et, pour chaque individu, on prend la valeur du médiateur qu'on obtiendrait *naturellement* sous l'effet de celle-ci. Mais on peut aussi parler d'effet direct contrôlé (Pearl (2001)), défini par :

$$\zeta^c(m) = \mathbb{E}[Y(1, m)] - \mathbb{E}[Y(0, m)] \quad (0.3.3)$$

Pour l'effet direct contrôlé, on fixe le médiateur affectant la variable de réponse à la même valeur  $m$ , le traitement affectant le premier terme est fixé à 1 et celui affectant le deuxième est fixé à 0, donnant ainsi accès à l'effet du traitement sur la variable de réponse pour une valeur de médiateur fixé à  $m$ .

Dans le cadre de cette thèse nous nous intéressons uniquement aux effets naturels. Ces effets ne sont pas directement identifiables, puisque définis avec des variables contrefactuelles inobservables.

Cependant, une hypothèse permet de rendre ces effets identifiables. C'est ce qu'ont montrés Imai et al. (2010b) avec deux hypothèses cruciales, appelées Hypothèse d'Ignorabilité Séquentielle (SIA), et données ci-dessous :

$$\{Y(t', m), M(t)\} \perp\!\!\!\perp T | X = x \quad \forall t, t', m \quad (0.3.4)$$

$$Y(t', m) \perp\!\!\!\perp M(t) | T = t, X = x \quad \forall t, t', m. \quad (0.3.5)$$

où  $P(T = t | X = x) > 0$  et  $P(M(t) = m | T = t, X = x) > 0$  pour toutes les valeurs possibles de  $t, m$  et  $x$ .

Ces conditions sont interprétées comme 1) l'absence de confusion sur les relations causales  $T - Y$ ,  $T - M$  et  $M - Y$  après ajustement sur les covariables de pré-traitement (c'est-à-dire les covariables qui ne sont pas affectées par  $T$ ), 2) l'absence de confusion due à des covariables de post-traitement (c'est-à-dire des covariables qui sont affectées par  $T$ ) entre  $M$  et  $Y$  quelles qu'elles soient, mesurées ou non.

En particulier, Imai et al. (2010b,a) ont démontré que sous les hypothèses d'Ignorabilité Séquentielle, l'effet indirect moyen était non-paramétriquement identifié, voir Théorème 0.3.1. De plus, ils ont proposé une analyse de sensibilité pour évaluer la robustesse des estimations aux violations de l'Ignorabilité Séquentielle. Notons que les covariables prises en compte dans les modèles seront notées  $X$ .

**Théorème 0.3.1** (Imai et al. (2010b)). *Sous SIA, les effets direct et indirect moyens sont identifiés non-paramétriquement et sont donnés par :*

$$\begin{aligned} \delta(t) &= \int \int \mathbb{E}[Y | M = m, T = t, X = x] dF_{M|T=1, X=x}(m) \\ &\quad - \int \int \mathbb{E}[Y | M = m, T = t, X = x] dF_{M|T=0, X=x}(m) dF_X(x) \\ \zeta(t) &= \int \int \mathbb{E}[Y | M = m, T = 1, X = x] dF_{M|T=t, X=x}(m) \\ &\quad - \int \int \mathbb{E}[Y | M = m, T = 0, X = x] dF_{M|T=t, X=x}(m) dF_X(x), \end{aligned}$$

pour  $t \in \{0, 1\}$ .

Dans le contexte des modèles linéaires, on déduit du théorème précédents les deux corollaires suivants, le premier pour une variable de réponse continue et le deuxième pour une variable de réponse binaire.

**Corollaire 0.3.2** (Imaï et al. (2010b)). *Sous SIA et en supposant les modèles linéaires à équations structurelles (LSEM) suivants :*

$$\begin{aligned} M &= \alpha_2 + \beta_2 T + \xi_2^\Gamma X + \epsilon_2 \\ Y &= \alpha_3 + \beta_3 T + \gamma M + \xi_3^\Gamma X + \epsilon_3, \end{aligned}$$

où  $\epsilon_i \sim \mathcal{N}(0, \sigma_i^2)$  pour  $i \in \{2, 3\}$ , les effets direct et indirect moyens sont identifiés par  $\delta(0) = \delta(1) = \beta_2 \gamma$  et  $\zeta(0) = \zeta(1) = \beta_3$ .

On remarque qu'on retrouve ainsi les résultats donnés précédemment dans le cadre simple des LSEM.

Lorsque la variable de réponse est binaire, sa distribution conditionnelle peut être modélisée de deux façons. D'une part avec un modèle probit

$$\mathbb{P}(Y = 1|T, M, X) = \Phi_{\mathcal{N}(0, \sigma_3^2)}(\alpha_3 + \beta_3 T + \gamma M + \xi_3^\Gamma X),$$

où  $\Phi_{\mathcal{N}(0, \sigma_3^2)}$  est la fonction de répartition d'une loi normale  $\mathcal{N}(0, \sigma_3^2)$ .

D'autre part, avec un modèle logit

$$\text{logit}(\mathbb{P}(Y = 1|T, M, X)) = \alpha_3 + \beta_3 T + \gamma M + \xi_3^\Gamma X.$$

**Corollaire 0.3.3** (Imaï et al. (2010b)). *Soit  $Y$  une variable binaire, on suppose les modèles suivants*

$$\begin{aligned} M &= \alpha_2 + \beta_2 T + \xi_2^\Gamma X + \epsilon_2 \\ Y &= \mathbb{1}_{\{Y^* > 0\}}, \text{ avec } Y^* = \alpha_3 + \beta_3 T + \gamma M + \xi_3^\Gamma X + \epsilon_3 \end{aligned}$$

où  $\epsilon_2 \sim \mathcal{N}(0, \sigma_2^2)$  et  $\epsilon_3 \sim \mathcal{N}(0, \sigma_3^2)$  (respectivement  $\epsilon_3 \sim \mathcal{L}(0, 1)$ ) dans le cas du modèle probit (respectivement modèle logit).  $\mathcal{L}(0, 1)$  désigne la fonction de répartition d'une loi logistique. Sous SIA, les effets direct et indirect moyens sont identifiés par

$$\begin{aligned} \delta(t) &= \mathbb{E}[F_u(h_{t,1}) - F_u(h_{t,0})] \\ \zeta(t) &= \mathbb{E}[F_u(h_{1,t}) - F_u(h_{0,t})] \end{aligned}$$

où

$$h_{t,t'} = \alpha_3 + \beta_3 t + \gamma(\alpha_2 + \beta_2 \times t' + \xi_2^\Gamma X) + \xi_3^\Gamma X$$

et pour un modèle probit la fonction  $F_u$  est donnée par :

$$F_u(z) = \Phi \left( \frac{z}{\sqrt{\gamma^2 \sigma_2^2 + 1}} \right)$$

alors que pour un modèle logit on a :

$$F_u(z) = \int_{-\infty}^{\infty} \Phi \left( \frac{z-y}{\gamma \sigma_2} \right) \frac{e^y}{(1+e^y)^2} dy.$$

## 0.4 Problématiques

Quand plusieurs médiateurs sont impliqués dans le modèle de médiation, il y a trois cas possibles de relation entre ces médiateurs. La Figure 6 illustre ces différentes situations avec l'exemple de deux médiateurs :

- Fig. 6(a) : Les médiateurs sont indépendants conditionnellement au traitement et aux covariables mesurées (non représentées).
- Fig. 6(b) : Les médiateurs sont dépendants conditionnellement au traitement et aux covariables mesurées sans pour autant que l'un affecte l'autre.
- Fig. 6(c) : Les médiateurs sont liés de façon causale, c'est-à-dire que l'un affecte l'autre.

Dans la figure 6(b), on parle de médiateurs *liés ou corrélés de façon non-causale* contrairement à la situation de la 6(c), où l'on parle de médiateurs corrélés de façon causale. De façon plus générale nous parlerons de *médiation avec médiateurs non-causalement liés* pour désigner les cas représentés dans les Figures 6(a) et 6(b).

Les modèles de médiation des Figures 6(a) et 6(c) ont été traités durant ces dernières années, notamment dans les articles suivants : VanderWeele and Vansteelandt (2014), Lange et al. (2014), Daniel et al. (2015).

Le modèle de médiation de la Figure 6(b) correspond à un Graphe Acyclic Mixte Dirigé (ADMG), notion introduite par Shpitser (2013) et Shpitser and Sherman (2018). De ces travaux on déduit que l'effet direct (médié ni par  $M$  ni par  $W$ ) et l'effet joint (médié par  $M$  et  $W$ ) peuvent être définis à partir d'observations. Cependant l'effet médié uniquement par l'un des médiateurs  $M$  et  $W$  n'a pas été défini dans ce contexte.

Pourtant l'estimation de ces effets indirects individuels, chacun spécifique à un médiateur donné, est d'une importance pratique. Dans le cadre de plusieurs médiateurs, Imai and Yamamoto (2013) proposent simplement d'étendre leur approche mentionnée

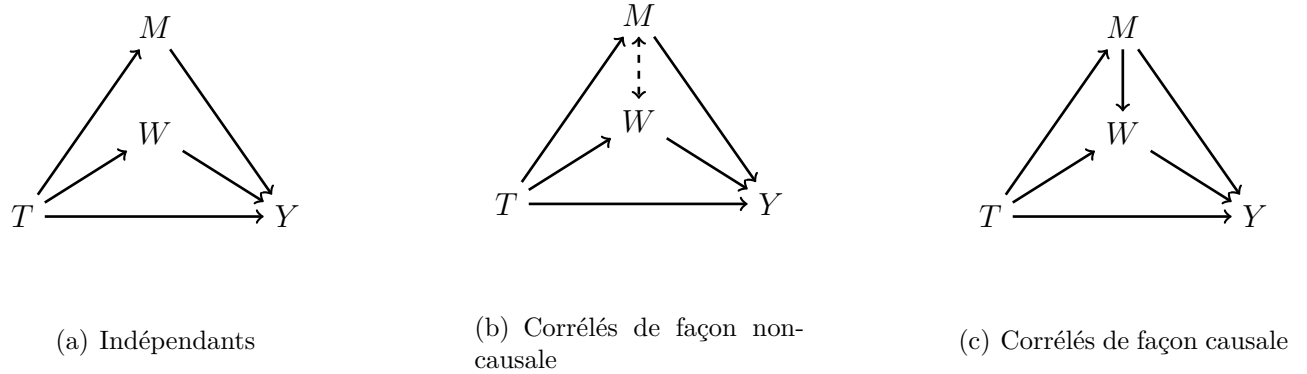
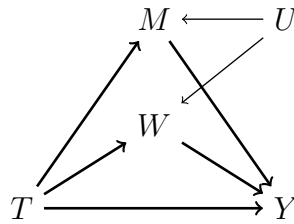


FIGURE 6 – Les trois relations possibles entre deux médiateurs  $M$  et  $W$ .

ci-dessus à plusieurs médiateurs. Lorsque les médiateurs n'ont aucun lien de causalité et que les hypothèses d'Ignorabilité Séquentielle sont vérifiées, ils suggèrent de faire plusieurs analyses de médiation simples indépendantes, un médiateur à la fois. De toute évidence, cette approche conduit à une estimation biaisée de l'effet direct, car elle oblige les effets indirects via tous les autres médiateurs à contribuer à l'effet direct. Plus subtilement, cette approche n'est pas appropriée lorsque les médiateurs sont corrélés de façon non-causale en raison d'une covariable non mesurée  $U$  affectant causalement les deux médiateurs  $M$  et  $W$  comme dans la figure 7. En effet, dans cette situation,  $U$  est un confondant non-observé de la relation entre  $M$  et  $Y$  et l'Ignorabilité séquentielle n'est alors pas vérifiée. Cette remarque importante a été faite par Imai and Yamamoto (2013) et VanderWeele and Vansteelandt (2014), mais aucune solution explicite au problème n'a été proposée autre que la réalisation de l'analyse de sensibilité susmentionnée. Un article récent de Kim et al. (2019) décrit une approche alternative dans laquelle la dépendance entre médiateurs est caractérisée par une copule gaussienne avec des modèles linéaires marginaux ; en utilisant une approche entièrement bayésienne, ils imputent les contrefactuels non observés pour estimer l'effet direct et les effets indirects à travers chaque médiateur. Cependant, cette approche a été spécifiquement développée pour des variables continues. Nous nous sommes alors demandés : *"Comment peut-on, en présence de médiateur liés de façon non-causale, calculer les effets indirects individuels de façon non-biaisée ? Quelles nouvelles hypothèses devons-nous introduire pour cela ?"* Dans la première partie de ce travail de thèse nous répondons à ces questions.

Après avoir proposé une modélisation permettant de répondre à ces premières questions, d'autres sont naturellement apparues concernant le nombre de médiateurs à prendre en compte dans le modèle. En effet, des modèles computationnellement coûteux sont-ils adaptés lorsque le nombre de médiateurs est très grand ? La réponse à cette question a fait l'objet de la seconde partie de ce travail de thèse. Dans le cadre de l'analyse de médiation, les données de grande dimension correspondent à des données où le nombre de médiateurs

FIGURE 7 – Corrélation entre les médiateurs induite par  $U$ .

*candidats* est important et en particulier supérieur au nombre d'observations (Figure 8). Notre méthode d'analyse de médiateurs multiples décrite au Chapitre 1 ne peut pas être appliquée au cas de la médiation en grande dimension car en pratique elle repose sur la régression, ce qui ne donne pas de résultats fiables lorsque le nombre d'observations est beaucoup plus grand que le nombre de variables. On peut alors se demander : *"Doit-on s'intéresser à l'ensemble des médiateurs ou en sélectionner certains ? Quels sont leurs effets ? Peut-on éventuellement classifier ces médiateurs ?"*

Dans les applications, il est raisonnable de faire l'hypothèse que seule une petite proportion des médiateurs candidats peut médier l'effet de l'exposition sur la variable de réponse. D'un point de vue théorique le modèle de la variable de réponse  $Y$  en fonction des médiateurs est identifiable lorsque le nombre de médiateurs est relativement petit, c'est-à-dire inférieur au nombre d'observations. Cette hypothèse nous permet d'étendre l'analyse de médiation multiple à un cadre de grande dimension en faisant de l'identification des médiateurs actifs un problème de sélection de variables.

## 0.5 Organisation de la thèse

Ce travail de thèse s'est déroulé en deux parties. Dans un premier temps, nous avons développé une méthode d'analyse de médiation multiple en présence de médiateurs liées de façon non-causale. Dans un deuxième temps, nous avons proposé une nouvelle méthode de sélection de médiateur dans le cadre de la grande dimension. Le manuscrit s'organise de la manière suivante :

Dans le Chapitre 1 de cette thèse, nous suggérons qu'une solution possible à l'estimation des effets d'intérêts en présence de médiateur corrélés de façon non-causale passe par l'estimation de la loi multivariée des médiateurs conditionnellement au traitement. Cela permet de prendre en compte la corrélation entre médiateurs induite par la variable non observée  $U$ . Nous présentons donc le Théorème 1.4.1 qui permet d'estimer les effets causaux de médiations de façon non-paramétrique, excepté pour les effets indirects individuels. Nous présentons aussi les corollaires 1.5.1 et 1.6.1 qui, pour des modèles paramétriques, permettent l'estimation et l'identification de tous les effets causaux de



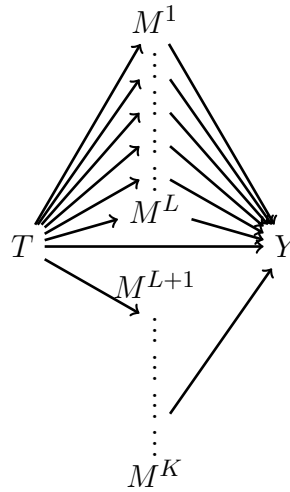


FIGURE 8 – Modèle de médiation en présence d’un grand nombre de médiateurs potentiels. Les variables  $(M^1, \dots, M^L)$  ont un rôle de médiation et ce sont ces variables que l’on cherche à sélectionner. Les variables  $(M^{L+1}, \dots, M^K)$  ne sont pas des médiateurs et on cherche à les exclure.

médiation.

Le Chapitre 2 aborde deux domaines, la question de la classification de médiateurs et celle de la sélection de médiateurs. Dans la première partie du Chapitre 2 nous proposons une méthode de classification basée sur les corrélations entre les médiateurs, corrélations conditionnelles au traitement et aux covariables. En se basant sur une approche de classification spectrale expliquée par von Luxburg (2007), cette méthode permet d’isoler des blocs de médiateurs indépendants entre eux. Ces blocs sont constitués de médiateurs corrélés entre eux de façon causale et/ou non-causale. Ainsi, en introduisant une analyse de médiation par bloc, on est ramené à la problématique du Chapitre 1.

La deuxième partie du Chapitre 2 décrit une nouvelle méthode de sélection de médiateurs reposant sur 2 étapes. La première se base sur une approche Group-LASSO de façon à réduire le nombre de médiateurs à un nombre inférieur au nombre de données. Nous utilisons ensuite notre méthode présentée au Chapitre 1 pour calculer les effets médiés et leur  $p$ -valeur permettant ainsi d’effectuer un test multiple sur les effets médiés. Nous nous comparons à trois méthodes de sélection de médiateurs existantes HIMA, HDMA et CMF. Notre méthode permet, de plus, la prise en compte simultanée de plusieurs traitements, sélectionnant des médiateurs communs pour plusieurs traitement identifiés sur une variable de réponse. L’ensemble de notre méthode est implémenté et disponible dans notre package présenté au Chapitre 3.

Dans le Chapitre 3 nous présentons l’algorithme d’estimation des effets causaux d’in-

térêts que nous utilisons dans notre package R `multimediate` disponible sur GitHub (<https://github.com/AllanJe/multimediate>). Nous présentons aussi notre méthode d'estimation de la corrélation entre les médiateurs conditionnellement au traitement et aux covariables pour différents type de médiateurs.

Finalement dans le Chapitre 4 nous avons appliqué notre méthode d'analyse de médiation multiple sur des données réelles. Premièrement en collaboration avec Laura Baglietto, Marco Fornili du Département de Médecine Clinique de l'Université de Pise et Gianluca Severi du Centre de Recherche en Epidémiologie et Santé des Populations de l'Inserm, nous avons travaillé sur l'effet de l'hormonothérapie substitutive sur le cancer du sein. Ce travail a permis la rédaction d'un article en soumission (Fornili et al. (2020)), dont nous présentons une partie des résultats. Deuxièmement en collaboration avec Enrica Marzola du Département de Neurosciences de l'Université de Turin, nous avons aussi travaillé sur l'effet des tempéraments affectifs sur l'anorexie mentale. Ce travail fait l'objet d'un article en cours de rédaction. Nous présentons ici les résultats qui figureront dans l'article.

Le Chapitre 1 correspond à l'article publié, Jérôlon et al. (2020), rédigé en collaboration avec Etienne Birmélé, Laura Baglietto, Vittorio Perduca et Flora Alarcon. L'article a été adapté de façon à être cohérent à un Chapitre de cette thèse. Notez que la sous-section 1.6.2 a été ajoutée car il s'agit d'un travail complémentaire ultérieure à la soumission de l'article Jérôlon et al. (2020). Le Chapitre 2 présente le travail d'un article en cours de rédaction. Le Chapitre 3 présente notre package R disponible sur Github implémentant les méthodes et approches de l'analyse de médiation présenté dans les Chapitres 1 et 2. Chaque Chapitre est introduit en reprenant et détaillant certains points de la présente introduction.

Chapitres	Contributions	Valorisation
Chapitre 1	<ul style="list-style-type: none"> <li>- Extension des hypothèses d'Ignorabilité Séquentielle au cas multiple</li> <li>- Extension des résultats d'identification non-paramétrique au cas multiple</li> <li>- Étude de simulation comparative</li> </ul>	Article publié Jérolon et al. (2020) Présentations orales Posters
Chapitre 2	<ul style="list-style-type: none"> <li>- Classification de médiateurs par leur corrélation</li> <li>- Analyse de médiation par bloc</li> </ul>	
	<ul style="list-style-type: none"> <li>- Sélection de médiateurs en grande dimension</li> </ul>	Article en cours de rédaction Présentation orale
Chapitre 3	<ul style="list-style-type: none"> <li>- Développement d'un algorithme pour l'estimation et l'inférence</li> </ul>	
	<ul style="list-style-type: none"> <li>- Présentation du Package R <code>multimediate</code></li> </ul>	Accessible sur Github
Chapitre 4	<ul style="list-style-type: none"> <li>- Effet de l'hormonothérapie substitutive sur le cancer du sein</li> </ul>	Article soumis Fornili et al. (2020)
	<ul style="list-style-type: none"> <li>- Effets des tempéraments sur l'anorexie mentale</li> </ul>	Article en cours de rédaction

TABLE 3 – Récapitulatif des contributions de chaque chapitre et leurs valorisations.

# Chapitre 1

## Médiateurs liés de façon non-causale

### Sommaire

---

<b>1.1</b>	<b>Introduction</b>	<b>27</b>
<b>1.2</b>	<b>Définition des effets</b>	<b>29</b>
<b>1.3</b>	<b>Hypothèses</b>	<b>30</b>
<b>1.4</b>	<b>Identifiabilité</b>	<b>31</b>
<b>1.5</b>	<b>Variable de réponse continue</b>	<b>33</b>
<b>1.6</b>	<b>Variable de réponse binaire</b>	<b>34</b>
1.6.1	Identification des effets causaux	34
1.6.2	Effets causaux à l'échelle des odds ratios	36
<b>1.7</b>	<b>Étude de simulation</b>	<b>37</b>
1.7.1	Méthode de simulation des données	37
1.7.2	Limites des analyses simples indépendantes en présence d'une cause commune affectant les médiateurs.	39
1.7.3	Étude empirique des propriétés des estimateurs.	41

---

### 1.1 Introduction

Dans ce chapitre nous nous concentrons sur l'analyse de médiation en présence de médiateurs multiples liés de façon non causale (c'est-à-dire soit indépendants, Fig. 6(a), soit corrélés de façon non-causale, Fig. 6(b)). Sur ces figures, les flèches en pointillés bidirectionnels indiquent une corrélation non causale, due par exemple à une cause commune latente, comme dans la Figure 7. Shpitser (2013) et Shpitser and Sherman (2018) définissent les districts comme les composants connectés du graphique limités aux arêtes et bidirectionnelles et décrivent une condition nécessaire et suffisante pour identifier les

effets, exprimée en termes de données d'observation. Dans le cas d'un cadre de médiation multiple, cette condition dit que l'effet médié par un ensemble  $\mathcal{S}$  de médiateurs peut être écrit en fonction des observations si et seulement si  $\mathcal{S}$  peut être écrit comme l'union de certains districts. Dans le cas de la Figure 6(b), cela signifie que l'effet direct (médié ni par  $M$  ni par  $W$ ) et l'effet joint (médié par  $M$  et  $W$ ) peuvent être écrits en termes d'observations, mais que l'effet médié uniquement par  $M$  ne le peut pas.

L'estimation de ces effets indirects individuels, chacun spécifique à un médiateur donné, est cependant d'une importance pratique. Pour ce faire, Imai and Yamamoto (2013) étend son approche mentionnée précédemment à plusieurs médiateurs. Lorsque les médiateurs n'ont aucun lien de causalité et que les hypothèses d'Ignorabilité Séquentielle sont vérifiées, ils suggèrent de faire plusieurs analyses de médiateur simple en parallèle, un médiateur à la fois. De toute évidence, cette approche conduit à une estimation biaisée de l'effet direct, car elle oblige les effets indirects via tous les autres médiateurs à contribuer à l'effet direct. Plus subtilement, cette approche n'est pas appropriée lorsque les médiateurs sont corrélés de façon non-causale en raison d'une covariable non mesurée  $U$  affectant causalement les deux médiateurs  $M$  et  $W$  comme dans la figure 7. En effet, dans cette situation,  $U$  est un confondant non-observé de la relation entre  $M$  et  $Y$  et l'Ignorabilité séquentielle n'est alors pas vérifiée. Cette remarque importante a été faite par Imai and Yamamoto (2013) et VanderWeele and Vansteelandt (2014), mais aucune solution explicite au problème n'a été proposée autre que la réalisation de l'analyse de sensibilité évoquée dans l'Introduction.

Dans l'Introduction de cette thèse nous avons commencé par revoir les résultats et définitions dans le cas de médiation simple obtenus par Imai et al. (2010b). Dans le présent Chapitre, nous étendons les définitions et les résultats théoriques au cas de médiateurs liés de façon non-causale. Pour ce faire, nous introduisons les hypothèses d'identification SIMMA et nous les comparons à celles de l'Ignorabilité Séquentiel (SIA) dans le cas multiple présenté dans Imai and Yamamoto (2013). Nous démontrons que sous les hypothèses SIMMA les effets direct, indirect et joint peuvent être exprimés par des formules n'impliquant que des variables observables, alors que l'effet indirect individuel est obtenu par une formule où interviennent des variables observables et contrefactuelles. Ces formules amènent à des estimations non-biaisées des effets direct and indirect joint, conformément à Shpitser and Sherman (2018).

Pour l'estimation non biaisée des effets indirects individuels, nous suggérons qu'une solution possible au problème d'identification passe par l'estimation de la loi multivariée des médiateurs conditionnellement au traitement. Cela permet de prendre en compte la corrélation entre médiateurs induite par la variable non observée  $U$ . Un article récent de Kim et al. (2019) décrit une approche alternative dans laquelle la dépendance entre médiateurs est caractérisée par une copule gaussienne avec des modèles linéaires marginaux; En utilisant une approche entièrement bayésienne, on impute des contrefactuels non observés pour estimer l'effet direct et les effets indirects à travers chaque médiateur. Cependant, cette approche a été spécifiquement développée pour les résultats continus,

alors que notre méthode est valable pour des combinaisons de modèles présentées dans le Chapitre 3.

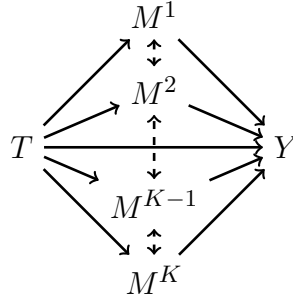


FIGURE 1.1 – Diagramme causal de médiation multiple avec des médiateurs éventuellement corrélés. Le vecteur des facteurs de confusion du prétraitement  $X$  n'apparaît pas dans la figure. Les lignes en pointillés représentent les possibles corrélations non causales et les traits pleins les relations causales. Une corrélation non causale est possible entre chaque paire de médiateurs mais cela n'est pas montré pour une meilleure lisibilité de la figure. Ainsi, chaque paire de médiateur peut être liés causalement.

Dans ce Chapitre et les suivants, nous adoptons le point de vue et le formalisme de Imai et al. (2010a) et Imai et al. (2010b), décrits en Introduction.

## 1.2 Définition des effets

Notons  $Z$  le vecteur de l'ensemble des  $K \geq 2$  médiateurs et  $M^k$  le médiateur d'intérêt. On note  $W^k$  le complément de  $M^k$  dans  $Z$ , c'est-à-dire l'ensemble des médiateurs qui ne sont pas d'intérêt. On note  $X$  le vecteur des covariables de pré-traitement.

L'effet moyen médié par  $M^k$  a été défini par Imai and Yamamoto (2013) et donné par :

$$\delta^k(t) = \mathbb{E} [Y(t, M^k(1), W^k(t))] - \mathbb{E} [Y(t, M^k(0), W^k(t))].$$

Comme mesure de l'effet moyen médié par l'ensemble des médiateurs, on définit :

$$\delta^Z(t) = \mathbb{E} [Y(t, Z(1))] - \mathbb{E} [Y(t, Z(0))].$$

Chacun des  $2^K$  effets directs est défini par :

$$\begin{aligned} \zeta(t_1, \dots, t_K) &= \mathbb{E} [Y(1, M^1(t_1), \dots, M^K(t_K))] \\ &\quad - \mathbb{E} [Y(0, M^1(t_1), \dots, M^K(t_K))] \end{aligned}$$

où  $t_k \in \{0, 1\}$  pour tout  $k \in \{1, \dots, K\}$ .

Nous ne considérerons que  $\zeta(0, \dots, 0)$  et  $\zeta(1, \dots, 1)$ , notés  $\zeta(t)$ ,  $t \in \{0, 1\}$ .  
L'effet total  $\tau$  est donné par :

$$\tau = \mathbb{E}[Y(1, Z(1))] - \mathbb{E}[Y(0, Z(0))].$$

Notons que  $\tau$  est la somme de l'effet indirect joint et de l'effet direct :

$$\tau = \delta^Z(t) + \zeta(1 - t).$$

**Remarque.** Notons que l'effet indirect joint peut se décomposer de la façon suivante :

$$\delta^Z(t) = \frac{\sum_{k=1}^K (\delta^k(t) + \eta^k(t))}{K}$$

où

$$\eta^k(t) = \mathbb{E}[Y(t, M^k(1-t), W^k(1))] - \mathbb{E}[Y(t, M^k(1-t), W^k(0))].$$

Voir le commentaire de l'annexe A.1 pour une preuve de ce résultat.

### 1.3 Hypothèses

À travers ce chapitre, nous adoptons les hypothèses SUTVA (Stable Unit Treatment Value Assumption, Rubin (1980)) qui sont nécessaires pour tout travail sur la causalité avec des contrefactuels. Ces hypothèses impliquent :

1) qu'il n'y a pas "d'interférence" dans le sens où pour l'individu  $i$  le médiateur potentiel et la variable de réponse ne dépendent pas du traitement des autres individus ( c'est-à-dire que  $M_i^k(T) = M_i^k(T_i)$  et  $Y_i(T, M^k, W^k) = Y_i(T_i, M_i^k, W_i^k)$ ) et

2) il n'y a pas "plusieurs versions" du traitement (c'est-à-dire.  $T_i = T'_i$  implique que  $M_i^k(T_i) = M_i^k(T'_i)$  et  $Y_i(T_i, M_i^k(T_i), W_i^k(T_i)) = Y_i(T'_i, M_i^k(T'_i), W_i^k(T'_i))$ ).

Nous complétons les hypothèses SUTVA standards en supposant qu'il n'y a pas "plusieurs versions" des médiateurs, c'est-à-dire que si  $M_i^k = M_i^{k'}$ , alors  $Y_i(T_i, M_i^k, W_i^k) = Y_i(T_i, M_i^{k'}, W_i^k)$  (Forastiere et al. (2016)).

Nos résultats sont basés sur les hypothèses suivantes, que nous avons appelés Hypothèses d'Ignorabilité Séquentielle pour des Médiateurs Multiples (SIMMA) :

$$\{Y(t, m, w), M(t'), W(t'')\} \perp\!\!\!\perp T | X = x, \quad (\text{A.1})$$

$$Y(t', m, w) \perp\!\!\!\perp (M(t), W(t)) | T = t, X = x \quad (\text{A.4})$$

$$Y(t, m, w) \perp\!\!\!\perp (M(t'), W(t)) | T = t, X = x \quad (\text{A.5})$$

où  $P(T = t | X = x) > 0$ ,  $P(M(t) = m, W(t') = w | X = x) > 0$  et  $P(Z = z | T = t, X = x) > 0$  pour toutes les valeurs possibles de  $t, t', m, w$  et  $x$ .

Une explication détaillée de SIMMA est donnée dans l'annexe A.2.

On rappelle ici que  $X$  est le vecteur de toutes les covariables de prétraitement observées, c'est-à-dire des variables qui ne sont pas affectées par le traitement. La première hypothèse implique qu'il ne doit pas y avoir de facteurs de confusion de prétraitement non observés entre le traitement et la variable de réponse et entre le traitement et les médiateurs individuels une fois conditionnés sur toutes les covariables observées. La deuxième et la troisième hypothèses excluent l'existence de deux types distincts de facteurs de confusion entre les médiateurs pris *conjointement* et la variable de réponse : la confusion par une variable de prétraitement non observée et la confusion par une variable de post-traitement observée ou non observée.

Fondamentalement, ces hypothèses remplacent la deuxième et la troisième hypothèse que Imai and Yamamoto (2013) font dans la situation de plusieurs médiateurs causalement indépendants, où une exigence similaire s'applique à chaque médiateur contrefactuel séparément et est interprétée comme la randomisation de chaque médiateur par rapport à la variable de réponse conditionnellement à l'action du traitement (voir annexe A.2). Cependant, il est important de souligner que l'hypothèse (A.4) n'est pas plus restrictive que l'hypothèse d'Imai en ce sens qu'elle ne les implique pas, comme nous le montrons dans l'annexe A.2. Cette hypothèse est la même que l'hypothèse 2) pour plusieurs médiateurs dans (VanderWeele and Vansteelandt (2014)). Notre troisième hypothèse (A.5) n'est pas incluse dans Imai et al. (2010a) ni dans VanderWeele and Vansteelandt (2014) et est nécessaire pour estimer l'effet indirect individuel de chaque médiateur.

La raison du remplacement de l'hypothèse d'Imai par (A.4) et (A.5) est que nous nous intéressons à la situation où  $M$  et  $W$  sont corrélés de manière non causale, généralement en raison d'une variable de prétraitement  $U$  affectant les deux médiateurs comme dans la Figure 1.2(a). Notez que si  $U$  n'est pas observé (c'est-à-dire qu'il n'est pas pris en compte par  $X$ ) les conditions (A.4) et (A.5) ne sont pas violées car la distribution conjointe des médiateurs incorpore l'influence de  $U$  sur les médiateurs individuels. Au contraire, un tel  $U$  violerait l'hypothèse correspondante de Imai and Yamamoto (2013) car il constitue un facteur de confusion non observé des relations entre  $W$  et  $Y$  et  $M$  et  $Y$ .

## 1.4 Identifiabilité

Dans ce qui suit, le médiateur d'intérêt  $M$  peut être l'un des  $K$  médiateurs, de sorte que les résultats ci-dessous peuvent être appliqués à chaque médiateur. En particulier, cela permettra d'exprimer l'effet indirect médié par chaque médiateur pris individuellement.

Notre premier résultat, donné par le Théorème 1.4.1 étend le Théorème 0.3.1 à plusieurs médiateurs, non seulement lorsque les médiateurs sont causalement indépendants comme le fait Imai and Yamamoto (2013), mais aussi lorsqu'ils sont corrélés de façon non causale.



**Théorème 1.4.1.** *On considère les  $K$  médiateurs qui peuvent être indépendants ou corrélés de manière non causale.*

*Sous SIMMA, l'effet indirect moyen du médiateur d'intérêt est défini par :*

$$\delta(t) = \int \left( \int_{\mathbb{R}^K} \mathbb{E}[Y|M = m, W = w, T = t, X = x] \left( \{dF_{(M(1), W(t))|X=x}(m, w) - dF_{(M(0), W(t))|X=x}(m, w)\} \right) dF_X(x). \right. \quad (1.4.1)$$

*De plus, l'effet indirect conjoint, l'effet direct et l'effet total sont identifiés de manière non paramétrique respectivement par :*

$$\delta^Z(t) = \int \left( \int_{\mathbb{R}^K} \mathbb{E}[Y|Z = z, T = t, X = x] dF_{Z|T=1, X=x}(z) - \int_{\mathbb{R}^K} \mathbb{E}[Y|Z = z, T = t, X = x] dF_{Z|T=0, X=x}(z) \right) dF_X(x),$$

$$\zeta(t) = \int \left( \int_{\mathbb{R}^K} \mathbb{E}(Y|Z = z, T = 1, X = x)(z) - \mathbb{E}(Y|Z = z, T = 0, X = x) dF_{Z|T=t, X=x}(z) \right) dF_X(x),$$

$$\tau = \int \left( \int_{\mathbb{R}^K} \mathbb{E}(Y|Z = z, T = 1, X = x) dF_{Z|T=1, X=x}(z) - \int_{\mathbb{R}^K} \mathbb{E}(Y|Z = z, T = 0, X = x) dF_{Z|T=0, X=x}(z) \right) dF_X(x).$$

En médiation multiple, le théorème 1.4.1 a le même rôle que le théorème 0.3.1 en médiation simple, car il montre que sous des hypothèses appropriées, les effets indirect joint et direct sont identifiés de manière non paramétrique. En particulier, à partir des équations ci-dessus, on peut dériver des estimateurs pour l'effet indirect joint et pour l'effet direct, comme déjà montré par Shpitser (2013). En général, cependant, l'équation (1.4.1) ne permet pas d'identifier un estimateur de l'effet indirect individuel du médiateur d'intérêt, car la distribution conditionnelle de  $(M(t'), W(t))$  n'est pas observable. Notez que dans le cas particulier où  $M$  est indépendant de  $W$ , l'équation (1.4.1) devient

$$\delta(t) = \int \int \mathbb{E}[Y|M = m, T = t, X = x] dF_{M|T=1, X=x}(m) - \int \int \mathbb{E}[Y|M = m, T = t, X = x] dF_{M|T=0, X=x}(m) dF_X(x),$$

qui est l'équation pour  $\delta(t)$  donnée par le théorème 0.3.1, permettant ainsi d'identifier l'effet indirect moyen de manière non paramétrique. Ce résultat a été donné par Imaï and Yamamoto (2013). Une preuve du théorème 1.4.1 est donnée dans l'annexe A.3.

Les deux corollaires suivants montrent des formules d'identification des effets indirects et directs dans le cadre du LSEM ou lorsque les variables médiatrices sont gaussiennes et  $Y$  est binaire. Fondamentalement, dans les corollaires suivants, nous supposons que les corrélations entre les médiateurs potentiels sont les mêmes quel que soit le traitement régissant les médiateurs :

$$\text{cor}(M^i(t), M^j(t')|T, X) = \rho_{ij}, \forall t, t' \in \{0, 1\}, \forall i, j \in [1, K]. \quad (1.4.2)$$

Cette hypothèse est en effet suffisante pour identifier les effets indirects individuels à travers  $M$  de l'équation (1.4.1) dans les modèles où la distribution jointe des médiateurs est complètement décrite par la matrice d'espérance et de covariance, telle que la gaussienne multivariée. Dans cette situation particulière, pour toutes les combinaisons de  $t \neq t'$ , l'espérance de  $(M(t), W(t'))|X = x$  est donnée par le vecteur  $(E[M|T = t, X = x], E[W|T = t', X = x])$  et la matrice de covariance est identifiée par la matrice de covariance de  $(M|T = t, X = x)$  et  $(W|T = t', X = x)$  soit de  $(M|T = t, X = x)$  et  $(W|T = t, X = x)$ .

## 1.5 Variable de réponse continue

**Corollaire 1.5.1.** *Considérons  $K$  médiateurs et  $P$  covariables. Supposons le modèle linéaire suivant*

$$Z = \alpha_2 + \beta_2^\Gamma T + \xi_2^\Gamma X + \Upsilon_2 \quad (1.5.1)$$

$$Y = \alpha_3 + \beta_3 T + \gamma^\Gamma Z + \xi_3^\Gamma X + \epsilon_3, \quad (1.5.2)$$

où  $\alpha_2 = (\alpha_2^k)_{1 \leq k \leq K}$ ,  $\beta_2 = (\beta_2^k)_{1 \leq k \leq K}$ ,  $\gamma = (\gamma_k)_{1 \leq k \leq K}$ ,  $\xi_2 = (\xi_2^{kp})_{1 \leq k \leq K, 1 \leq p \leq P}$ ,  $\xi_3 = (\xi_3^p)_{1 \leq p \leq P}$ , et  $\Upsilon_2 = (\epsilon_2^k)_{1 \leq k \leq K} \sim \mathcal{N}(0, \Sigma_2)$  est le vecteur des résidus qui a pour matrice de covariance  $\Sigma_2 \in \mathbb{R}^K \times \mathbb{R}^K$  et  $\epsilon_3 \sim \mathcal{N}(0, \sigma_3^2)$ , avec  $\sigma_3 \in \mathbb{R}$ .

Si les  $K$  médiateurs sont non causalement corrélés, supposons que les corrélations par paires entre les médiateurs potentiels ne dépendent pas des traitements qui les régissent, c'est-à-dire que nous supposons la condition (1.4.2). Alors, sous SIMMA, l'effet indirect du  $k$ -ième médiateur est identifié et donné par :

$$\delta^k(0) = \delta^k(1) = \gamma_k \beta_2^k.$$

De plus, l'effet indirect joint est la somme des effets indirects moyens de chaque médiateur :

$$\delta^Z(t) = \sum_{k=1}^K \delta^k(t).$$

L'effet direct est également identifié et donné par

$$\zeta(0) = \zeta(1) = \beta_3.$$

Une preuve du Corollaire 1.5.1 est donnée dans l'annexe A.4. Notez qu'un résultat équivalent pour l'effet indirect joint est présenté dans VanderWeele and Vansteelandt (2014). Notez également que l'additivité des effets indirects individuels dans l'effet direct conjoint (c'est-à-dire  $\delta^Z(t) = \sum_k \delta^k(t)$ ) n'est valable que dans le contexte du corollaire 1.5.1, sinon ce n'est pas le cas.

Nous avons déjà observé que si les  $K$  médiateurs sont indépendants, l'équation de l'effet indirect marginal donné par le théorème 1.4.1 (analyse multiple) se réduit à l'équation donnée par le théorème 0.3.1 (analyse simple). Dans cette situation, le corollaire 1.5.1 implique que dans le cadre des LSEM, les effets indirects donnés par des analyses simples peuvent être sommés pour obtenir l'effet indirect joint. Évidemment, des analyses simples ne permettent pas d'évaluer un effet direct global, car selon le médiateur d'intérêt, chaque analyse simple conduira à un effet direct différent. Tous ces aspects seront illustrés par des simulations dans la section 1.7.

## 1.6 Variable de réponse binaire

### 1.6.1 Identification des effets causaux

Dans cette partie, nous supposons la variable de réponse binaire, comme c'est fréquemment le cas en épidémiologie. Comme pour la médiation simple, on considère soit la régression probit

$$\mathbb{P}(Y = 1|T, Z, X) = \Phi_{\mathcal{N}(0, \sigma_3^2)}(\alpha_3 + \beta_3 T + \gamma^\Gamma Z + \xi_3^\Gamma X),$$

soit la régression logistique

$$\text{logit}(\mathbb{P}(Y = 1|T, Z, X)) = \alpha_3 + \beta_3 T + \gamma^\Gamma Z + \xi_3^\Gamma X.$$

**Corollaire 1.6.1.** *Supposons le modèle suivant avec une variable de réponse binaire :*

$$Z = \alpha_2 + \beta_2^\Gamma T + \xi_2^\Gamma X + \Upsilon_2, \quad (1.6.1)$$

$$Y^* = \alpha_3 + \beta_3 T + \gamma^\Gamma Z + \xi_3^\Gamma X + \epsilon_3, \quad (1.6.2)$$

$$Y = \mathbb{1}_{\{Y^* > 0\}} \quad (1.6.3)$$

où  $\Upsilon_2 \sim \mathcal{N}(0, \Sigma_2)$  et où soit  $\epsilon_3 \sim \mathcal{N}(0, \sigma_3^2)$ , soit  $\epsilon_3 \sim \mathcal{L}(0, 1)$ . Supposons les  $K$  médiateurs indépendants, ou non corrélés de manière causale. Supposons de plus que les corrélations par paires entre médiateurs potentiels ne dépendent pas des traitements qui les régissent comme dans la condition (1.4.2). Sous SIMMA, les effets d'intérêt sont donnés par :

$$\begin{aligned} \delta^k(t) &= \int F_U \left( (\alpha_3 + \sum_{j=1}^K \gamma_j \alpha_2^j) + (\beta_3 + \sum_{j=1, j \neq k}^K \gamma_j \beta_2^j) t + \gamma_k \beta_2^k \times 1 + (\xi_3 + \sum_{j=1}^K \gamma_j \xi_2^{\Gamma j}) x \right) \\ &\quad - F_U \left( (\alpha_3 + \sum_{j=1}^K \gamma_j \alpha_2^j) + (\beta_3 + \sum_{j=1, j \neq k}^K \gamma_j \beta_2^j) t + \gamma_k \beta_2^k \times 0 + (\xi_3 + \sum_{j=1}^K \gamma_j \xi_2^{\Gamma j}) x \right) dF_X(x), \\ \delta^Z(t) &= \int F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times t + \sum_{k=1}^K \gamma_k \beta_2^k \times 1 + (\xi_3 + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right) \\ &\quad - F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times t + \sum_{k=1}^K \gamma_k \beta_2^k \times 0 + (\xi_3 + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right) dF_X(x), \\ \zeta(t) &= \int F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times 1 + (\sum_{k=1}^K \gamma_k \beta_2^k) \times t + (\xi_3 + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right) \\ &\quad - F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times 0 + (\sum_{k=1}^K \gamma_k \beta_2^k) \times t + (\xi_3 + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right) dF_X(x), \end{aligned}$$

où pour une régression probit nous avons

$$F_U(z) = \Phi \left( \frac{z}{\sqrt{\sigma_3^2 + \sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right),$$

et pour une régression logit nous avons

$$F_U(z) = \int_{\mathbb{R}} \Phi \left( \frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) \frac{e^{e_3}}{(1 + e^{e_3})^2} de_3.$$

Quand les médiateurs sont indépendants on a pour une régression probit

$$F_U(z) = \Phi \left( \frac{z}{\sqrt{\sigma_3^2 + \sum_{k=1}^K \gamma_k^2 \sigma_2^2}} \right),$$

et pour une régression logistique

$$F_U(z) = \int_{\mathbb{R}} \Phi \left( \frac{z - e_3}{\sqrt{\sum_{k=1}^K \gamma_k^2 \sigma_2^2}} \right) \frac{e^{e_3}}{(1 + e^{e_3})^2} de_3.$$

Une preuve du Corollaire 1.6.1 peut être trouvée dans l'annexe A.5.

## 1.6.2 Effets causaux à l'échelle des odds ratios

Les effets causaux à l'échelle des odds ratios sont plus facilement interprétable lorsque la variable de réponse est binaire, notamment par la communauté épidémiologique. L'utilisation de cette échelle est avantageuse notamment dans le cas de données cas témoins, souvent utilisées lorsque la variable de réponse représente une maladie dont la prévalence est faible (VanderWeele and Vansteelandt (2010)).

Les effets causaux à l'échelle des odds ratios sont définis par :

$$\begin{aligned} OR_Z^\delta(t) &= \frac{\mathbb{P}[Y(t, Z(1))]}{1 - \mathbb{P}[Y(t, Z(1))]} \times \frac{1 - \mathbb{P}[Y(t, Z(0))]}{\mathbb{P}[Y(t, Z(0))]} \\ OR^\zeta(t) &= \frac{\mathbb{P}[Y(1, Z(t))]}{1 - \mathbb{P}[Y(1, Z(t))]} \times \frac{1 - \mathbb{P}[Y(0, Z(t))]}{\mathbb{P}[Y(0, Z(t))]} \\ OR^\tau &= \frac{\mathbb{P}[Y(1, Z(1))]}{1 - \mathbb{P}[Y(1, Z(1))]} \times \frac{1 - \mathbb{P}[Y(0, Z(0))]}{\mathbb{P}[Y(0, Z(0))]} \end{aligned}$$

L'odds ratio de l'effet total se décompose avec les odds ratio des effets direct et indirect joint :

$$OR^\tau = OR_Z^\delta(t) OR^\zeta(1-t)$$

En présence de plusieurs médiateurs nous définissons les odds ratio des effets indirects individuels par :

$$OR_k^\delta(t) = \frac{\mathbb{P}[Y(t, M^k(1), W^k(t))]}{1 - \mathbb{P}[Y(t, M^k(1), W^k(t))]} \times \frac{1 - \mathbb{P}[Y(t, M^k(0), W^k(t))]}{\mathbb{P}[Y(t, M^k(0), W^k(t))]}$$

Nous remarquons que nous pouvons établir une relation entre les effets et les effets à l'échelle des odds ratios. Pour un effet causal défini par la différence de deux termes  $Eff_{caus} = A - B$ , l'effet causal correspondant à l'échelle des odds ratios se définit par  $OR^{Eff_{caus}} = \frac{A}{B} \times \frac{1-B}{1-A}$ .

On peut donc écrire  $Eff_{caus}$  en fonction de  $OR^{Eff_{caus}}$  :

$$Eff_{caus} = (OR^{Eff_{caus}} \frac{1-B}{1-A} - 1)B$$

Réciproquement on a :

$$(OR^{Eff_{caus}} = (Eff_{caus}B + 1) \frac{1-A}{1-B})$$

## 1.7 Étude de simulation

Dans cette section, nous validons nos résultats méthodologiques par des études empiriques. En particulier, nous comparons nos estimations des effets causaux de la médiation aux effets réels et aux estimations obtenues en exécutant de simples analyses de médiation, une pour chaque médiateur.

### 1.7.1 Méthode de simulation des données

À l'exception du cadre LSEM, il n'est en général pas simple d'obtenir les vraies valeurs de l'effet de médiation à partir d'un modèle génératif causal, c'est-à-dire un ensemble d'équations structurelles causales. Pour surmonter cette difficulté, nous commençons par simuler une grande base de données de valeurs pour le traitement  $T$  et pour tous les médiateurs contrefactuels  $M^k(t)$ , et les variables de réponses  $Y(t, M^1(t_1), \dots, M^K(t_K))$ , voir Table 1.1 pour un exemple. Ensuite, nous calculons simplement les effets indirects  $\delta^k(t)$  et  $\delta^Z(t)$  et l'effet direct  $\zeta(t)$  comme des moyennes, selon les définitions données dans la section 1.2. La grande taille de l'ensemble de données garantit que ces estimations de Monte-Carlo peuvent être considérées comme les vraies valeurs. Dans cette étude, nous générons un ensemble de données de  $10^6$  observations, de sorte que l'erreur d'estimation est aussi petite que 0,2% de l'écart type de l'effet d'intérêt.

Afin d'obtenir un sous-ensemble d'observations pour tester les méthodes d'estimation, nous échantillonnons  $N$  individus (i.e. lignes)  $i = 1, \dots, N$  et pour chacun d'eux nous ne sélectionnons que les valeurs  $Y(T_i, Z_i(T_i))$  et  $Z_i(T_i)$  correspondant à la valeur spécifique de  $T_i$ . Plus précisément :

- si  $T_i = 0$  on prend  $Z_i = (M_i^1, \dots, M_i^k) = (M_i^1(0), \dots, M_i^k(0)) = Z_i(0)$  et  $Y_i = Y_i(0, Z_i(0))$ ,
- si  $T_i = 1$  on prend  $Z_i = (M_i^1, \dots, M_i^k) = (M_i^1(1), \dots, M_i^k(1)) = Z_i(1)$  et  $Y_i = Y_i(1, Z_i(1))$ .

Les tableaux 1.1 et 1.2 illustrent la procédure de simulation.

Pour plusieurs modèles de simulations, nous estimons les différents effets d'intérêt au moyen de l'algorithme général pour plusieurs médiateurs décrit dans la section 3.2.3. Nous comparons nos estimations avec les vraies valeurs et les estimations de deux analyses simples (une pour chaque médiateur) obtenues avec le package `mediation`. Parce qu'en toute généralité  $\delta^k(1) \neq \delta^k(0)$  et  $\zeta(1) \neq \zeta(0)$ , et par souci de simplicité, nous nous concentrons sur des effets moyens tels que  $\delta = (\delta(0) + \delta(1))/2$  et  $\zeta = (\zeta(1) + \zeta(0))/2$ . Notez que pour une variable de réponse continue et en l'absence d'interaction entre le traitement et les médiateurs, les corollaires 0.3.2 et 1.5.1 impliquent que  $\delta^k(1) = \delta^k(0)$  et  $\zeta(1) = \zeta(0)$ . Pour chaque médiateur la proportion médiée est définie par  $PM^k = \delta^k/\tau$ .

$T$	$M(0)$	$M(1)$	$W(0)$	$W(1)$	$Y(1, M(1), W(1))$	$Y(1, M(1), W(0))$	$Y(1, M(0), W(1))$
<b>0</b>	<b>0.28</b>	1.08	<b>0.53</b>	1.43	2.42	1.79	1.94
<b>0</b>	<b>0.42</b>	1.22	<b>-1.80</b>	-0.90	1.41	0.78	0.93
<b>1</b>	0.63	<b>1.43</b>	0.03	<b>0.93</b>	<b>1.87</b>	1.24	1.39
<b>1</b>	0.75	<b>1.55</b>	2.24	<b>3.14</b>	<b>2.95</b>	2.32	2.47
$Y(0, M(1), W(1))$		$Y(1, M(0), W(0))$		$Y(0, M(1), W(0))$		$Y(0, M(0), W(1))$	
2.02		1.31		1.39		1.54	
1.01		0.30		0.38		<b>0.53</b>	
1.47		0.76		0.84		0.99	
2.55		1.84		1.92		2.07	

TABLE 1.1 – Données contrefactuelles simulées avec deux médiateurs indépendants.

$T$	$M$	$W$	$Y$
0	0.28	0.53	0.91
0	0.42	-1.80	-0.09
1	1.43	0.93	1.87
1	1.55	3.14	2.95

TABLE 1.2 – Données observées simulées avec deux médiateurs indépendants. Les observations ont été extraites de la table 1.1.

À des fins de comparaison, nous analysons les simulations avec notre méthode de médiation multiple, avec l'approche consistant à exécuter une analyse simple en parallèle

(Imai and Yamamoto (2013)) et avec la méthode décrite par VanderWeele and Vansteelandt (2014) que nous appelons V&V dans les résultats. Dans ce dernier cas, nous rapportons non seulement les estimations des effets indirects et directs conjoints, mais aussi les estimations des effets indirects spécifiques aux médiateurs, même si les auteurs avaient déjà clairement expliqué que la corrélation entre médiateurs conduirait à un biais.

### 1.7.2 Limites des analyses simples indépendantes en présence d'une cause commune affectant les médiateurs.

Dans cette section, les données sont générées sous le modèle décrit dans la figure 1.2(a), où la dépendance entre les deux médiateurs est induite par la variable de prétraitement  $U$ . Plus précisément, les variables sont simulées selon les distributions suivantes ( $N = 1000$ ) :

- $T$  suit une distribution de Bernoulli  $\mathcal{B}(0.3)$
- $U$  suit une distribution normale  $\mathcal{N}(0, 1)$
- la distribution conditionnelle des deux médiateurs contrefactuels sachant  $T$  et  $U$  est

$$M^1(t, u) \sim \mathcal{N}(1 + 4t + 2u, 1)$$

$$M^2(t, u) \sim \mathcal{N}(2 + 6t + 3u, 1)$$

- la variable de réponse contrefactuelle suit la distribution normale

$$Y(t, M^1(t'), M^2(t'')) \sim \mathcal{N}(1 + 10t + 5M^1(t') + 4M^2(t''), 1).$$

Notons que la corrélation entre les deux médiateurs conditionnellement au traitement (et non à  $U$ , Figure 6(b)), est égale à 0,7.

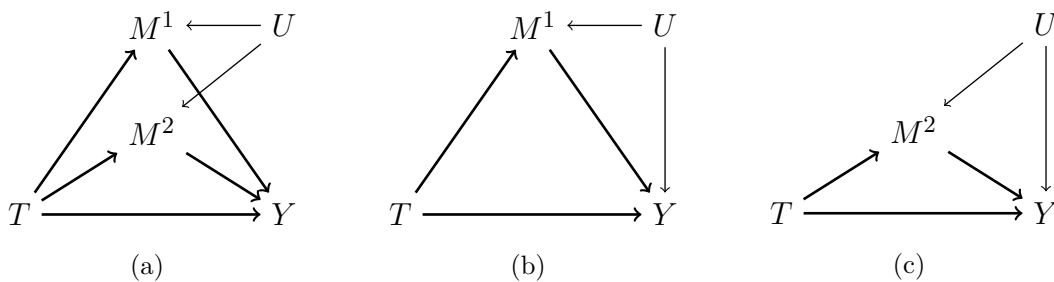


FIGURE 1.2 – Analyses de médiation multiples et simples,  $U$  étant observé. Les données sont simulées selon le modèle en (a).

Lorsque nous avons deux médiateurs causalement indépendants et que  $U$  est observé, l'approche de Imai and Yamamoto (2013) consiste à effectuer deux analyses simples



comme dans les figures 1.2(b) et 1.2(c). Cependant, lorsque  $U$  n'est pas observé, la situation est comme dans la Figure 6(b) avec des médiateurs montrant une corrélation résiduelle. Dans ce cas, la réalisation d'analyses simples séparées n'est pas appropriée car les hypothèses d'ignorabilité séquentielle (A.2) et (A.3) sont violées (Imai and Yamamoto (2013)).

Nous illustrons ici ce problème par des simulations. Les intervalles de confiance à 95% des analyses simples sont donnés par la méthode d'Imai. Ceux de la méthode de VanderWeele et Vansteelandt sont obtenues avec la delta-méthode. Et ceux de l'analyse de médiation multiple sont obtenus avec notre méthode (voir 3.2.3).

Effets	Valeur	Analyse Simple avec $M^1$	Analyse Simple avec $M^2$	V&V	Analyse Multiple
$\delta^Z$	44	NA	NA	45.20 [40.30 ; 50.10]	44.40 [43.30 ; 45.50]
$PM^Z$	0.81	NA	NA	0.82	0.81 [0.81 ; 0.82]
$\delta^1$	20	19.40 [17.70 ; 21.10]	NA	20.40 [15.90 ; 24.90]	20.50 [19.60 ; 21.40]
$PM^1$	0.37	0.36 [0.33 ; 0.39]	NA	0.37	0.38 [0.36 ; 0.40]
$\delta^2$	24	NA	21.60 [18.80 ; 24.60]	25.30 [18.90 ; 31.60]	23.90 [23.10 ; 24.66]
$PM^2$	0.44	NA	0.40 [0.34 ; 0.45]	0.45	0.44 [0.42 ; 0.45]
$\zeta$	10	35 [33.40 ; 36.60]	32.80 [29.70 ; 35.65]	9.90 [9.20 ; 10.60]	9.90 [9.70 ; 10.20]
$\tau$	54	54.40 [53.40 ; 55.40]	54.40 [53.30 ; 55.50]	55.10 [50.20 ; 59.90]	54.43 [53.20 ; 55.50]

TABLE 1.3 – Comparaison des méthodes en ajustant pour  $U$  lorsque toutes les variables de la figure 1.2(a) sont observées. Pour rappel  $\delta^a$  : effet médié par  $a$ ,  $PM^a$  : proportion médiée par  $a$ ,  $\zeta$  : effet direct et  $\tau$  : effet total.

Effets	Valeur	Analyse Simple avec $M^1$	Analyse Simple avec $M^2$	V&V	Analyse Multiple
$\delta^Z$	44	NA	NA	46.80 [40.30 ; 53.40]	43.20 [40.20 ; 46.20]
$PM^Z$	0.81	NA	NA	0.80	0.81 [0.80 ; 0.82]
$\delta^1$	20	38.40 [34.50 ; 42.30]	NA	41.60 [32.40 ; 50.70]	20 [18.00 ; 22.20]
$PM^1$	0.37	0.72 [0.68 ; 0.75]	NA	0.73	0.37 [0.33 ; 0.42]
$\delta^2$	24	NA	40.86 [37 ; 44.90]	44.70 [36.60 ; 52.80]	23.20 [20.90 ; 25.50]
$PM^2$	0.44	NA	0.76 [0.73 ; 0.79]	0.78	0.43 [0.38 ; 0.48]
$\zeta$	10	14.70 [13.20 ; 16.40]	12.46 [10.90 ; 13.90]	10.10 [9.90 ; 10.40]	9.90 [9.70 ; 10.20]
$\tau$	54	53.20 [49.20 ; 57]	53.30 [49.50 ; 57.20]	57 [50.40 ; 63.60]	53.20 [50.20 ; 56.20]

TABLE 1.4 – Comparaison des méthodes en n'ajustant pas pour  $U$  : les données sont générées comme dans la figure 1.2(a) mais analysées comme si  $U$  n'était pas observé. Pour rappel  $\delta^a$  : effet médié par  $a$ ,  $PM^a$  : proportion médiée par  $a$ ,  $\zeta$  : effet direct et  $\tau$  : effet total.

Comme prévu, les tableaux 1.3 et 1.4 montrent que des analyses simples ajustées pour  $U$  donnent des estimations précises des effets indirects (mais évidemment pas de l'effet

direct), alors qu'elles donnent un biais d'estimation lorsque  $U$  n'est pas prise en compte. Au contraire, notre méthode donne des estimations précises de tous les effets avec ou sans prise en compte de  $U$ , montrant qu'il est toujours possible de mener une analyse de médiation pour estimer tous les effets même lorsque  $U$  n'est pas observé. Notez que l'ajustement pour  $U$  se fait en prenant en compte  $U$  dans les modèles de régression des médiateurs et de la variable de réponse.

Dans la sous-section suivante, nous supposons que  $U$  n'est pas observée, comme c'est souvent le cas en pratique.

### 1.7.3 Étude empirique des propriétés des estimateurs.

La section précédente a illustré notre méthode sur une seule simulation. Dans cette section, nous effectuons une étude basée sur la simulation pour évaluer les propriétés des estimateurs proposés. Plus précisément, nous calculons le biais, la probabilité de couverture de l'intervalle de confiance, l'erreur quadratique moyenne (MSE) et la variance de nos estimateurs en faisant des moyennes sur les 200 simulations pour chaque paramètre considéré. Nous comparons les estimations de plusieurs analyses simples, une pour chaque médiateur, aux estimations obtenues avec notre analyse de médiation multiple pour différents niveaux de corrélation. Nous considérons deux modèles de simulation causale tenant compte de deux types de variable de réponse (continue et logit binaire), et deux médiateurs continus sans relation causale. Les médiateurs corrélés de manière non causale, Figure 6(b), sont simulés à partir d'une distribution normale multivariée avec une matrice de covariance fixe. Les détails des modèles de simulation se trouvent dans l'annexe A.6.

Des simulations selon le modèle 1 (variable de réponse continue) sont exécutées pour différentes valeurs de corrélation entre les médiateurs et l'augmentation de la taille de l'échantillon de données observées (50, 200, 500 et 1000). Les résultats pour le biais et la probabilité de couverture peuvent être vus dans les figures 1.3, 1.4, 1.5 et 1.6. Cette figure montre clairement que notre approche permet une estimation non biaisée, contrairement aux analyses simples, tant pour les effets directs qu'indirects.

L'intervalle de confiance empirique à 95% donné par notre méthode contient la valeur réelle dans environ 95% des essais, tant pour les effets indirects que directs et quelle que soit la corrélation entre les médiateurs. Au contraire, des analyses simples n'obtiennent des couvertures satisfaisantes que lorsque la corrélation est presque nulle.

Comme prévu, les estimateurs des effets indirects individuels obtenus avec la méthode de VanderWeele and Vansteelandt (2014) ont le même comportement que les estimations d'analyse simples. De plus, les estimateurs des effets indirects et directs conjoints par la méthode de VanderWeele and Vansteelandt (2014) se comportent de la même manière que les nôtres, sauf que la probabilité de couverture est constante pour leur méthode.

Notons par ailleurs que nos estimateurs ont une faible variance et une faible MSE pour des échantillons de plus de 200 observations.

Des simulations ont également été effectuées pour le modèle 2 (variable de réponse binaire) pour différentes valeurs de corrélation entre les médiateurs avec 1000 données d'observations. Comme l'illustre la figure 4.5 de l'annexe, les résultats pour le biais, la probabilité de couverture, la variance et la MSE confirment que nos estimateurs sont sans biais et ont une faible variance. De plus, la probabilité de couverture attendue est supérieure à celle obtenue dans le cas de plusieurs analyses simples. Notez que seule la probabilité de couverture des effets indirects par l'intermédiaire des médiateurs individuels lorsque la corrélation est positive n'est pas satisfaisante. Cela est probablement dû aux estimations de variance très faibles.

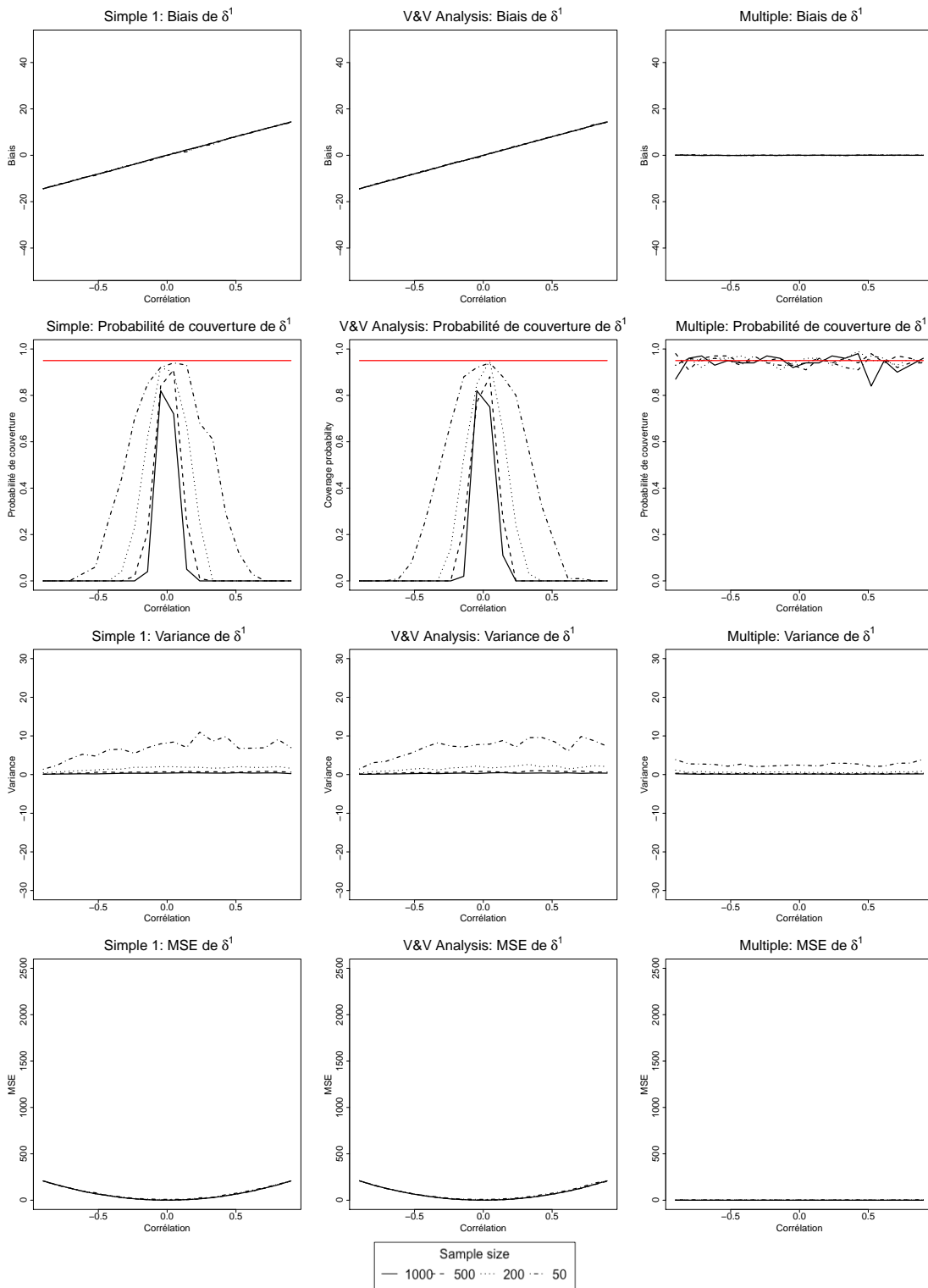


FIGURE 1.3 – Modèle 1 (variable de réponse continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance des estimateurs à effet indirect  $\hat{\delta}^1$  calculés comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est  $Bias = \Theta - \mathbb{E} \left[ \hat{\Theta} \right]$ .

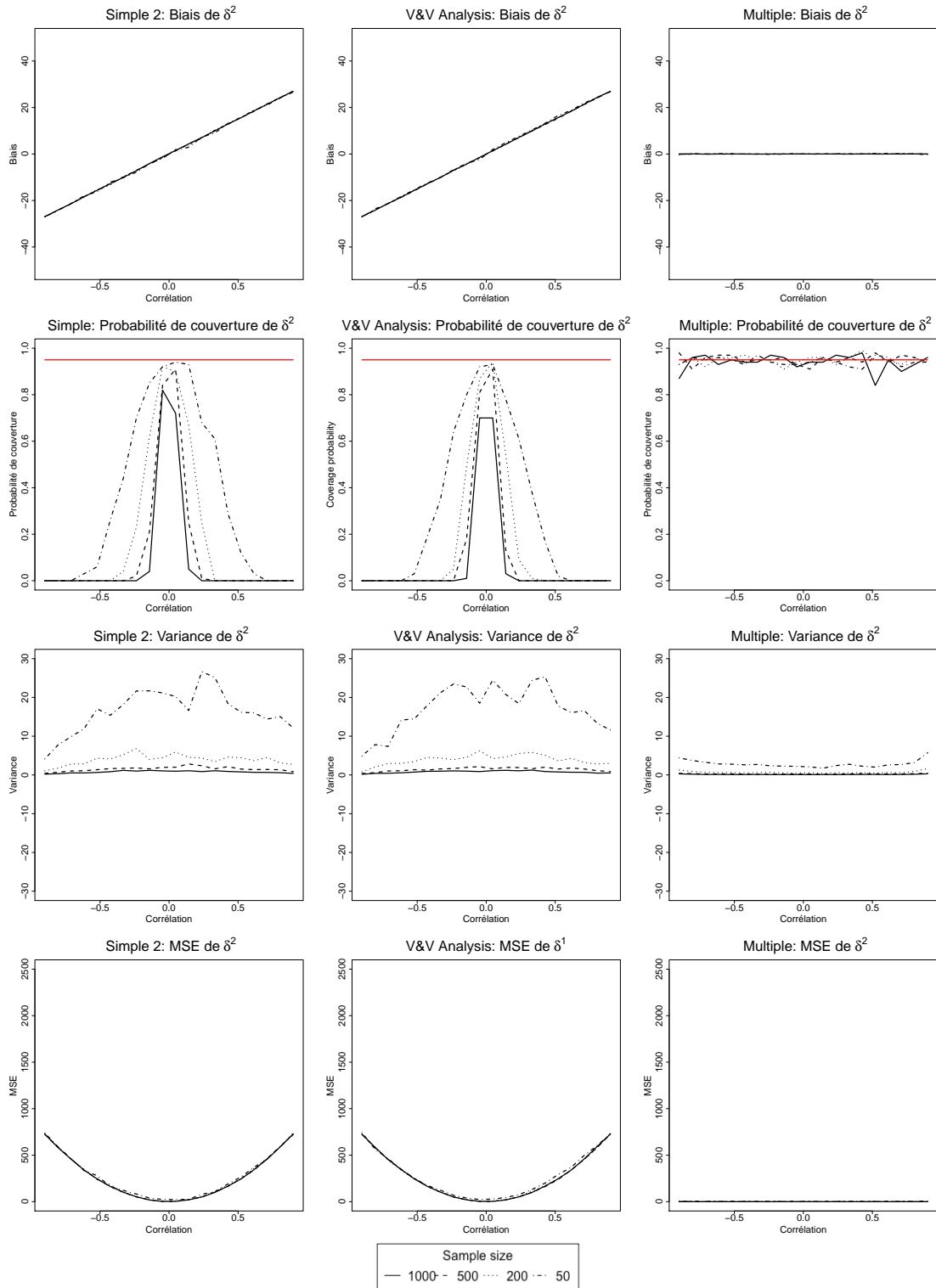


FIGURE 1.4 – Modèle 1 (résultat continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance des estimateurs des effets indirects  $\hat{\delta}^2$  calculés comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est  $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ .

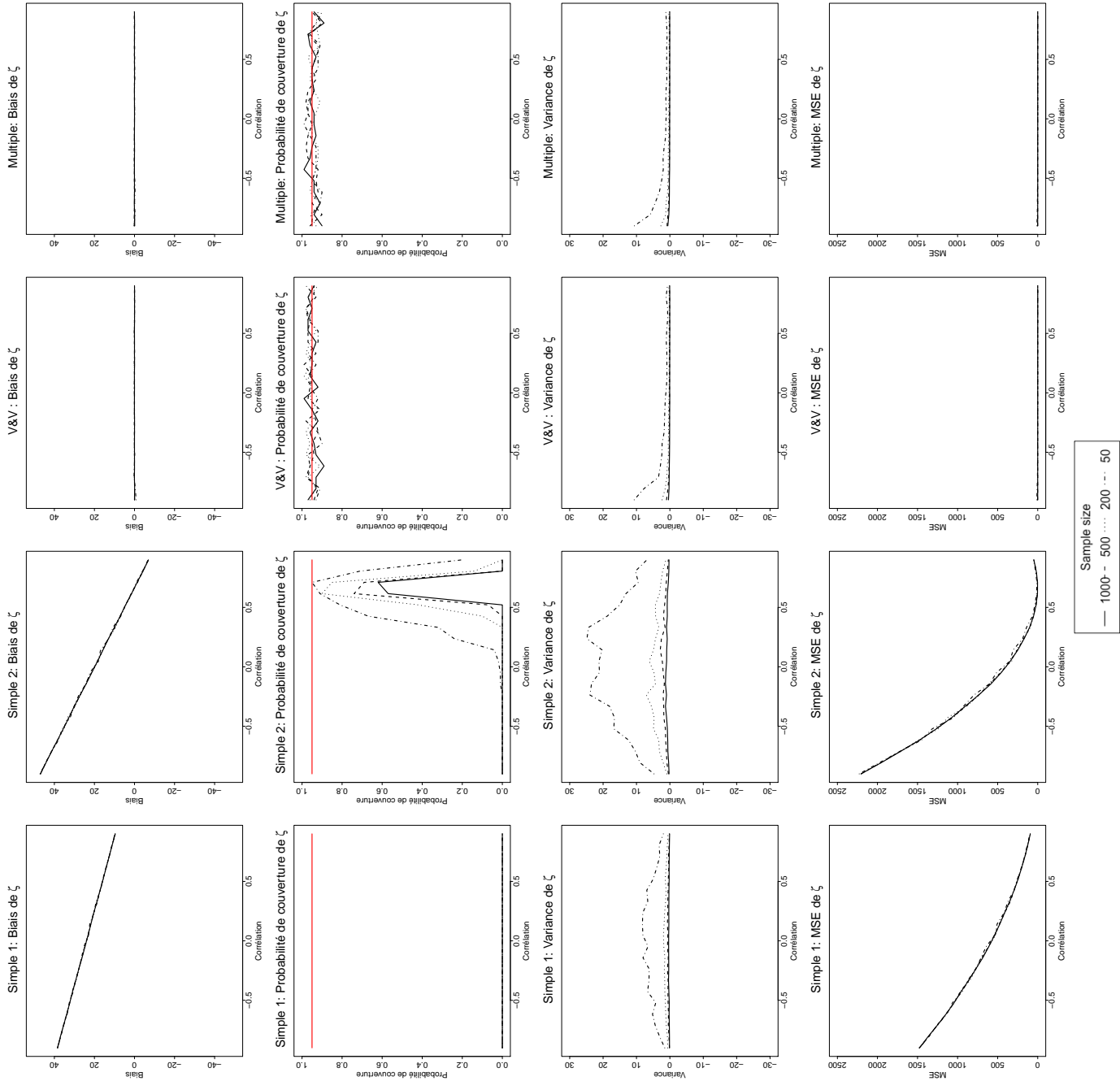


FIGURE 1.5 – Modèle 1 (variable de réponse continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance de l'estimateur de effet direct  $\hat{\zeta}$  calculé comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est  $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ .

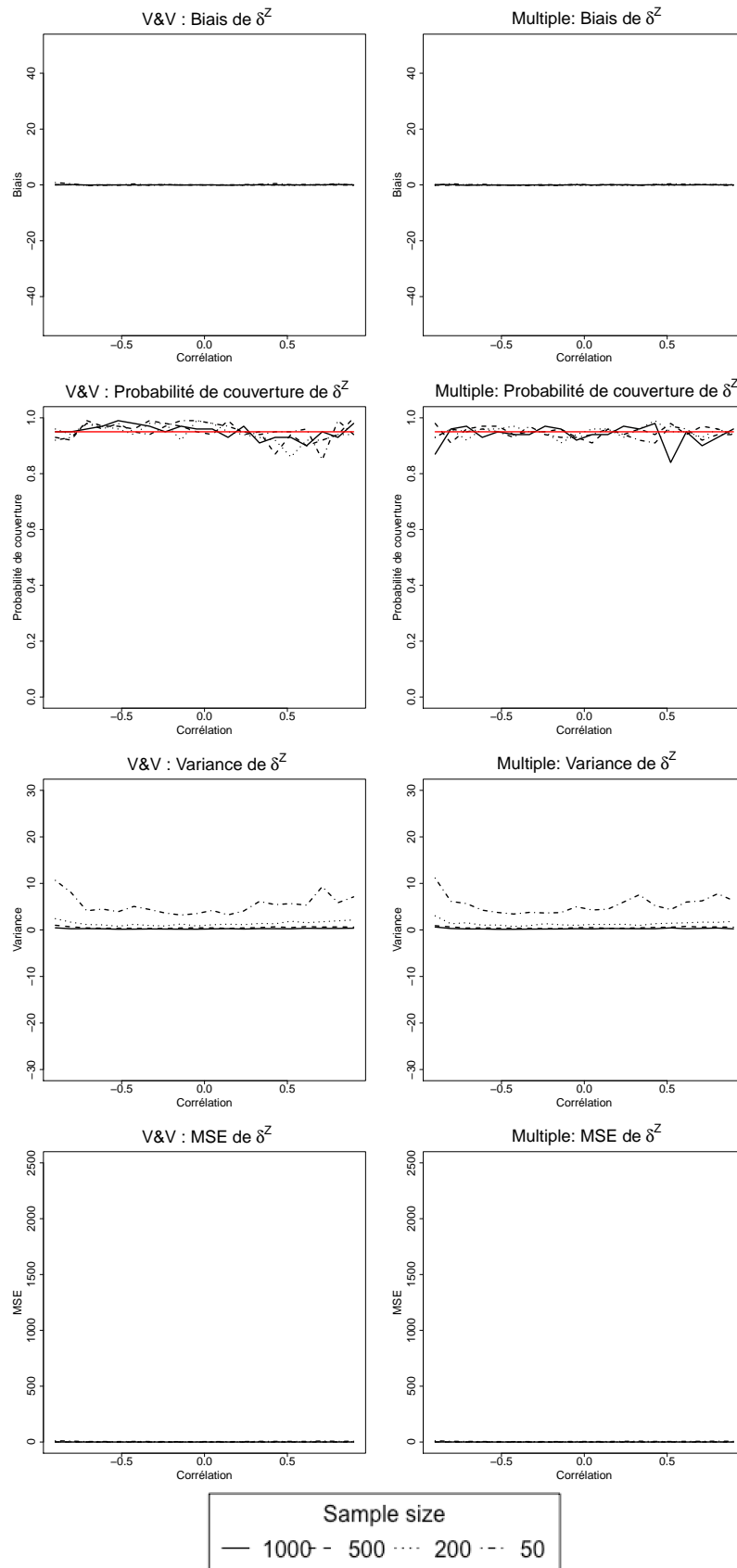


FIGURE 1.6 – Modèle 1 (variable de réponse continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance de l'estimateur de l'effet indirect joint  $\delta^Z$  calculé comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est  $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ .

# Chapitre 2

## Grande dimension

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>47</b>
<b>2.2</b>	<b>Classification des médiateurs</b>	<b>48</b>
2.2.1	Classification spectrale	49
2.2.2	Analyse de médiation par blocs	51
2.2.3	Étude de simulation	53
<b>2.3</b>	<b>Sélection de médiateur en grande dimension.</b>	<b>55</b>
2.3.1	MAHI en présence d'un seul traitement	57
2.3.2	Quelques méthodes existantes	61
2.3.3	Comparaison des méthodes	64

---

## 2.1 Introduction

L'amélioration de la collecte et de la mesure des données a fait que les données de grande dimension sont largement utilisées dans de nombreux domaines de la recherche scientifique. Dans le cadre de l'analyse de médiation, les données de grande dimension correspondent à des données où le nombre de médiateurs *potentiels* (qui ici ne renvoie pas à la notion de contrefactuelle) est important et en particulier supérieur au nombre d'observations (Figure 8). On peut alors se demander : *"Doit-on s'intéresser à l'ensemble des médiateurs ou en sélectionner certains ? Quels sont leurs effets ? Peut-on éventuellement classifier ces médiateurs ?"*

La première partie de ce chapitre aborde la question de la classification de médiateurs. Nous proposons une méthode de classification basée sur les corrélations entre les médiateurs, corrélations conditionnelles au traitement et aux covariables. Cette méthode



de classification se base sur une approche de classification spectrale expliquée par von Luxburg (2007).

La deuxième partie de ce chapitre décrit une nouvelle méthode de sélection reposant sur deux étapes. La première se base sur une approche Group-LASSO de façon à réduire le nombre de médiateurs à un nombre inférieur au nombre de données. Nous utilisons ensuite notre méthode présentée aux Chapitre 1 pour calculer les effets médiés et leur p-valeur permettant ainsi d'effectuer un test multiple sur les effets médiés. Nous nous comparons à trois méthodes de sélection de médiateurs existantes HIMA, HDMA et CMF. Notre méthode permet de plus la prise en compte simultanée de plusieurs traitements. Sélectionnant des médiateurs communs pour plusieurs traitement identifiés sur une variable de réponse. L'ensemble de notre méthode est implémenté et disponible dans notre package présenté au Chapitre 3.

## 2.2 Classification des médiateurs

Pour réduire le nombre d'effet à considérer quand le nombre de médiateur  $P$  est grand, nous proposons de considérer les médiateurs par groupe et de considérer alors les effets médiés par ces groupes de médiateurs. On choisit de définir ces groupes par rapport à la corrélation entre les médiateurs au vu du biais qu'il peut y avoir dans l'estimation des effets causaux, comme illustré au Chapitre 1.

On souhaite donc, à partir d'un nombre  $P$  de médiateurs potentiels, définir différents blocs de médiateur par rapport à leur corrélation. On se propose ici d'utiliser la théorie des graphes et le clustering spectral pour déterminer les différents blocs puis de faire de l'analyse de médiation par blocs, c'est-à-dire quantifier l'effet médié par chaque bloc. En pratique on se retrouve pour chaque bloc  $B_k$  avec un nombre  $P_k$  de médiateurs sans savoir à quels blocs appartiennent ces médiateurs ni le nombre  $K$  de blocs.

On se place dans le modèle suivant illustré dans la Figure 2.1 :

—  $T$  suit une loi de Bernoulli.

— Les médiateurs du bloc  $B_k$  suivent la loi  $M_k^i \sim \mathcal{N}(\mu_k^i + \alpha_k^i T + \kappa_k^i U_k, \sigma_k^i)$  pour  $i \in \{1, \dots, P_k\}$ .

— La variable de réponse suit une loi normale  $Y \sim \mathcal{N}(\mu_0 + \alpha_0 T + \sum_k^K \beta_k B_k, \sigma_0)$

où  $B_k = (M_k^1, \dots, M_k^{P_k})^\Gamma$  et  $\beta_k = (\beta_k^1, \dots, \beta_k^{P_k})^\Gamma$ , avec  $P_k$  le nombre de médiateurs dans le bloc  $B_k$ .

Ici on observe le traitement  $T$ , un nombre  $P = \sum_k P_k$  de médiateurs et la variable de réponse  $Y$ . Les variables de pré-traitement  $U_k$  ne sont pas observées. Dans un même bloc  $B_k$  on retrouve les médiateurs qui ont  $U_k$  pour cause commune et/ou des médiateurs qui

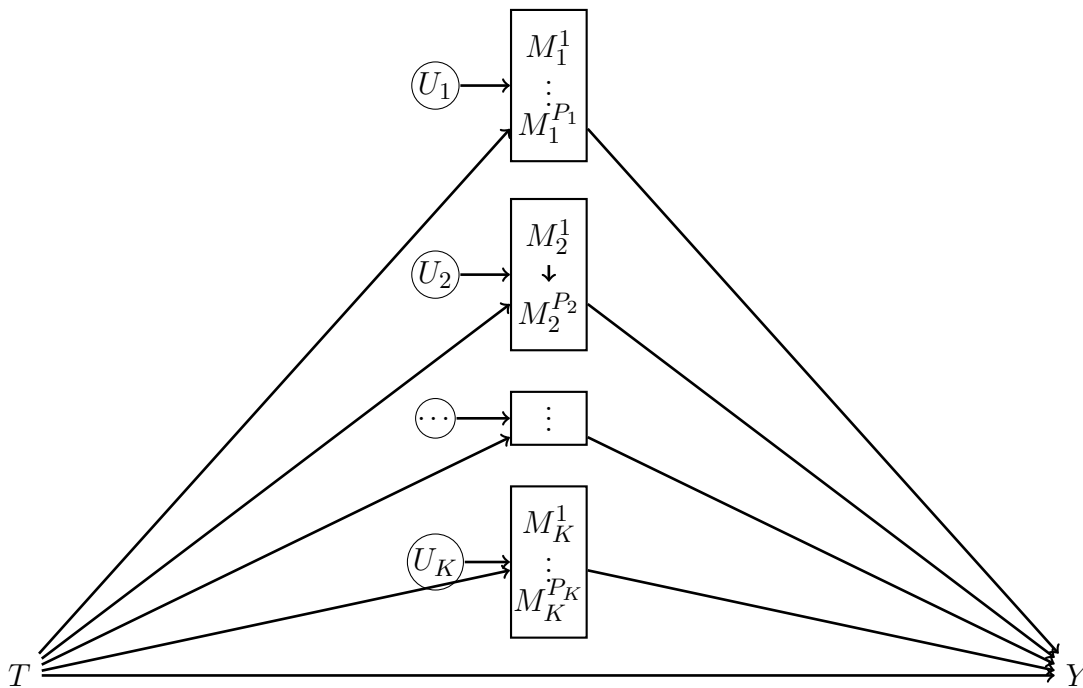


FIGURE 2.1 – Médiation par bloc

sont causalement liés. On suppose que les blocs sont indépendants, c'est-à-dire que les  $U_k$  sont indépendants entre eux et qu'il n'y a pas de relation causale entre les blocs. La matrice de variance-covariance entre les médiateurs  $\Sigma$  a, par définition du modèle, une structure diagonale par bloc :

$$\Sigma = \begin{pmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & \ddots & \\ & & & \Sigma_K \end{pmatrix}$$

### 2.2.1 Classification spectrale

Considérons le graphe non orienté dont les nœuds sont les médiateurs et dont les arêtes relient les médiateurs, avec un poids lié à leur corrélation ou autre mesure de similarité. Sous le modèle étudié, ce graphe est une union de  $K$  groupes de taille  $P_1, \dots, P_k$ , les médiateurs issus de deux groupes différents étant indépendants. Une manière de déterminer les groupes de médiateurs, tels que décrits dans la Figure 2.1, est alors de considérer l'estimation  $G$  de ce graphe obtenue en considérant le graphe complet des médiateurs avec des arêtes pondérées par leur similarité observée. On réalise ensuite une classification de ce graphe, c'est-à-dire un partitionnement des sommets en groupes tel que les connections

de poids fort aient préférentiellement lieu entre les sommets d'un même groupe. Pour cela, nous proposons d'utiliser la classification spectrale (von Luxburg (2007)), ce qui donne lieu à l'algorithme suivant. On présente l'ensemble des étapes de l'algorithme que nous adaptions à notre problème.

**Algorithme.** *Classification des médiateurs*

1. **Matrice de similarité S.** Notons  $\rho_{ij}$  la corrélation entre  $M_i$  et  $M_j$ ,  $\rho_{ij} = \text{cor}(M_i, M_j)$ . La matrice de similarité  $S = (s_{ij})_{1 \leq i, j \leq Q}$  permet de modéliser les relations entre les médiateurs  $M_i$  et  $M_j$ . Dans notre cas les médiateurs faisant partie du même bloc auront une mesure de similarité élevée, et une mesure plus petite entre deux médiateurs appartenant à des blocs différents. La mesure de similarité,  $s_{ij}$ , devant être positive, on choisit l'information mutuelle,  $I(A, B)$ , qui est une mesure d'indépendance entre deux variables donnée par :

$$I(A, B) = \int_{\mathbb{R}} \int_{\mathbb{R}} f_{(A,B)}(a, b) \log \frac{f_{(A,B)}(a, b)}{f_A(a) f_B(b)} da db$$

Dans le cas où les variables, ici les médiateurs, suivent une loi normale on a :

$$s_{ij} = \begin{cases} -\frac{1}{2} \log(1 - \rho_{ij}^2) & \text{si } \rho_{ij} \leq 0.99 \\ -\frac{1}{2} \log(1 - 0.99^2) & \text{sinon} \end{cases}$$

Lorsque  $\rho_{ij} \geq 0.99$  nous choisissons de fixer l'information mutuelle entre  $M_i$  et  $M_j$  à  $I(M_i, M_j) = -\frac{1}{2} \log(1 - 0.99^2)$  pour éviter d'avoir des valeurs qui explosent à l'infini.

2. **Matrice d'adjacence W.** La matrice d'adjacence pondérée  $W = (w_{ij})_{1 \leq i, j \leq Q}$ , contient des valeurs positives. Comme  $G$  est un graphe non-orienté,  $w_{ij} = w_{ji}$ . De plus si  $M_i$  et  $M_j$  n'appartiennent pas au même block alors  $w_{ij} = 0$ . La matrice  $W$  sera construite à partir de la matrice de similarité  $S$  :

$$w_{ij} = \begin{cases} s_{ij} & \text{si } s_{ij} \geq \text{seuil} \\ 0 & \text{sinon} \end{cases}$$

On décide d'avoir un graphe totalement connecté, on choisit donc 0 comme seuil. Donc  $W = S$ .

3. **Matrice de degré D** La matrice de degrés  $D = (d_i)_{1 \leq i \leq Q}$ , est une matrice diagonale et est déterminée par

$$d_i = \sum_{j=1}^Q w_{ij}$$

$d_i$  est le degré du médiateur  $M_i$ , il est obtenu en sommant le poids des arêtes liant  $M_i$  au médiateurs qui lui sont adjacents.

4. **Graphe laplacien non normalisé.** La matrice  $L$  est obtenue avec  $L = D - W$ .
5. **Matrice des vecteurs propres  $U$ .** On détermine les  $K$  plus petites valeurs propres de  $L$  et les  $K$  vecteurs propres correspondants. On concatène en colonne ces vecteurs propres dans la matrice  $U$ .
6.  **$K$ -means.** La matrice  $U$  contient  $P$  lignes, chaque ligne correspond à un médiateur. En appliquant l'algorithme des  $k$ -means sur les lignes de  $U$  on peut classifier les médiateurs dans les  $K$  blocs.

Une fois les blocs déterminés on modifie  $\Sigma$ , la matrice de variance-covariance entre les médiateurs, en fixant à 0 les covariances entre les médiateurs appartenant à des blocs différents. On a ainsi indépendance entre les blocs. La structure de la matrice  $\Sigma$  est donc donnée par les données.

Nous nous sommes placés dans un cadre où les médiateurs sont normalement distribués et seule la construction de la matrice de similarité  $S$  à partir des corrélations utilise cette hypothèse. Dans tous les autres cas, il est possible d'utiliser la classification spectrale en adaptant la définition de la similarité comme on peut le voir au Chapitre 3 dans la sous-section 3.2.2.

Les propriétés de la matrice  $L$  sont données dans l'annexe B.1

### 2.2.2 Analyse de médiation par blocs

Une fois les blocs définis on peut calculer les effets médiés par les différents blocs. Ces effets sont définis de façon analogue aux quantités d'intérêt définis au Chapitre 1. On note  $B_k$  l'ensemble des médiateurs appartenant au  $k$ -ième bloc et  $W_k$  le complémentaire de  $B_k$ . De plus, on note  $\delta^{B_k}$  l'effet médié par le bloc  $B_k$ .

**Théorème 2.2.1.** *Sous les hypothèses SIMMA et sous l'hypothèse que les  $K$  blocs de médiateurs sont indépendants, on a le résultat suivant :*

*L'effet médié par le bloc  $B_k$  est identifiable non-paramétriquement et donné par :*

$$\begin{aligned}
\delta^{B_k}(t) &= \mathbb{E}[Y(t, B_k(1), W_k(t))] - \mathbb{E}[Y(t, B_k(0), W_k(t))] \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y|B_k = b, W_k = w, T = t] \\
&\quad \{dF_{(B_k(1), W_k(t))}(b, w) - dF_{(B_k(0), W_k(t))}(b, w)\} \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y|B_k = b, W_k = w, T = t] \\
&\quad \{dF_{(B_k(1))}(b)dF_{(W_k(t))}(w) - dF_{(B_k(0))}(b)dF_{(W_k(t))}(w)\} \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y|B_k = b, T = t] \{dF_{(B_k|T=1)}(b) - dF_{(B_k|T=0)}(b)\},
\end{aligned}$$

**Corollaire 2.2.2.** *Soit le modèle avec  $K$  blocs de médiateurs suivant :*

$$M_k^i = \mu_k^i + \alpha_k^i T + \gamma_k^i X + \epsilon_k^i, \quad (2.2.1)$$

$$Y = \mu_0 + \alpha_0 T + \sum_k \beta_k Z + \gamma_0 X + \epsilon_0, \quad (2.2.2)$$

où  $1 \leq i \leq P_k, 1 \leq k \leq K, B_k = (M_k^1, \dots, M_k^{P_k})^\Gamma$  et  $\beta_k = (\beta_k^1, \dots, \beta_k^{P_k})^\Gamma$ , avec  $P_k$  le nombre de médiateurs dans le bloc  $B_k$ . On suppose  $\epsilon_k^i \sim \mathcal{N}(0, \sigma_k^i)$ , avec  $\sigma_k^i \in \mathbb{R}$  et  $\epsilon_k^i \sim \mathcal{N}(0, \sigma_0)$ , avec  $\sigma_0 \in \mathbb{R}$ .

On suppose que les  $K$  blocs de médiateurs sont indépendants. Sous les hypothèses SIMMA l'effet médié par le bloc  $B_k$  est non-paramétriquement identifié et donné par :

$$\delta^{B_k}(0) = \delta^{B_k}(1) = \sum_i^{P_k} \alpha_k^i \beta_k^i.$$

De plus, l'effet indirect joint est la somme des effets indirects médiés par chaque bloc :

$$\delta^Z(t) = \sum_{k=1}^K \delta^{B_k}(t).$$

L'effet direct est aussi identifié et donné par

$$\zeta(0) = \zeta(1) = \alpha_0.$$

**Corollaire 2.2.3.** *Soit le modèle avec la variable de réponse binaire suivante :*

$$Z = \alpha_2 + \beta_2^\Gamma T + \Upsilon_2, \quad (2.2.3)$$

$$Y^* = \alpha_3 + \beta_3 T + \gamma^\Gamma Z + \epsilon_3, \quad (2.2.4)$$

$$Y = \mathbb{1}_{\{Y^* > 0\}} \quad (2.2.5)$$

où  $\Upsilon_2 \sim \mathcal{N}(0, \Sigma)$  et où  $\epsilon_3 \sim \mathcal{N}(0, \sigma_3^2)$  ou  $\mathcal{L}(0, 1)$ .

Nous supposons que les  $K$  blocs de médiateurs sont indépendants. Sous les hypothèses SIMMA, les effets d'intérêts sont donnés par :

$$\begin{aligned}
\delta^{B_k}(t) &= \int F_U \left( (\alpha_3 + \sum_{j=1}^K \gamma_j \alpha_2^j) + (\beta_3 + \sum_{j=1, j \notin B_k}^K \gamma_j \beta_2^j) t + \sum_{j=1, j \in B_k}^K \gamma_j \beta_2^j \times 1, P_k \right) \\
&\quad - F_U \left( (\alpha_3 + \sum_{j=1}^K \gamma_j \alpha_2^j) + (\beta_3 + \sum_{j=1, j \notin B_k}^K \gamma_j \beta_2^j) t + \sum_{j=1, j \in B_k}^K \gamma_j \beta_2^j \times 0, P_k \right) dX, \\
\delta^Z(t) &= \int F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times 1 + \sum_{k=1}^K \gamma_k \beta_2^k t, K \right) \\
&\quad - F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times 0 + \sum_{k=1}^K \gamma_k \beta_2^k t, K \right) dX, \\
\zeta(t) &= \int F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times 1 + (\sum_{k=1}^K \gamma_k \beta_2^k) \times t \right) \\
&\quad - F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + \beta_3 \times 0 + (\sum_{k=1}^K \gamma_k \beta_2^k) \times t \right) dX,
\end{aligned}$$

où pour une régression probit on a :

$$F_U(z, K) = \Phi \left( \frac{z}{\sqrt{\sigma_3^2 + \sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right),$$

et pour une régression logistique :

$$F_U(z, K) = \int_{\mathbb{R}} \Phi \left( \frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) \frac{e^{e_3}}{(1 + e^{e_3})^2} de_3.$$

### 2.2.3 Étude de simulation

L'analyse de médiation par blocs a pour avantage de considérer moins de paramètres d'intérêt que l'analyse de médiation multiple quand le nombre de médiateur augmente. En effet, l'analyse de médiation par blocs considère l'effet total, l'effet direct, l'effet médié joint et les effets médiés par les blocs de médiateurs quelque soit le nombre de médiateurs.

Comme elle considère moins de paramètres, l'analyse de médiation par blocs est plus facilement interprétable et plus intéressante en terme de temps de calcul. À titre d'illustration, nous comparons dans la Figure 2.2 le temps de calcul de l'analyse de médiation multiple et de l'analyse de médiation par blocs.

On considère un traitement binaire, des médiateurs continus et une variable de réponse continue. Les médiateurs sont simulés de façon à avoir des blocs de deux médiateurs corrélés de façon non-causale. On augmente le nombre de médiateurs et on fait une analyse de médiation multiple à l'aide de notre fonction `multimediate` et une analyse de médiation par blocs à l'aide de notre fonction `blockmediate` et on mesure le temps écoulé.

La Figure 2.2, illustre bien que lorsque que la structure du modèle est adéquat et que les médiateurs sont nombreux sans pour autant être supérieur au nombre d'observations, il est plus intéressant d'utiliser l'analyse de médiation par blocs.

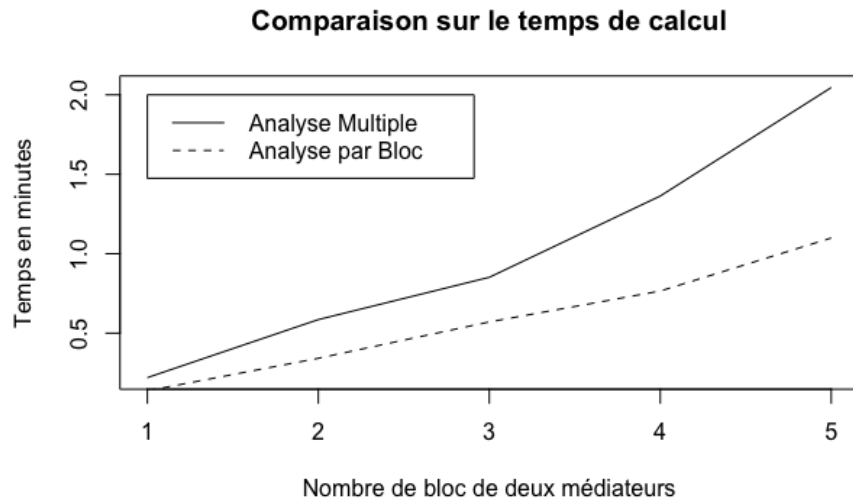


FIGURE 2.2 – Comparaison du temps d'estimation des paramètres pour l'analyse de médiation multiple et pour l'analyse de médiation par blocs. Le temps est donné en minutes. Les blocs de médiateurs sont constitués de deux médiateurs corrélés de façon non-causale.

## 2.3 Sélection de médiateur en grande dimension.

Dans les applications réelles, il est raisonnable de faire l'hypothèse que seule une petite proportion de variables médie l'effet de l'exposition sur la variable de réponse. De plus, d'un point de vue théorique le modèle de la variable de réponse  $Y$  en fonction des médiateurs n'est identifiable que lorsque le nombre de médiateurs est petit, c'est-à-dire inférieur au nombre d'observations. Étendre l'analyse de médiation multiple à un cadre de grande dimension nécessite donc l'identification des médiateurs actifs à travers un problème de sélection variable.

La sélection de médiateurs est un problème récent auquel plusieurs auteurs se sont intéressés.

Une première famille de méthodes repose sur la succession de trois étapes dont la première est l'application de la *sure independent screening* (SIS), décrite dans le travail de Fan and Lv (2008). Cette méthode permet de réduire le nombre de médiateur sur la base de leur corrélation avec la variable de sortie.

Zhang et al. (2016) proposent la méthode de sélection HIMA, où la deuxième étape permet de sélectionner un ensemble de médiateurs à l'aide de la pénalité concave minimax décrite dans le travail de Zhang (2010), avant d'être validé par un test multiple sur les coefficients de médiation. L'ensemble de la procédure est implémentée et disponible dans le package R `hima`.

Luo et al. (2020) proposent une procédure qui correspond à l'adaptation de Zhang et al. (2016) pour les modèles de survie. La différence réside dans la dernière étape où l'on décompose des effets de façon à obtenir l'effet médié par chaque médiateur, pour finalement tester leur significativité avec un test de Sobel multiple.

Gao et al. (2019) proposent la méthode HDMA. La première étape SIS est suivie d'une étape de de-biaisement des estimateurs et de dé-sparsification qui permet de prendre en compte les corrélations entre les médiateurs. Et enfin une étape de test multiple sur les coefficients de médiation est mise en oeuvre. L'ensemble de la procédure HDMA est implémentée dans le package R `hdma` disponible sur Gao (2019).

Une deuxième famille de méthode utilise une approche bayésienne :

Dans le cadre des données génotypiques les méthodes généralement utilisées ne peuvent pas distinguer la pléiotropie de la médiation causale et sont limitées par des hypothèses trop fortes sur les données. Pour surmonter ces limitations, Park et al. (2017) développent la médiation causale multivariée dans le déséquilibre de liaison étendue (CaMMEL), un nouveau cadre d'inférence bayésien pour modéliser conjointement plusieurs effets médiés et non médiés en s'appuyant uniquement sur des statistiques sommaires. Leur méthode est implémentée et disponible sur Park (2017).

Avec la croissance des données omiques, l'analyse conjointe des données de génomique au niveau moléculaire avec les données épidémiologiques grâce à l'analyse de médiation est



de plus en plus courante. Cependant, une telle analyse conjointe nécessite des méthodes qui peuvent simultanément prendre en compte des données avec un grand nombre de médiateurs et des données qui en ont peu. Song et al. (2018) développent une méthode d'inférence bayésienne utilisant des *shrinkage priors* continus pour étendre les techniques usuelles d'analyse de la médiation causale à un cadre de grande dimension. Ainsi qu'un test multiple pour les effets indirects naturels en utilisant une procédure de permutation.

D'autres méthodes existent sans s'inscrire dans une famille particulière : Boca et al. (2014), guidés par les méthodes développées pour tester les associations directes, proposent une approche par permutation pour tester plusieurs médiateurs et contrôler la *family wise error rate* (FWER). Leur méthode est implémentée et disponible sur Boca (2013).

Zhao and Luo (2016) propose une nouvelle pénalité appelée *Pathway LASSO*. Cette fonction de pénalité est une relaxation convexe de la fonction produit non convexe, et elle permet grâce à un critère d'optimisation calculable d'estimer et de sélectionner des médiateurs.

van Kesteren and Oberski (2019) développe une approche hybride qui utilise à la fois des composants de filtrage et de régularisation : le *coordinate-wise mediation filter* (CMF). Cette approche converge en cas de données de grande dimension et le filtrage conditionnel aux autres médiateurs sélectionnés permet de prendre en compte la corrélation entre médiateurs. L'ensemble de la procédure CMF est implémentée dans le package R `cmfilter`, disponible sur van Kesteren (2013).

Les approches d'estimation actuelles nécessitent de spécifier correctement un modèle pour les relations causales entre les médiateurs, ce qui peut être difficile en présence d'un grand nombre de médiateur continu. Loh et al. (2020) évite d'avoir à modéliser ces relations, en s'appuyant sur une définition des effets interventionnels suggérés par VanderWeele and Tchetgen Tchetgen (2017) pour la médiation longitudinale.

## Problème

Le problème que nous désirons aborder a lieu en présence d'un(des) traitement(s), une variable de réponse et d'un grand nombre de médiateurs potentiels qui, comme illustré sur la Figure 2.3, est constitué de trois types de variables. Premièrement des vrais médiateurs (illustrés par  $M_1, \dots, M_9$  dans la figure 2.3), deuxièmement des faux médiateurs (illustrés par  $M_{10}, \dots, M_{13}$  dans la figure 2.3), que nous appelons ainsi étant donné que la relation de causalité avec ces variables est *soit* avec le(s) traitement(s) *soit* avec la variable de réponse, et troisièmement des variables n'ayant aucun lien ni avec le(s) traitement ni avec la variable de réponse (illustrés par  $M_{14}, \dots, M_K$  dans la figure 2.3).

Nous nous plaçons sous l'hypothèse que la distribution des médiateurs et de la variable de réponse sont normales, que nous avons identifiés  $P$  variables d'expositions et d'un grand

nombre  $K$  de médiateurs potentiels. On a donc le modèle de médiation suivant :

$$\begin{aligned} M_k &= \mu_k + \sum_{p=1}^P \alpha_{p,k} T_p + \epsilon_k \\ Y &= \mu_0 + \sum_{p=1}^P \alpha_p T_p + \sum_{k=1}^K \beta_k M_k + \epsilon_0 \end{aligned} \quad (2.3.1)$$

où  $\alpha_p$  l'effet direct de  $T_p$  sur  $Y$  après ajustement sur les médiateurs,  $\boldsymbol{\alpha}_k = (\alpha_{1,k}, \dots, \alpha_{P,k})^\Gamma$  représente l'effet de chaque traitement sur le médiateur  $M_k$  et  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)^\Gamma$  l'effet des médiateurs sur la variable de réponse.

En présence de beaucoup de médiateurs potentiels, il est plus crédible de supposer que seul un petit nombre  $K^*$  joue réellement un rôle de médiation. Il est donc nécessaire de réaliser une sélection de médiateurs, de façon à isoler les médiateurs qui médient un effet significatif.

Nous développons une méthode permettant de sélectionner les médiateurs communs (illustrés par  $M_1, M_2, M_3$  dans la figure 2.3) à plusieurs traitements identifiés. L'idée est de favoriser la sélection de mêmes médiateurs potentiels pour les différentes expositions : si un  $\beta_k$  ou un  $\alpha_{p,k}$  sont non nuls, on a envie de moins pénaliser les autres éléments du vecteur  $\boldsymbol{\alpha}_k$ . Pour cela nous cherchons à résoudre le problème suivant :

$$\hat{\theta} = \underset{\theta \in \mathbb{R}^{P \times K+1}}{\operatorname{argmin}} \sum_k \left\| M_k - \sum_p \alpha_{p,k} T_p \right\|^2 + \left\| Y - M\boldsymbol{\beta} - \sum_p \alpha_p T_p \right\|^2 + \lambda \sum_k \operatorname{pen}(\boldsymbol{\alpha}_k, \beta_k), \quad (2.3.2)$$

où  $\theta$  est le vecteur des coefficients  $(\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_K, \beta_1, \dots, \beta_K, \mu_0)^\Gamma$  et  $\operatorname{pen}(\boldsymbol{\alpha}_k, \beta_k)$  étant une pénalité à définir. Ainsi un médiateur  $M_k$  est sélectionné lorsque les coefficients du vecteur  $\boldsymbol{\alpha}_k$  et le coefficient  $\beta_k$  sont non nuls.

### 2.3.1 MAHI en présence d'un seul traitement

Pour présenter notre méthode MAHI simplement et pour nous comparer aux autres méthodes existantes, nous nous plaçons dans le cas d'une exposition unique, c'est à dire que  $P = 1$ . On note  $K$  le nombre de médiateurs et  $I$  le nombre d'observations.

Notre méthode de sélection se décompose en une première sélection un peu grossière et un affinement par la suite. C'est aussi le cas dans la plupart des méthodes existantes, mais notre méthode de sélection se base sur les effets médiés dès la première étape.

En effet la première étape des méthodes connues commencent par déterminer les liens entre la variable de sortie et les médiateurs. Pour déterminer ces liens, l'intuition première serait de calculer l'estimateur des moindres carrés  $\hat{\boldsymbol{\beta}}^{MC}$  de la régression linéaire des médiateurs sur la variable de réponse, défini par :

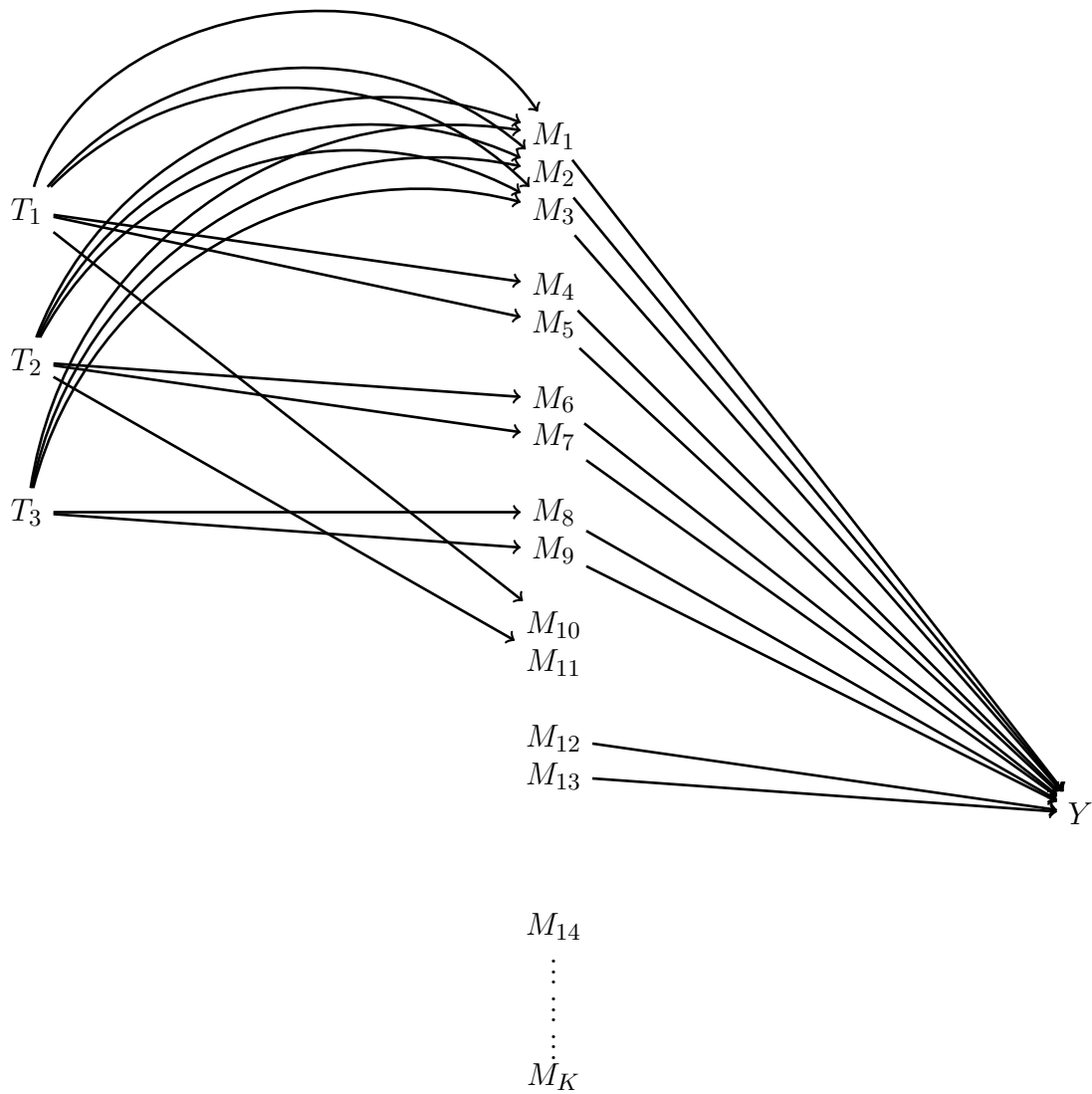


FIGURE 2.3 – Modèle de médiation en présence de trois expositions, d'un grand nombre de médiateurs potentiels et d'une variable de réponse. Les effets directs des expositions ( $T_1, T_2, T_3$ ) sont présents dans le modèle mais non explicités sur la figure pour sa lisibilité.

$$\widehat{\boldsymbol{\beta}}^{MC} = \operatorname{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^K} \left\{ \frac{1}{2} \|Y - M\boldsymbol{\beta} - \alpha T\|_2^2 \right\}$$

Mais dans le cas où  $K > I$ , la matrice  $M^\Gamma M$  n'est pas de plein rang et donc non inversible. La solution des moindres carrés est alors mal définie. On ne peut donc pas l'utiliser pour déterminer les vrais médiateurs.

La méthode LASSO permet de pallier à ce problème en contraignant la norme  $l_1$  des coefficients du vecteur  $\boldsymbol{\beta}$ , ce qui force un certain nombre de coefficients de  $\boldsymbol{\beta}$  à être nuls. L'estimateur LASSO  $\widehat{\boldsymbol{\beta}}_\lambda^L$  est défini par :

$$\widehat{\boldsymbol{\beta}}_\lambda^L = \operatorname{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^K} \left\{ \frac{1}{2} \|Y - M\boldsymbol{\beta} - \alpha T\|_2^2 + \lambda \|\boldsymbol{\beta}\|_1 \right\}, \quad (2.3.3)$$

avec  $\lambda \geq 0$  le paramètre de régularisation. Plus  $\lambda$  sera grand et plus le problème sera contraint et donc moins de coefficients de  $\widehat{\boldsymbol{\beta}}_\lambda^L$  seront non nuls. La méthode LASSO permet de sélectionner les médiateurs un à un, en fonction de leur lien avec  $Y$  sans prendre en compte le lien avec l'exposition.

Pour définir une méthode basée sur les effets médiés, la méthode LASSO n'est donc pas adaptée. Nous nous sommes alors intéressés à la méthode Group-LASSO (Grimonprez et al. (2015)) qui, à l'aide d'une partition donnée a priori, permet de sélectionner un certain nombre de groupes en fonction du paramètre de régularisation. En particulier dans notre cas, les groupes doivent correspondre aux paires de coefficients  $(\boldsymbol{\alpha}_k, \beta_k)$ , chacune correspondant à un médiateur potentiel.

Définissons  $\mathcal{G}$  un ensemble de groupes formant une partition en  $|\mathcal{G}|$  groupes des  $K$  variables et  $\boldsymbol{\beta}_g \in \mathbb{R}^{|g|}$  avec  $g \in \mathcal{G}$ , le vecteur  $\boldsymbol{\beta}$  restreint aux variables du groupe  $g$ . L'estimateur du Group-LASSO  $\widehat{\boldsymbol{\beta}}_\lambda^{GL}$  est défini par :

$$\widehat{\boldsymbol{\beta}}_\lambda^{GL} = \operatorname{argmin}_{\boldsymbol{\beta} \in \mathbb{R}^K} \left\{ \frac{1}{2} \|Y - M\boldsymbol{\beta} - \alpha T\|_2^2 + \lambda \sum_{g \in \mathcal{G}} \sqrt{|g|} \|\boldsymbol{\beta}_g\|_2 \right\}$$

Ainsi on retient ou rejette des groupes de variables déjà définis. Le poids  $|g|$  permet de pénaliser plus fortement les groupes de grande taille qui auront tendance à être favorisés pour la sélection par rapport à des groupes de petite taille.

Rappelons que dans notre cas nous ne cherchons pas à former des groupes de médiateurs corrélés qui seront gardés ou non mais nous cherchons plutôt à choisir des médiateurs par rapport à leur effet médié. Nous devons donc construire, pour  $K$  médiateurs,  $K$  groupes contenant deux variables, où le  $k$ -ième groupe est composé du  $k$ -ième du médiateur et du traitement. De plus la significativité des médiateurs ne se mesurent pas seulement à la significativité des  $(\beta_1, \dots, \beta_K)^\Gamma$  mais à la significativité des effets médiés, qui dans le cas présenté dans (2.3.1) et pour  $P = 1$  correspond à la significativité des produits  $(\alpha_{1,1}\beta_1, \dots, \alpha_{1,K}\beta_K)^\Gamma$ .

Les questions suivantes se posent alors : *Comment construire  $K$  groupes différents avec une variable récurrente (l'exposition) dans chacun des groupes ? Comment définir une pénalité qui permette la prise en compte simultanée pour un médiateur  $M_k$  des coefficients  $\alpha_{1,k}$  et  $\beta_k$  ?*

**Méthode Group-LASSO** Pour répondre à ces deux questions nous devons réorganiser les données de la façon suivante :

—  $\mathbf{z}$  le vecteur de toutes les observations, de taille  $(K+1)I$  :  $\mathbf{z} = \begin{pmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_K \\ \mathbf{Y} \end{pmatrix}$ ,

avec  $\mathbf{M}_k = \begin{pmatrix} M_{k,1} \\ \vdots \\ M_{k,I} \end{pmatrix}$  le vecteur des observations du  $k$ -ième médiateur et  $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_I \end{pmatrix}$

le vecteur des observations de la variable de réponse.

—  $X$  la matrice des variables explicatives de taille  $(K+1)I \times 2K+1$  définie par :

$$X = \begin{pmatrix} \mathbf{T} & \mathbf{0}_I & \dots & \mathbf{0}_I & \mathbf{0}_{I \times K} & \mathbf{0}_I \\ \mathbf{0}_I & \mathbf{T} & \dots & \mathbf{0}_I & & \\ \vdots & & & \vdots & \vdots & \vdots \\ \mathbf{0}_I & \mathbf{0}_I & \dots & \mathbf{T} & \mathbf{0}_{I \times K} & \mathbf{0}_I \\ \mathbf{0}_I & \mathbf{0}_I & \dots & \mathbf{0}_I & \mathbf{M} & \mathbf{T} \end{pmatrix},$$

avec  $\mathbf{T} = \begin{pmatrix} T_1 \\ \vdots \\ T_I \end{pmatrix}$  le vecteur des observations du traitement,  $\mathbf{0}_I$  le vecteur nul de taille  $I$ ,  $\mathbf{0}_{I \times K}$  la matrice nulle de taille  $I \times K$  et  $\mathbf{M}$  la matrice de terme général  $M_{k,i}$ .

Avec les données réorganisées on remarque que  $\mathbf{z} - X\theta = \begin{pmatrix} \mathbf{M}_1 - \alpha_1 \mathbf{T} \\ \vdots \\ \mathbf{M}_K - \alpha_K \mathbf{T} \\ \mathbf{Y} - \mathbf{M}\beta - \mu \mathbf{T} \end{pmatrix}$ . On a donc

l'égalité suivante :

$$\sum_k \|\mathbf{M}_k - \alpha_k \mathbf{T}\|^2 + \|\mathbf{Y} - \mathbf{M}\beta - \mu \mathbf{T}\|^2 = \|\mathbf{z} - X\theta\|^2$$

On peut maintenant appliquer la méthode de Group-LASSO, on obtient :

$$\hat{\theta}_\lambda^{GL} = \underset{\theta \in \mathbb{R}^{2K+1}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{z} - X\theta\|^2 + \lambda \sum_{1 \leq k \leq K} \sqrt{2} \|\theta_k\| \right\}, \quad (2.3.4)$$

avec  $\theta_k = (\alpha_k, \beta_k)$ .

Résoudre l'équation (2.3.4) revient à résoudre l'équation (2.3.2). Nous avons donc une méthode permettant de sélectionner des médiateurs lorsque le nombre de médiateurs est supérieur au nombre d'observations et cela en considérant simultanément les coefficients  $\alpha_k$  et  $\beta_k$ .

Afin de diminuer le taux d'erreur de sélection des médiateurs par la méthode Group-LASSO on utilise la *stability selection* (Meinshausen and Bühlmann (2010)), qui est décrite dans la première étape de notre algorithme de sélection de médiateurs, MAHI, décrit ci-dessous.

**Algorithme.** *L'algorithme MAHI se déroule en deux étapes :*

1. *Stability selection.* À partir des données on génère un nombre  $n_{boot}$  de données bootstrap. Pour chaque jeu de données bootstrap on utilise la méthode Group-LASSO telle que définie dans l'équation (2.3.4). On compte alors  $N_k$ , le nombre de fois où le médiateur  $M_k$  a été sélectionné sur les  $n_{boot}$  fois et on classe les médiateurs par rapport à  $N_k$ . On choisit alors les  $K_{max}$  premiers médiateurs, où  $K_{max}$  est un nombre choisi arbitrairement.
2. *Test multiple.* On utilise notre fonction  $R$  multimediate pour réaliser une analyse de médiation multiple avec les  $K_{max}$  médiateurs. La fonction multimediate nous permet d'avoir une  $p$ -valeur pour chaque effet médié. La méthode utilisée pour calculer cette  $p$ -valeur est présentée dans le Chapitre 3. On a donc un vecteur de  $p$ -valeur de longueur  $K_{max}$ , noté  $P_{vect}$ , auquel on applique la correction de Bonferroni :

$$P_{corr,k} = \min(P_{km} \times K_{max}, 1),$$

où  $P_{km}$  est la  $p$ -valeur du  $k$ -ième élément de  $P_{vect}$ .

Si  $P_{corr,k} < 0,05$ , nous pouvons conclure qu'il existe un effet de médiation significatif passant par  $M_k$ .

L'ensemble de la procédure est implémenté dans le package R multimediate, disponible sur <https://github.com/AllanJe/multimediate/>. Notons que la correction utilisée et la  $p$ -valeur seuil peuvent être spécifiés par l'utilisateur permettant ainsi de choisir des corrections moins conservatives que celle de Bonferroni.

Cette méthode se généralise aisément à plusieurs expositions en définissant les groupes avec  $\theta_k = ((\alpha_{1,k}, \dots, \alpha_{P,k})^\Gamma, \beta_k)$ . Ces résultats sont présentés dans l'appendice B.2.

### 2.3.2 Quelques méthodes existantes

Nous présentons maintenant les méthodes CMF, HIMA et HDMA, trois méthodes disponibles dans la littérature récente de l'analyse de médiation en grande dimension pour traiter le modèle (2.3.1). Ces méthodes ne s'appliquent que pour une exposition à la fois.

## Méthode CMF

van Kesteren and Oberski (2019) propose une approche qui permet de faire de l'analyse de médiation en grande dimension, appelée la méthode du *coordinate-wise mediation filter* (CMF, filtre de médiation par coordonnées). L'implémentation CMF se compose de deux composants : un algorithme interne qui effectue la sélection des médiateurs par descente de coordonnées à l'aide d'une fonction de décision  $\mathcal{D}$ , et un algorithme externe qui effectue plusieurs lancement aléatoires de l'algorithme interne et l'agrégation ultérieure des multiples sorties.

**Algorithme Interne** L'objectif de l'algorithme interne est le suivant. Dans un premier temps, un vecteur de décision  $P$  de longueur  $K$  (c'est-à-dire le nombre total de médiateurs) contenant la sélection actuelle de médiateurs est initialisé au hasard avec des valeurs 0 et 1. Puis à chaque étape de l'algorithme interne,  $\sqrt{K}$  médiateurs sont sélectionnés au hasard parmi tous les médiateurs et le vecteur de décision est mis à jour comme suit. Pour chaque médiateur  $M_k$  parmi les  $\sqrt{K}$  sélectionnés, une matrice de données  $M_*$ , qui contient tous les médiateurs actuellement sélectionnés (ce sont des médiateurs de valeur correspondante 1 dans le vecteur de décision  $P$ ) à l'exclusion de la variable  $M_k$  considérée, est créée. Ensuite, les résidus de l'exposition  $T$  et de la variable de réponse  $Y$  par rapport à  $M_*$  sont calculés. Enfin la décision binaire pour savoir si  $M_k$  est sélectionné par rapport à la matrice  $M_*$ , est prise en utilisant le test de Sobel. Notons que c'est à cette étape que  $\alpha_k$  et  $\beta_k$  sont estimés.

**Algorithme Externe** La valeur du vecteur de décision résultant de l'algorithme interne dépend dans une certaine mesure des valeurs initiales, en raison de la nature discrète de ses coordonnées. Par conséquent, l'algorithme est intégré dans une boucle externe qui effectue un nombre prédéfini de boucles. Lorsque toutes les boucles sont terminées, un taux de sélection  $S_k$  est calculé pour chaque médiateur  $M_k$ , qui représente la proportion de fois où le médiateur potentiel  $M_k$  a été sélectionné par l'algorithme interne. Ces proportions conduisent à un classement des médiateurs potentiels. Ce classement est ensuite à nouveau dichotomisé en utilisant une valeur seuil qui doit être spécifiée par l'utilisateur. Tous les médiateurs dont le taux de sélection est strictement supérieur à la valeur seuil sont classés comme vrais médiateurs. Il n'y a pas de procédure proposée pour le choix du seuil dans l'article qui présente la méthode CMF (van Kesteren and Oberski (2019)), celui-ci est fixé arbitrairement par l'utilisateur.

Cette méthode est présentée de façon plus détaillée dans l'appendice B.4. L'ensemble de la procédure CMF est implémentée dans le package R `cmfilter`, disponible sur van Kesteren (2013).

### Méthodes HIMA et HDMA

Zhang et al. (2016) propose l'approche HIMA (High dimensional Mediation Analysis), qui repose sur des régressions pénalisées et utilise une fonction de pénalité de type LASSO appelée pénalité concave minimax (MCP) (Zhang (2010)). Cette méthode est présentée de façon plus détaillée dans l'appendice B.3.

**Algorithme.** *L'algorithme HIMA se déroule en trois étapes :*

1. Réduire le nombre initial de médiateurs potentiels d'un très grand à un nombre raisonnable (c'est-à-dire plus petit que la taille de l'échantillon) en utilisant le *sure independence screening (SIS)* décrite dans Fan and Lv (2008). Les données sont normalisées par médiateur et seuls les  $\frac{n}{\log(n)}$  médiateurs les plus corrélés à  $Y$  sont sélectionnés.
2. Sélectionner les médiateurs. Les estimations  $\widehat{\beta}_k$ ,  $k \in \mathcal{I}$  sont calculées en minimisant le critère de pénalisation MCP (Zhang (2010)),

$$\mathcal{Q}^{ols} = \sum_{i=1}^n (Y_i - \mu_0 - \alpha_0 T_i - \sum_{k \in \mathcal{I}} \beta_k M_{ik})^2 + \sum_{k \in \mathcal{I}} p_{\lambda, \delta}(\beta_k), \quad (2.3.5)$$

où  $p_{\lambda, \delta}(\cdot)$  est la fonction MCP défini par :

$$p_{\lambda, \delta}(\beta_k) = \lambda \left[ |\beta_k| - \frac{|\beta_k|^2}{2\delta\lambda} \right] \mathbb{1}_{\{0 \leq |\beta_k| < \delta\lambda\}} + \frac{\lambda^2 \delta}{2} \mathbb{1}_{\{|\beta_k| \geq \delta\lambda\}},$$

où  $\lambda > 0$  est le paramètre de régularisation, et  $\delta > 0$  détermine la concavité du MCP.

3. Effectuer des test multiples pour les effets médiés par les médiateurs obtenus par la procédure de pénalisation MCP. La correction de Bonferroni est appliquée pour tester la nullité des coefficients  $(\alpha_k, \beta_k)$ , c'est-à-dire que l'hypothèse nulle d'absence d'effet de médiation avec  $M_k$  n'est rejetée que si  $\alpha_k$  et  $\beta_k$  sont tous deux significatifs.

L'ensemble de la procédure est implémenté dans le package R `hima`.

La méthode HDMA est très semblable mais diffère au niveau de la deuxième étape. Cette deuxième étape consiste à obtenir un estimateur débiaisé  $\widehat{\beta}_k$  par :

$$\widehat{\beta}_k = \frac{Z_k^\Gamma Y}{Z_k^\Gamma M_k} - \sum_{l \neq k} \frac{Z_k^\Gamma M_l}{Z_k^\Gamma M_k} \widehat{\beta}_{\lambda k}^L$$

où  $\widehat{\beta}_{\lambda k}^L$  est le coefficient du  $k$ -ième médiateur  $M_k$  de la régression LASSO (2.3.3). Par ailleurs,  $Z_k = M_k - W_k \widehat{\gamma}^L$ , où  $\widehat{\gamma}^L$  est le coefficient de la régression LASSO expliqué par l'ensemble des médiateurs retenus par la première étape excepté le  $k$ -ième, noté  $W_k$ , sur  $M_k$ .

L'ensemble de la procédure est implémentée et disponible sur Gao (2019).



### 2.3.3 Comparaison des méthodes

Pour comparer les méthodes MAHI, CMF, HIMA et HDMA nous utilisons le modèle (2.3.6) pour simuler 100 jeux de données.

$$\begin{aligned} M_{ik} &= \mu_k + \alpha_k T_i + \epsilon_{ik} \\ Y_i &= 20 + 50T_i + \sum_k \beta_k M_{ik} + \epsilon_{i0} \end{aligned} \quad (2.3.6)$$

où  $\mu_k$  est tiré aléatoirement pour chaque médiateurs,  $\epsilon_k \sim \mathcal{N}(0,1)$  pour  $k \in (0, \dots, 500)$  et  $T \sim \mathcal{B}(0.3)$ .

Chaque jeu de données comporte n=100 observations et K=500 médiateurs dont 30 vrais médiateurs, 20 faux médiateurs et 450 liés ni au traitement ni à la variable de réponse. Nous avons intentionnellement défini très peu de médiateurs comme de vrais médiateurs afin d'être cohérent avec des applications réelles dans le cadre génomique, où le sous-ensemble de vrais médiateurs que nous recherchons est très petit par rapport au nombre total de médiateurs potentiels (c'est-à-dire tous les médiateurs qui sont considérés).

<b>Indices</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>	<b>10</b>
$\alpha_{\dots}$	-95	90	95	95	-100	95	-95	85	-95	-100
$\beta_{\dots}$	185	-195	190	185	-190	185	195	-190	100	185
<b>Indices</b>	<b>11</b>	<b>12</b>	<b>13</b>	<b>14</b>	<b>15</b>	<b>16</b>	<b>17</b>	<b>18</b>	<b>19</b>	<b>20</b>
$\alpha_{\dots}$	-2.75	3.25	-3.50	2.50	-3.75	3.00	-3.25	2.75	3.00	3.50
$\beta_{\dots}$	3.25	-2.50	3.75	-3.00	3.50	-2.75	3.75	-3.25	3.25	-2.75
<b>Indices</b>	<b>21</b>	<b>22</b>	<b>23</b>	<b>24</b>	<b>25</b>	<b>26</b>	<b>27</b>	<b>28</b>	<b>29</b>	<b>30</b>
$\alpha_{\dots}$	-0.875	0.625	-0.375	-0.25	0.50	-0.875	0.125	-1.125	0.375	-0.25
$\beta_{\dots}$	0.375	-0.625	0.625	-1.125	0.75	0.25	-0.50	0.375	-0.625	0.75
<b>Indices</b>	<b>31</b>	<b>32</b>	<b>33</b>	<b>34</b>	<b>35</b>	<b>36</b>	<b>37</b>	<b>38</b>	<b>39</b>	<b>40</b>
$\alpha_{\dots}$	25	25	25	25	25	25	25	25	25	25
$\beta_{\dots}$	0	0	0	0	0	0	0	0	0	0
<b>Indices</b>	<b>41</b>	<b>42</b>	<b>43</b>	<b>44</b>	<b>45</b>	<b>46</b>	<b>47</b>	<b>48</b>	<b>49</b>	<b>50</b>
$\alpha_{\dots}$	0	0	0	0	0	0	0	0	0	0
$\beta_{\dots}$	45	45	45	45	45	45	45	45	45	45

TABLE 2.1 – 50 premières valeurs des paramètres  $\alpha_k$  et  $\beta_k$ , les 450 autres étant à 0.

#### Métriques de comparaison

Nous présentons les métriques que nous avons utilisé pour comparer les méthodes de sélection de médiateurs. On note  $VP$ ,  $VN$ ,  $FP$  et  $FN$  respectivement le nombre de vrais positifs, le nombre de vrais négatifs, le nombre de faux positifs et le nombre de faux négatifs.

		Prédit	
		Vrai	Faux
Observé	Vrai	VP	FN
	Faux	FP	VN

TABLE 2.2 – Exemple de table de confusion

- Le rappel ou taux de vrais positifs  $TVP$  (puissance) est défini comme :

$$TVP = \frac{VP}{VP + FN}$$

Le rappel est le pourcentage de médiateurs qui ont été classés comme vrais parmi ceux qui sont effectivement de vrais médiateurs. Le rappel permet de mesurer à quel point le test est performant lorsqu'il est utilisé sur des vrais médiateurs.

- Le taux de faux négatif  $TFN$  est défini comme :

$$TFN = \frac{FN}{FN + VP} = 1 - TVP$$

Le taux de faux négatif est le pourcentage de médiateurs qui ont été classés comme faux parmi ceux qui sont en fait de vrais médiateurs.

- La spécificité ou le taux de vrais négatifs  $TVN$  est défini comme :

$$TVN = \frac{VN}{FP + VN}$$

La spécificité est la proportion de médiateurs négatifs effectivement bien détectés par le test. Autrement dit, la spécificité permet de mesurer à quel point le test est performant lorsqu'il est utilisé sur des médiateurs négatifs.

- Le taux de faux positif  $TFP$  (erreur de type I) est

$$TFP = \frac{FP}{VN + FP} = 1 - TVN$$

Le taux de faux positifs est le pourcentage de médiateurs qui ont été classés comme vrais parmi ceux qui ne le sont pas.

- La précision ou la valeur prédictive positive  $VPP$  est définie comme

$$VPP = \frac{VP}{VP + FP}$$

La précision représente la probabilité qu'un médiateur qui a été classé comme vrai soit un vrai médiateur.

- Le taux d'erreur  $TE$  se définit comme :

$$TE = \frac{FN + FP}{K}$$

Le taux d'erreur est la proportion de médiateurs mal classés.

### Paramétrage des méthodes

Pour notre méthode MAHI nous avons choisi  $n_{boot} = 30$  bootstraps. À l'aide de la fonction R `lambdaseq`, pour un jeu de donnée donné, on génère un vecteur de valeurs possibles pour  $\lambda$ , allant de la valeur où tous les médiateurs sont rejetés à la valeur où tous les médiateurs sont sélectionnés. Nous avons choisi  $\lambda = 0.008522981$  aléatoirement dans ce vecteur. Dans un premier temps nous sélectionnons les  $\frac{100}{\log(100)} = 21$  premiers médiateurs, car cette valeur est proposée par la méthode SIS. Dans un deuxième temps on choisit les 43 premiers médiateurs. Cette valeur 43 est arbitraire et permet de voir si les résultats des comparaisons sont dépendants du choix de  $K_{max}$ . Pour la deuxième étape la correction appliquée aux p-valeur est celle de Bonferroni avec pour seuil de p-valeur 0.05

Pour la méthode CMF la fonction de décision est le test de Sobel. L'algorithme externe est lancé 1000 fois. Une fois que le taux de sélection pour chaque médiateur est calculé, un médiateur est choisit si son taux de sélection est supérieur à 0.079, valeur recommandée par van Kesteren and Oberski (2019).

La première étape des méthodes HIMA et HDMA est celle de la *sure independent screening*. Dans un premier temps on choisit les  $\frac{100}{\log(100)} = 21$  premiers médiateurs qui est la valeur proposée par la méthode SIS. Dans un deuxième temps on choisit les 43 premiers médiateurs.

### Résultats

Nous avons évalué les métriques énumérées précédemment pour chaque méthode sur les 100 jeux de données.

Les méthodes MAHI et HIMA/HDMA ont une première étape de sélection (respectivement MGLASSO et SIS) pour ramener le nombre de médiateurs à un nombre inférieur au nombre d'observations. Pour comparer la méthode MGLASSO et la méthode SIS, dans un premier temps, on montre dans les figures 2.4 les résultats obtenus en choisissant les 21 premiers médiateurs obtenus par ces méthodes, puis dans la 2.5 les résultats obtenus en choisissant les 43 premiers médiateurs.

Lors de la première étape, que le nombre de médiateurs retenus soit inférieur ou qu'il soit supérieur au nombre de vrai médiateurs, les Figures 2.4 et 2.5 montrent qu'en moyenne notre méthode a de meilleurs résultats pour chacun des critères. En augmentant le nombre de médiateurs retenus, la précision diminue pour les deux méthodes car le nombre de

faux médiateurs retenus augmentent. Le rappel augmente pour les deux méthodes car on sélectionne plus de vrais médiateurs.

On compare ensuite dans les Figures 2.6 et 2.7 les méthodes complètes. On voit que pour notre méthode la précision est proche de 1. Cela montre que parmi les médiateurs sélectionnés, pratiquement la totalité sont des vrais pour notre méthode. Le rappel de notre méthode est à 34%. Ainsi 34% des vrais médiateurs sont en moyenne sélectionnés et 66% d'entre eux ne sont donc pas sélectionnés. Ce taux est éloigné de 1 mais est significativement meilleur que les autres méthodes.

La méthode CMF a une précision de plus de 80%; les médiateurs sélectionnés sont principalement des vrais médiateurs mais beaucoup de vrais médiateur ne sont pas sélectionnés. En effet la méthode CMF a un rappel d'environ 17% (le rappel le plus petit parmi les quatre méthodes).

Les méthodes HIMA et HDMA sélectionnent toutes deux 30% des vrais médiateurs mais la précision de HDMA est plus élevée, 50% pour HDMA contre 40% pour la méthode HIMA. La méthode HIMA sélectionne donc plus de variables non médiatrices.

La proportion de vrai médiateur étant petite, il est normal d'observer que la spécificité est proche de 1 et que le taux d'erreur est proche de 0 pour l'ensemble des méthodes.

Notre méthode a de meilleurs résultats pour l'ensemble des critères. On observe aussi que la méthode du Chapitre 1 élimine les faux positifs gardés par MGLASSO. Au vu des résultats entre la première étape et la deuxième étape de notre méthode on peut choisir un nombre relativement grand de médiateurs retenus lors de la première étape, sachant que la deuxième étape identifie et élimine clairement les faux médiateurs.

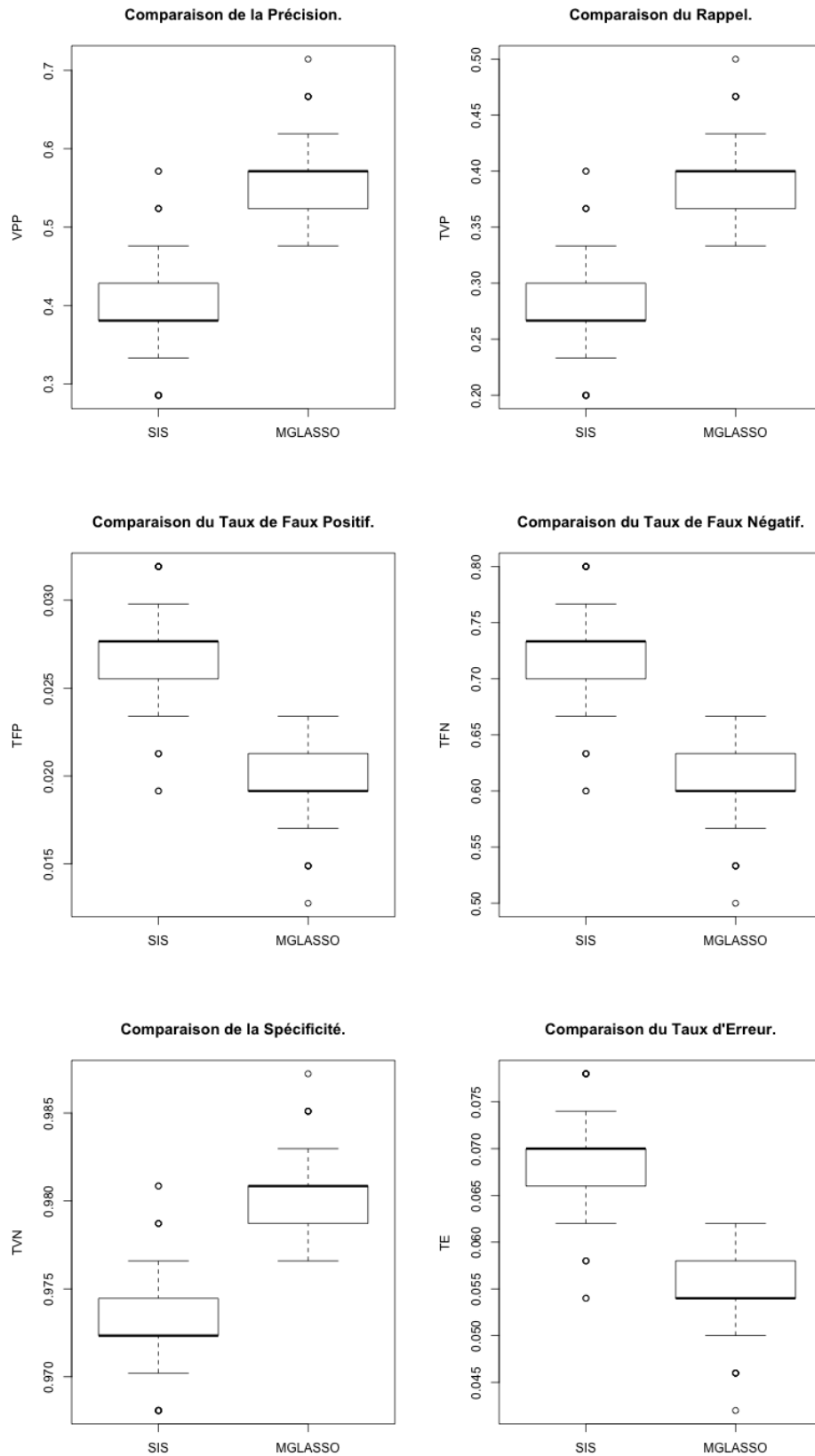


FIGURE 2.4 – Comparaison des boxplots des méthodes (SIS, MGLASSO) pour la précision, du rappel, du taux de faux positifs, du taux d'erreurs, du taux de faux négatifs et de la spécificité pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. Suite au rang obtenus par les méthodes, on choisit les 21 premiers médiateurs.

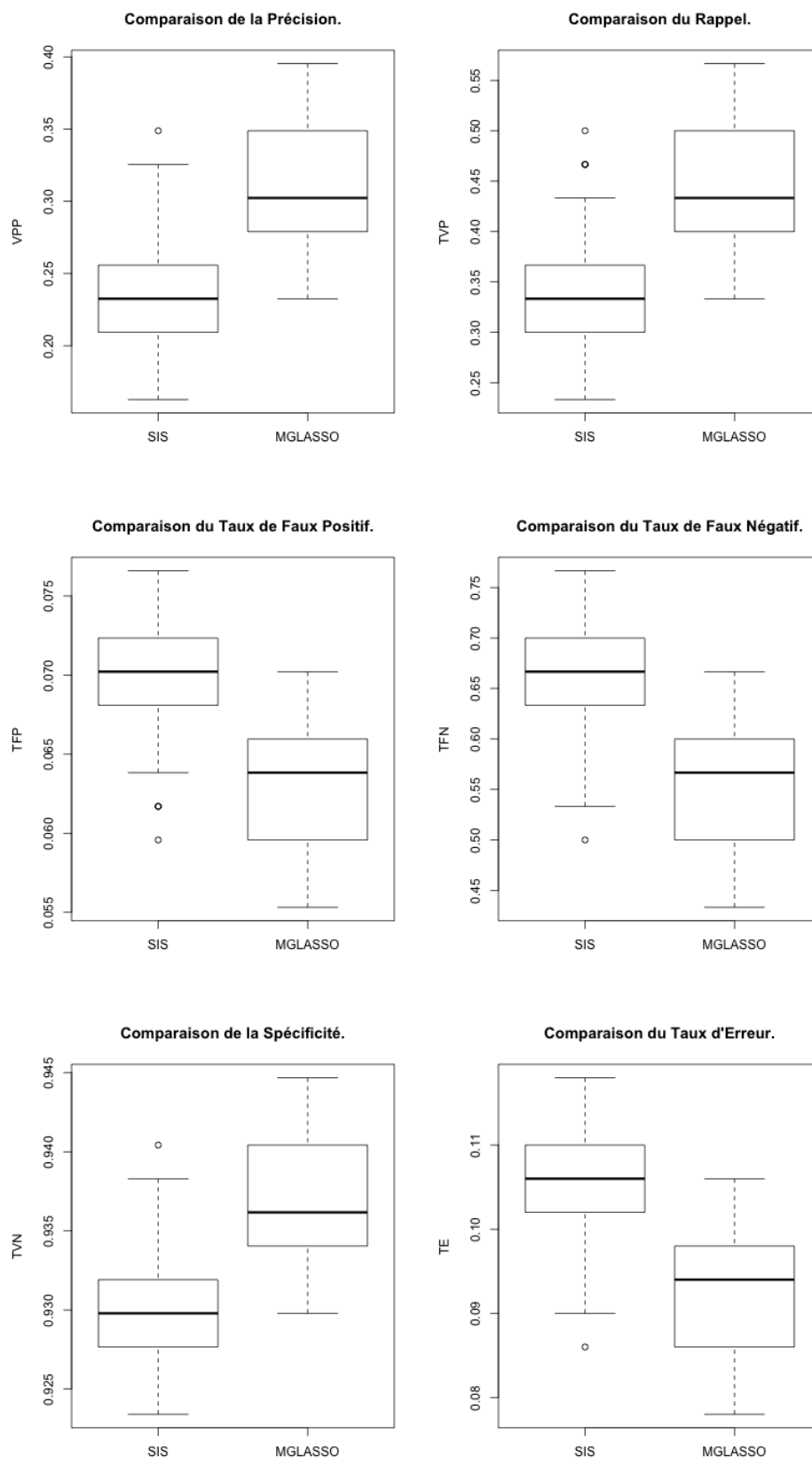


FIGURE 2.5 – Comparaison des boxplots des méthodes (SIS, MGLASSO) pour la précision, du rappel, du taux de faux positifs, du taux d'erreurs, du taux de faux négatifs et de la spécificité pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. Suite au rang obtenu par les méthodes, on choisit les 43 premiers médiateurs.

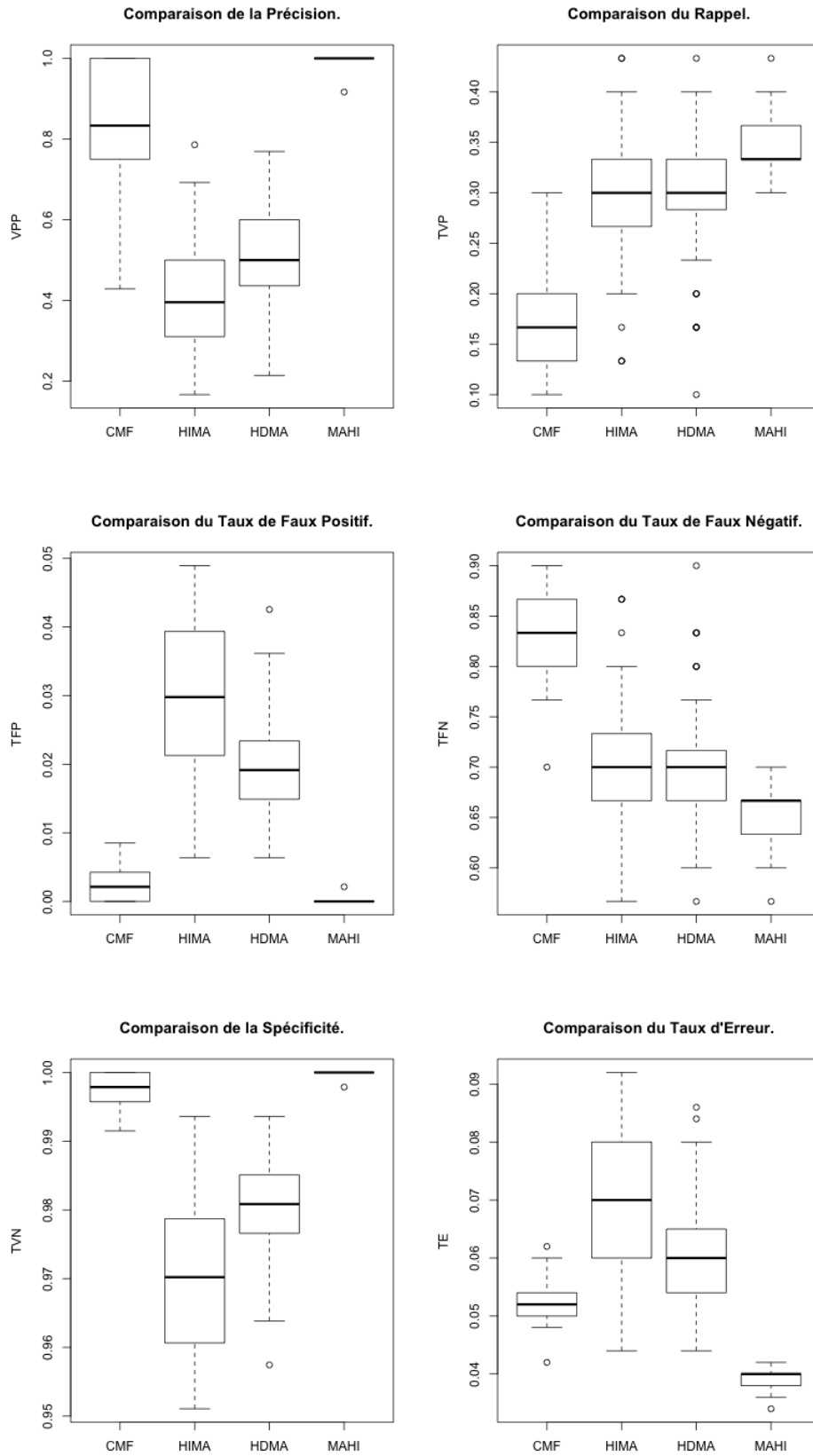


FIGURE 2.6 – Comparaison des boxplots des méthodes pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. À la première étape des méthodes HIMA, HDMA et MAHI on retient 21 médiateurs.

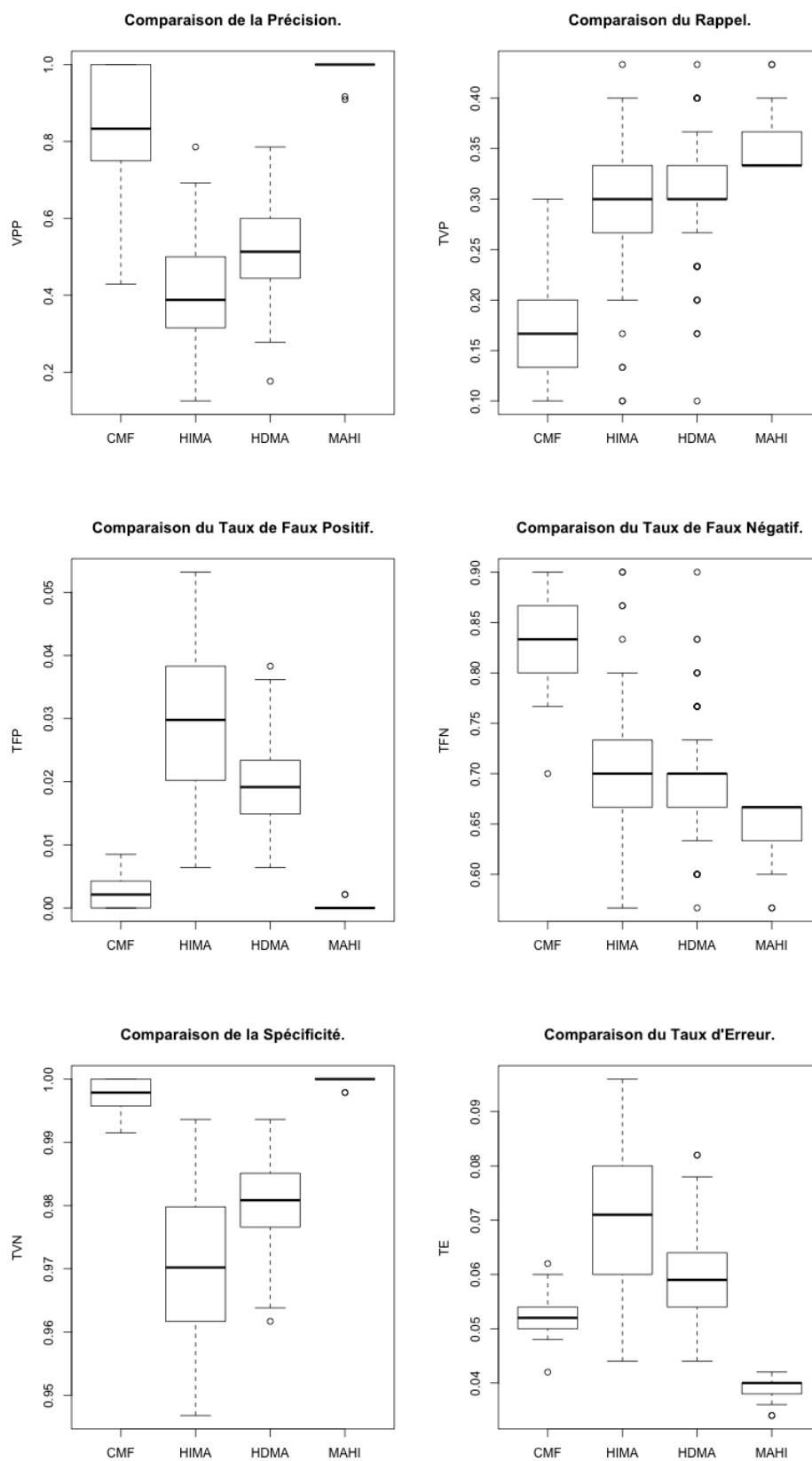


FIGURE 2.7 – Comparaison des boxplots des méthodes pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. À la première étape des méthodes HIMA, HDMA et MAHI on retient 43 médiateurs





# Chapitre 3

## Package R : multimediate

### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>73</b>
<b>3.2</b>	<b>Algorithmes d'estimation</b>	<b>74</b>
3.2.1	Algorithme d'Imai	74
3.2.2	Estimation de la corrélation entre médiateurs	75
3.2.3	Algorithme pour médiateurs multiples non causalement corrélés	78
<b>3.3</b>	<b>Présentation du package</b>	<b>80</b>
3.3.1	Analyse de médiation en présence de médiateur corrélés de façon non-causale	80
3.3.2	Analyse de médiation par bloc	83
3.3.3	Sélection de médiateur en grande dimension	88

---

### 3.1 Introduction

Imai et al. (2010a) présentent des algorithmes d'estimation des effets d'intérêts qui sont mis en œuvre dans le package R `mediation` (Tingley et al. (2014)). Pour permettre l'analyse de médiation en présence de plusieurs médiateurs liés de façon non-causale, présentée au Chapitre 1, nous avons adapté l'un de ces algorithmes. Nous avons implémenté cet algorithme dans un package R `multimediate` disponible sur GitHub, <https://github.com/AllanJe/multimediate>. Dans ce Chapitre nous présentons dans un premier temps l'algorithme d'Imai et al. (2010a). Dans un deuxième temps nous proposons une manière de prendre en compte la corrélation entre des médiateurs linéaires, binaires et catégoriels ordonnés. Dans un troisième temps nous exposons notre algorithme qui s'applique en présence de plusieurs médiateurs liés de façon non-causale. Enfin nous présentons des exemples d'applications de notre package avec des données simulées. Ces

simulations illustrent l'utilisation du package pour l'analyse de médiation multiple, l'analyse de médiation par blocs ainsi que la sélection de médiateurs en grande dimension.

## 3.2 Algorithmes d'estimation

### 3.2.1 Algorithme d'Imai

Nous présentons ici l'algorithme d'Imai dans le cas d'un seul médiateur tel qu'il apparaît dans Imai et al. (2010a). L'algorithme d'Imai se définit par une succession d'étapes présentée ci-dessous :

**Algorithme.** *Supposons que la quantité d'intérêt soit l'effet de médiation causale moyen,  $\delta(t)$*

1. *Faire les régressions des modèles de la variable de réponse et du médiateur. Estimer un modèle paramétrique,  $f_{\Theta_M}(M_i|T_i, X_i)$ , pour le médiateur et un autre modèle paramétrique  $f_{\Theta_Y}(Y_i|T_i, M_i, X_i)$ , pour la variable de sortie où  $\Theta_M$  et de  $\Theta_Y$  représentent les paramètres des modèles.*
2. *Simuler les paramètres à partir de leur distribution d'échantillonnage. Tirer  $J$  copies de  $\Theta_M$  et  $\Theta_Y$  à partir de leurs distributions d'échantillonnage notées  $\Theta_M^{(j)}$  et  $\Theta_Y^{(j)}$ , respectivement. On utilise l'approximation basée sur la distribution normale multivariée (avec moyenne et variance égales aux paramètres estimés et à leur matrice de covariance asymptotique, respectivement).*
3. *Pour chaque  $j = 1, \dots, J$  répéter les trois étapes suivantes :*
  - (a) *Simuler les médiateurs potentiels. Pour chaque  $t = 0, 1$  et chaque  $i = 1, \dots, n$ , tirer  $K$  copies de  $M_i(t)$  à partir de  $f_{\Theta_M^{(j)}}(M_i|t, X_i)$  que l'on note  $M_i^{(jk)}(t)$  pour  $k = 1, \dots, K$ . Ainsi on génère deux échantillons de médiateurs potentiels chacun basé sur le modèle paramétrique du médiateur, l'un conditionné par la valeur de traitement et l'autre par la valeur de contrôle.*
  - (b) *Simuler les variables de réponse potentielles conditionnellement aux valeurs du médiateur. Pour chaque  $t = 0, 1$  et chaque  $i = 1, \dots, n$ , tirer une copie de  $Y_i(t, M_i^{(jk)}(t'))$  à partir de  $f_{\Theta_Y^{(j)}}(Y_i|t, M_i^{(jk)}(t'), X_i)$  que l'on note  $Y_i^{(jk)}(t, M_i^{(jk)}(t'))$  pour  $k = 1, \dots, K$ . Ainsi pour chaque statut de traitement, on génère deux échantillons de variables de réponse potentielles, aussi en utilisant les modèles de prédiction. L'une utilise le médiateur généré sous le statut de traitement et l'autre utilise le médiateur généré sous le statut de contrôle.*
  - (c) *Calculer les effets causaux de médiation.*

$$\widehat{\delta}^{(j)}(t) = \frac{1}{nK} \sum_{i=1}^n \sum_{k=1}^K \{Y_i^{(jk)}(t, M_i^{(jk)}(1)) - Y_i^{(jk)}(t, M_i^{(jk)}(0))\}.$$

On fait simplement la différence entre les variables de sortie potentielles sous le statut de traitement et sous le statut de contrôle, ceci pour les  $n$  individus. On obtient alors  $\widehat{\delta}^{(j)}(t)$ , qui est un tirage de Monte Carlo de l'effet de médiation causale.

4. Calculer les statistiques descriptives. Finalement, on calcule les estimations  $\widehat{\delta}(t)$  et leur significativité à partir de la distribution de l'effet de médiation causale :  $\widehat{\delta}^{(j)}(t)$ . Par exemple, la médiane et la variance de la distribution peuvent être utilisées comme estimation de  $\widehat{\delta}(t)$  et de sa variance. De plus, les quantiles de la distribution peuvent servir d'intervalle de confiance  $\widehat{\delta}(t)$ .

### 3.2.2 Estimation de la corrélation entre médiateurs

Pour étendre l'algorithme précédent au cas de médiateurs liés de façon non-causale, il faut à l'étape 3(a) simuler les médiateurs en prenant en compte leur corrélations conditionnellement au traitement et covariables observées. Dans cette section, nous illustrons les méthodes utilisées pour estimer la matrice de variance-covariance  $\Sigma$  pour des médiateurs linéaires ou catégoriels ordonnés. Nous entendons par médiateurs linéaires, des variables continues pouvant être modélisées par un modèle linéaire. Par soucis de simplification, on considère deux médiateurs. Cependant la démarche est généralisable à un nombre  $K > 2$  de médiateurs.

— **Estimation de la corrélation entre médiateurs linéaires.**

Soit les modèles linéaires suivants :

$$\left( M_1, M_2 \mid T = t, X = x \right) \sim \mathcal{N} \left( \mu = \begin{pmatrix} \alpha_1 + \beta_1 t + \gamma_1 x \\ \alpha_2 + \beta_2 t + \gamma_2 x \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix} \right)$$

où  $\sigma_{1,2} = \rho \sigma_1 \sigma_2$ , avec  $\rho$  la corrélation entre  $M_1$  et  $M_2$  ;

$T$	$X$	$M_1$	$M_2$
1	0	6.4	8.2
1	1	10.9	10.8
0	1	-1.8	-3.4
0	0	-0.63	1
0	1	-5.7	-14.9
1	0	5.9	3

TABLE 3.1 – Exemple de données observées avec des médiateurs linéaires.

On estime  $\Sigma$  en calculant la matrice des variances-covariances entre les résidus obtenus après régression.

- **Estimation de la corrélation entre médiateurs catégoriels ordonnés.** Lorsque les médiateurs sont catégoriels ordonnés on ne peut pas estimer les corrélations

aussi simplement. Une solution possible est de se placer dans le cadre d'une modélisation probit multivariée.

Pour  $M_1$  et  $M_2$ , des variables catégorielles ordonnées, on peut considérer deux variables auxiliaires linéaires  $M_1^*, M_2^*$  et des constantes  $a_1, \dots, a_{Q+1}, a'_1, \dots, a'_{P+1}$  telles que (3.2.2). On considère les modèles linéaires (3.2.1).

$$\left( M_1^*, M_2^* \mid T = t, X = x \right) \sim \mathcal{N} \left( \mu = \begin{pmatrix} \alpha_1 + \beta_1 t + \gamma_1 x \\ \alpha_2 + \beta_2 t + \gamma_2 x \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right) \quad (3.2.1)$$

$M_1, M_2$  s'obtiennent à partir des variables linéaires  $M_1^*, M_2^*$  en appliquant la discrétisation suivante :

$$\begin{aligned} M_1 &= \sum_{q=1}^Q q \mathbb{1}_{\{a_q < M_1^* \leq a_{q+1}\}} \\ M_2 &= \sum_{p=1}^P p \mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}} \end{aligned} \quad (3.2.2)$$

avec  $a_1 = a'_1 = -\infty$  et  $a_{Q+1} = a'_{P+1} = \infty$ . Notons que les coefficients et les constantes peuvent être estimés à l'aide de la fonction `polr` sur R.

Comme on peut le voir dans la Table 3.2, en pratique on ne peut pas observer  $M_1^*$  et  $M_2^*$  du modèle (3.2.1), mais seulement  $M_1$  et  $M_2$  du modèle (3.2.2).

$T$	$X$	$M_1$	$M_2$
1	0	2	2
1	1	3	4
0	1	2	3
0	0	1	2
0	1	3	3
1	0	1	1

TABLE 3.2 – Exemple de données observables avec des médiateurs catégoriels ordonnés.

Les régressions sur  $M_1$  et  $M_2$  ne permettent pas d'avoir les résidus et d'estimer  $\rho = Cov(M_1^*, M_2^* \mid T = t, X = x)$ . Nous démontrons dans l'annexe C.2 que nous pouvons tout de même approcher la valeur de  $\rho$  en approchant la valeur qui annule la fonction suivante :

$$\begin{aligned} f(\rho) &= Cov(M_1, M_2 \mid T = t, X = x) - \sum_{q=1}^Q \sum_{p=1}^P qp \Phi_{(0, \Sigma)}(BS_q, BI_q, BS_p, BI_p) \\ &\quad + \sum_{q=1}^Q q \left( \Phi_{(0,1)}(BS_q) - \Phi_{(0,1)}(BI_q) \right) \times \sum_{p=1}^P p \left( \Phi_{(0,1)}(BS_p) - \Phi_{(0,1)}(BI_p) \right) \end{aligned}$$

avec

$$\begin{aligned} BI_q &= a_q - (\alpha_1 + \beta_1 t + \gamma_1 x) \\ BS_q &= a_{q+1} - (\alpha_1 + \beta_1 t + \gamma_1 x) \\ BI_p &= a_p - (\alpha_2 + \beta_2 t + \gamma_2 x) \\ BS_p &= a_{p+1} - (\alpha_2 + \beta_2 t + \gamma_2 x) \end{aligned}$$

Notons que les médiateurs binaires sont un cas particulier des médiateurs catégoriels. Avec des médiateurs binaires la fonction devient :

$$f(\rho) = Cov(M_1, M_2 | T = t) - \Phi_{(0, \Sigma)}(\alpha_1 + \beta_1 t, \alpha_2 + \beta_2 t) + \Phi_{(0, 1)}(\alpha_1 + \beta_1 t) \Phi_{(0, 1)}(\alpha_2 + \beta_2 t)$$

— **Estimation de la corrélation entre un médiateur linéaire et un médiateur catégoriel ordonné.**

En présence d'un médiateur linéaire  $M_1$  et d'un médiateur catégoriel ordonné  $M_2$ , on peut considérer comme précédemment une variable auxiliaire linéaire  $M_2^*$  et des constantes  $a_1, \dots, a_{P+1}$  telle que (3.2.4). On considère les modèles linéaires (3.2.3).

$$\left( M_1, M_2^* \mid T = t, X = x \right) \sim \mathcal{N} \left( \mu = \begin{pmatrix} \alpha_1 + \beta_1 t + \gamma_1 x \\ \alpha_2 + \beta_2 t + \gamma_2 x \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & 1 \end{pmatrix} \right) \tag{3.2.3}$$

$M_2$  s'obtient à partir de la variable linéaire  $M_2^*$  en appliquant la discrétisation suivante :

$$M_2 = \sum_{p=1}^P p \mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}} \tag{3.2.4}$$

$T$	$X$	$M_1$	$M_2$
1	0	8.2	2
1	1	10.8	3
0	1	-3.4	2
0	0	1	1
0	1	-14.9	3
1	0	3	1

TABLE 3.3 – Exemple de données observables avec un médiateur linéaire et un médiateur catégoriel.

On peut estimer  $\sigma_1^2$  à l'aide des résidus de la régression de  $M_1$ . Nous démontrons dans l'annexe C.3 que  $\rho$  peut être obtenu par :

$$\rho = \frac{\text{Cov}(M_1, M_2 | T = t, X = x)}{\sum_{p=1}^P p \int_{BI_p}^{BS_p} e_2 f_{(0,1)}(e_2) de_2}$$

$$\begin{aligned} BI_p &= a_p - (\alpha_2 + \beta_2 t + \gamma_2 x) \\ BS_p &= a_{p+1} - (\alpha_2 + \beta_2 t + \gamma_2 x) \end{aligned}$$

### 3.2.3 Algorithme pour médiateurs multiples non causalement corrélés

La preuve du Théorème 1.4.1 peut être généralisée pour prouver que, sous les hypothèses SIMMA, les densités des variables de réponse contrefactuelles peuvent s'exprimer de la façon suivante :

$$f(Y(t, M(t'), W(t')) | X = x) = \int_{\mathbb{R}^K} f(Y | T = t, M = m, W = w, X = x) dF_{(M, W) | T=t', X=x}(m, w) \quad (3.2.5)$$

$$f(Y(t, M(t'), W(t)) | X = x) = \int_{\mathbb{R}^K} f(Y | T = t, M = m, W = w, X = x) dF_{(M(t'), W(t)) | X=x}(m, w). \quad (3.2.6)$$

L'équation (3.2.5) justifie une estimation de Monte-Carlo de l'espérance  $E[Y(t, Z(t'))] = E[Y(t, M(t'), W(t'))]$ , et donc des effets direct, indirect joint et total. De plus sous la condition supplémentaire (1.4.2) (et en supposant que la loi jointe des médiateurs est complètement déterminée par son espérance et sa matrice de covariance), l'équation (3.2.6) permet également d'échantillonner  $Y(t, M(t'), W(t))$  et donc d'estimer son espérance à travers  $M$ . En particulier, SIMMA et (1.4.2) permettent d'estimer la matrice de covariance conditionnelle des médiateurs contrefactuels pour chaque combinaison possible d'interventions du traitement et des covariables de prétraitement.

En conséquence, nous adaptons l'algorithme quasi-bayésien présenté par Imai et al. (2010a) dans la section 3.2.1, à la situation de plusieurs médiateurs indépendants et corrélés de façon non causale.

**Algorithme.** *Estimer les effets d'intérêt :*

1. *Faire les régressions des modèles de la variable de réponse (compte tenu de tous les médiateurs, du traitement et des covariables) et des médiateurs (compte tenu du traitement et des covariables), désignés respectivement par  $\hat{\Theta}^Y$  et  $\hat{\Theta}^Z = (\hat{\Theta}^1, \dots, \hat{\Theta}^K)$ . Estimer  $\hat{\Sigma}_2$  la matrice de covariance entre les médiateurs (compte tenu du traitement et des covariables).*

2. Pour chaque modèle, échantillonner les  $J$  valeurs pour chacun de ses paramètres en fonction de leur distribution d'échantillonnage multivariée, notées  $\widehat{\Theta}_{(j)}^Y$ ,  $j = 1, \dots, J$  et  $\widehat{\Theta}_{(j)}^Z = (\widehat{\Theta}_{(j)}^1, \dots, \widehat{\Theta}_{(j)}^K)$ . Comme dans (Imai et al. (2010a)), nous utilisons l'approximation basée sur la distribution normale multivariée, avec une moyenne et une variance égales aux paramètres estimés et à leur matrice de covariance asymptotique estimée, respectivement.
3. Pour chaque  $j = 1, \dots, J$ , répéter les étapes suivantes :
- (a) Simuler les valeurs potentielles de chaque médiateur. Pour chaque paire différente de  $t, t' \in \{0, 1\}$  et chaque individu  $i = 1, \dots, n$ , simuler pour chaque  $k$  médiateurs  $R$  copies de  $Z_{(ji)}^{(kr)}(t, t') = \left( M_{(ji)}^{(kr)}(t), W_{(ji)}^{(kr)}(t') \right)$ . Lorsque tous les médiateurs ont la même valeur de traitement, le vecteur du médiateur sera noté  $Z_{(ji)}^{(r)}(t) = Z_{(ji)}^{(kr)}(t, t)$ . À noter que c'est à cette étape que l'on prend en compte la corrélation entre médiateurs, en simulant le vecteur d'erreurs suivant une loi gaussienne, avec  $\widehat{\Sigma}_2$  comme matrice de covariance.
- (b) Simuler les variables de sorties potentielles étant donné les valeurs simulées des médiateurs potentiels, notées  $Y_{(ji)}^{(r)}\left(t, Z_{(ji)}^{(kr)}(t', t'')\right)$  pour  $k$  et  $t, t', t'' \in \{0, 1\}$ .
- (c) Estimer les effets de médiation causale :

$$\begin{aligned} \widehat{\delta}_{(j)}^k(t) &= \frac{1}{nR} \sum_{i=1}^n \sum_{r=1}^R \left\{ Y_{(ji)}^{(r)}\left(t, Z_{(ji)}^{(kr)}(1, t)\right) - Y_{(ji)}^{(r)}\left(t, Z_{(ji)}^{(kr)}(0, t)\right) \right\} \\ \widehat{\delta}_{(j)}^Z(t) &= \frac{1}{nR} \sum_{i=1}^n \sum_{r=1}^R \left\{ Y_{(ji)}^{(r)}\left(t, Z_{(ji)}^{(r)}(1)\right) - Y_{(ji)}^{(r)}\left(t, Z_{(ji)}^{(r)}(0)\right) \right\} \\ \widehat{\zeta}_{(j)}(t) &= \frac{1}{nR} \sum_{i=1}^n \sum_{r=1}^R \left\{ Y_{(ji)}^{(r)}\left(1, Z_{(ji)}^{(r)}(t)\right) - Y_{(ji)}^{(r)}\left(0, Z_{(ji)}^{(r)}(t)\right) \right\} \\ \widehat{\tau}_{(j)}(t) &= \frac{1}{nR} \sum_{i=1}^n \sum_{r=1}^R \left\{ Y_{(ji)}^{(r)}\left(1, Z_{(ji)}^{(r)}(1)\right) - Y_{(ji)}^{(r)}\left(0, Z_{(ji)}^{(r)}(0)\right) \right\}. \end{aligned}$$

4. En prenant la moyenne des distributions empiriques de chaque effet ci-dessus, on obtient des estimations. On peut utiliser ces distributions pour calculer les  $p$ -valeurs, comme expliqué dans l'annexe C.1. Les intervalles de confiance peuvent être obtenus en calculant les quantiles de ces distributions empiriques.

Il s'ensuit de l'algorithme qu'il est de complexité  $\mathcal{O}(J \times n \times R \times K)$ . Notons que cet algorithme n'implémente pas les formules données pour les modèles spécifiques des Corollaires 1.5.1 et 1.6.1.



### 3.3 Présentation du package

Nous avons implémenté, dans notre package R `multimediate`, nos méthodes d'analyse de médiation en présence de médiateurs corrélés de façon non causale, d'analyse de médiation par bloc et de sélection de médiateur en grande dimension. Les modèles supportés par notre package R `multimediate` sont montrés dans la Table 3.4.

<i>Médiateurs</i>	<i>Variable de réponse</i>			
	Linéaire	Binaire	Catégoriel ordonné	Catégoriel non-ordonné
1) Linéaire	✓	✓	✓	✗
2) Binaire (Probit)	✓	✓	✓	✗
3) Catégoriel ordonné (Probit)	✓	✓	✓	✗
4) Catégoriel non-ordonné	✗	✗	✗	✗
Mélange de 1), 2) et 3)	✓	✓	✓	✗

TABLE 3.4 – Les modèles supportés par notre package R `multimediate`.

#### 3.3.1 Analyse de médiation en présence de médiateur corrélés de façon non-causale

##### Présentation de la fonction `multimediate`

`multimediate` est utilisé pour estimer diverses quantités pour l'analyse de médiation causale, c'est-à-dire les effets médiés moyens (effets indirects), les effets directs moyens, les proportions médiées et l'effet total, en présence de plusieurs médiateurs liés de façon non causale. Avec une variable de réponse binaire, `multimediate` estime également les effets de médiation causale moyens à l'échelle OR et à l'échelle logOR.

La fonction `multimediate` s'utilise de la manière suivante :

```
multimediate( lmodel.m, correlated = FALSE, model.y, treat, treat.value
= 1, control.value = 0, J = 1000, conf.level = 0.95 )
```

- `lmodel.m` : liste des objets de régression des médiateurs. Ils peuvent être de classe 'lm', 'polr', 'glm'.
- `correlated` : une valeur logique. Si la valeur est 'FALSE', une matrice d'identité est utilisée pour la matrice de corrélation des médiateurs ; si la valeur est 'TRUE', la matrice de corrélation est estimée. La valeur par défaut est 'FALSE'.

- `model.y` : un objet de régression pour la variable de réponse. Il peut être de classe 'lm', 'polr', 'glm'.
- `treat` : une chaîne de caractères indiquant le nom de la variable de traitement utilisée dans les modèles. Le traitement peut être binaire (entier ou à deux valeurs) ou continu.
- `treat.value` : valeur de la variable de traitement utilisée comme condition de traitement. La valeur par défaut est 1.
- `control.value` : valeur de la variable de traitement utilisée comme condition de contrôle. La valeur par défaut est 0.
- `J` : nombre de tirages de Monte Carlo pour une approximation quasi-bayésienne. La valeur par défaut est 1000.
- `conf.level` : niveau des intervalles de confiance bilatéraux renvoyés. La valeur par défaut renvoie les 2.5 et 97.5 centiles des quantités simulées.

`multimediate` renvoie un objet de classe "mm", une liste qui contient des composants générés pendant l'utilisation de la fonction. La fonction `summary` (c'est-à-dire `summary.mm`) peut être utilisée pour obtenir un tableau des résultats.

### Exemple d'utilisation de la fonction `multimediate`

On se place dans le contexte du graphe causal représenté dans la Figure 3.1, avec un traitement  $T$  binaire, un outcome continu, trois médiateurs continus  $M_1, M_2, M_3$  et des covariables binaires.

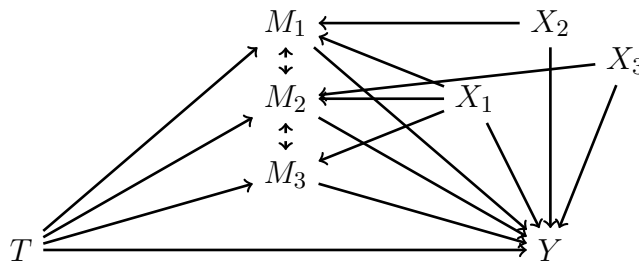


FIGURE 3.1 – DAG du premier exemple d'utilisation du package.

Les données ont été simulées avec le modèle suivant :

$$\begin{aligned}
 X_1 &\sim \mathcal{B}(0.5) \\
 X_2 &\sim \mathcal{B}(0.4) \\
 X_3 &\sim \mathcal{B}(0.6) \\
 (M_1, M_2, M_3 \mid T = t, X = x) &\sim \mathcal{N}\left(\mu = \begin{pmatrix} 1 + 2t + 1x_1c + 5x_2 \\ 2 + 3t + 5x_1 + 3x_3 \\ 2 + 6t + 3x_1 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.9 & 0.6 \\ 0.9 & 1 & 0.3 \\ 0.6 & 0.3 & 1 \end{pmatrix}\right) \\
 Y &\sim \mathcal{N}(7 + 10T + 5M_1 + 6M_2 + 9M_3 + 2X_1 + 3X_2 + 1X_3, 1)
 \end{aligned}$$

```

data(data1)
data1$Treatment=as.factor(data1$Treatment)
data1$X1=as.factor(data1$X1)
data1$X2=as.factor(data1$X2)
data1$X3=as.factor(data1$X3)
data1$M1=as.numeric(data1$M1)
data1$M2=as.numeric(data1$M2)
data1$M3=as.numeric(data1$M3)
data1$Outcome=as.numeric(data1$Outcome)
summary(data1)
#> Treatment C1      C2      C3      M1      M2
#> 0:480      0:501  0:417  0:596  Min.   :-1.410  Min.   :-0.8455
#> 1:520      1:499  1:583  1:404  1st Qu.: 2.850  1st Qu.: 4.7567
#>                                     Median : 5.809  Median : 7.2685
#>                                     Mean   : 5.431  Mean   : 7.2333
#>                                     3rd Qu.: 7.921  3rd Qu.: 9.7710
#>                                     Max.   :12.054  Max.   :15.6520
#>      M3      Outcome
#> Min.   :-0.8298  Min.   : 12.88
#> 1st Qu.: 3.5296  1st Qu.:100.61
#> Median : 6.8410  Median :145.90
#> Mean   : 6.6268  Mean   :145.54
#> 3rd Qu.: 9.7148  3rd Qu.:190.45
#> Max.   :13.5039  Max.   :295.79

```

La Table 3.5 indique les valeurs des effets causaux avec ce modèle.

Pour réaliser l'analyse de médiation multiple, on commence par faire la régression de chaque médiateur compte tenu du traitement et des covariables ainsi que la régression de l'outcome compte tenu du traitement, des médiateurs et des covariables à l'aide de la fonction `lm`. Les sorties de ces appels à `lm` sont alors utilisées comme argument de `multimediate`.

$\tau$	$\zeta$	$\delta^Z$	$\delta^{M_1}$	$\delta^{M_2}$	$\delta^{M_3}$
92	10	82	10	18	54

TABLE 3.5 – Valeur des effets causaux pour le premier jeu de donnée. Pour rappel  $\delta^a$  : effet médié par  $a$ ,  $\zeta$  : effet direct et  $\tau$  : effet total.

```

M1reg=lm(M1~ Treatment + X1 + X2, data=data1)
M2reg=lm(M2~ Treatment + X1 + X3, data=data1)
M3reg=lm(M3~ Treatment + X1      , data=data1)

Yreg=lm(Outcome~ Treatment + M1 + M2 + M3 + X1 + X2 + X3, data=data1)

med.analysis=multimediate(lmodel.m=list(M1reg,M2reg,M3reg),correlated=TRUE,
model.y=Yreg,treat="Treatment",treat.value=1,control.value=0,J=1000,
conf.level=0.95)
summary(med.analysis,opt="avg")
#>           . Estimation IC.inf IC.sup P.val
#> 1  ACME.joint    81.5677 79.9315 83.2901    0
#> 2   PM.joint     0.8860 0.8780 0.8945    0
#> 3  ACME.M1     10.3892  9.5890 11.1634    0
#> 4   PM.M1       0.1127 0.1041 0.1214    0
#> 5  ACME.M2     17.5371 16.6758 18.5698    0
#> 6   PM.M2       0.1904 0.1805 0.2015    0
#> 7  ACME.M3     53.6414 52.3379 54.9294    0
#> 8   PM.M3       0.5827 0.5661 0.5996    0
#> 9    ADE       10.4827  9.7151 11.2201    0
#> 10 Total Effect  92.0504 90.4867 93.5400    0

```

On obtient une bonne estimation significative de chaque effet causal.

### 3.3.2 Analyse de médiation par bloc

#### Présentation de la fonction `blockmediate`

`blockmediate` est utilisée pour estimer diverses quantités pour l'analyse de médiation causale comme `multimediate`, non pas pour plusieurs médiateurs mais pour différents blocs de médiateurs.

La fonction `blockmediate` s'utilise de la manière suivante :

```

blockmediate( lmodel.m, model.y, treat, treat.value = 1, control.value =
0, J = 1000, conf.level = 0.95, clust = NULL )

```

- `lmodel.m` : liste des objets de régression des médiateurs. Ils peuvent être de classe `'lm'`, `'polr'`, `'glm'`.
- `model.y` : un objet de régression pour la variable de réponse. Il peut être de classe `'lm'`, `'polr'`, `'glm'`.
- `treat` : une chaîne de caractères indiquant le nom de la variable de traitement utilisé dans les modèles. Le traitement peut être binaire (entier ou à deux valeurs) ou continu.
- `treat.value` : valeur de la variable de traitement utilisée comme condition de traitement. La valeur par défaut est 1.
- `control.value` : valeur de la variable de traitement utilisée comme condition de contrôle. La valeur par défaut est 0.
- `J` : nombre de tirages de Monte Carlo pour une approximation quasi-bayésienne. La valeur par défaut est 1000.
- `conf.level` : niveau des intervalles de confiance bilatéraux renvoyés. La valeur par défaut renvoie les 2.5 et 97.5 centiles des quantités simulées.
- `clust` : vecteur de longueur égale au nombre de médiateurs. L'ordre des numéros correspond à l'ordre des médiateurs dans `lmodel.m` et les médiateurs appartenant au même cluster ont le même numéro. La valeur par défaut est `'NULL'`, les clusters sont définis en utilisant le clustering spectral.

`blockmediate` renvoie un objet de classe "bm", une liste qui contenant des composants générés pendant l'utilisation de la fonction. La fonction `summary` (c'est-à-dire `summary.bm`) peut être utilisée pour obtenir un tableau des résultats.

**Exemple d'utilisation de la fonction blockmediate**

On considère des données simulées avec le modèle suivant :

$$\begin{aligned}
 X_1 &\sim \mathcal{B}(0.7) \\
 \begin{pmatrix} M_1 \\ M_2^* \\ M_3^* \\ M_4 \\ M_5^* \\ M_6^* \\ M_7 \\ M_8^* \\ M_9^* \end{pmatrix} \Bigg|_{T=t, X_1=x_1} &\sim \mathcal{N} \left( \mu = \begin{pmatrix} -3+4t+2x_1 \\ -0.6+0.5t+0.15x_1 \\ 2+5t+3x_1 \\ -3.5+3t+5x_1 \\ -0.4+0.8t+0.13x_1 \\ 8+3t+4x_1 \\ 1+0t \\ -0.6+0.7t \\ -3+2t \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \Sigma_2 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \Sigma_3 \end{pmatrix} \right) \\
 M_i &= \sum_{q_i=1}^{Q_i} q_i \mathbb{1}_{\{a_{q_i} < M_i^* \leq a_{q_i+1}\}} \forall i \in \{2, 3, 5, 6, 8, 9\} \\
 Y &\sim \mathcal{N}(5+6T+7M_1+8M_2+0M_3+10M_4+11M_5+12M_6+13M_7 \\
 &\quad +14M_8+15M_9, 1)
 \end{aligned}$$

$$\text{avec } \Sigma_i = \begin{pmatrix} 1 & \rho & \rho \\ \rho & 1 & \rho \\ \rho & \rho & 1 \end{pmatrix} \mathbf{0} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

La Table 3.6 indique les valeurs des effets causaux avec ce modèle.

$\tau$	$\zeta$	$\delta^Z$	$\delta^{B_1}$	$\delta^{B_2}$	$\delta^{B_3}$
169.7	6	163.7	77.3	53.1	33.3

TABLE 3.6 – Valeur des effets causaux pour le deuxième jeu de données. Pour rappel  $\delta^a$  est l'effet médié par  $a$ ,  $\zeta$  est l'effet direct et  $\tau$  est l'effet total.

On réalise les régressions linéaires à l'aide de `lm`, les régressions binaires à l'aide de `glm` et les régressions catégorielles ordonnées à l'aide de `polr`.

```

M1reg=polr(M1~ Treatment + X1, data=data5, method="probit")
M2reg=polr(M2~ Treatment + X1, data=data5, method="probit")
M3reg=polr(M3~ Treatment, data=data5, method="probit")

```

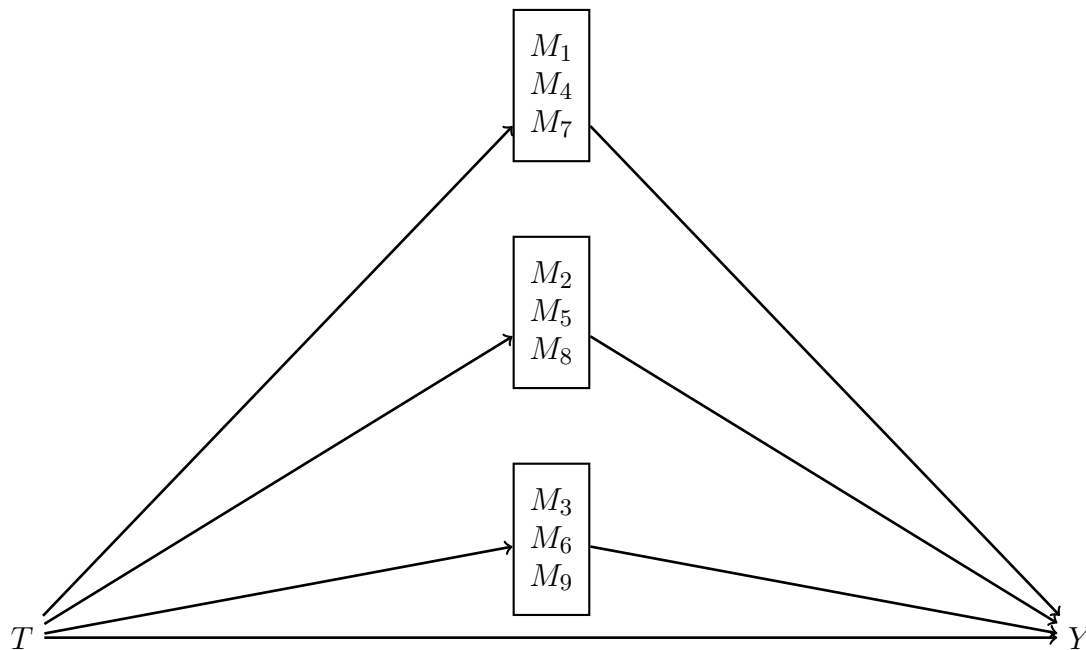


FIGURE 3.2 – DAG du deuxième exemple d'utilisation du package.

```
M4reg=glm(M4~ Treatment + X1, data=data5 family=binomial("probit"))
M5reg=glm(M5~ Treatment + X1, data=data5 family=binomial("probit"))
M6reg=glm(M6~ Treatment, data=data5 family=binomial("probit"))

M7reg=lm(M7~ Treatment + X1      , data=data5)
M8reg=lm(M8~ Treatment + X1      , data=data5)
M9reg=lm(M9~ Treatment, data=data5)

Yreg=lm(Outcome~ Treatment + M1 + M2 + M3 + M4 + M5 + M6 + M7 + M8 + M9,
data=data5)
```

On utilise ensuite les sorties des fonctions `lm`, `glm` et `polr` comme arguments de la fonction `blocmediate`.

```
bloc.analysis=blocmediate(lmodel.m=list(M1reg,M2reg,M3reg,M4reg,M5reg,
M6reg,M7reg,M8reg,M9reg),model.y=Yreg,treat="Treatment"
,treat.value=1,control.value=0,J=1000,conf.level=0.95)
```

Une fois l'algorithme de classification de médiateurs lancé, les valeurs propres de la matrice  $U$  sont calculées. On peut voir l'ébouilissement des valeurs propres dans la Figure 3.3. On voit qu'il y en a trois qui sont proches de 0, ce qui signifie qu'il y a trois blocs de médiateurs identifiés.

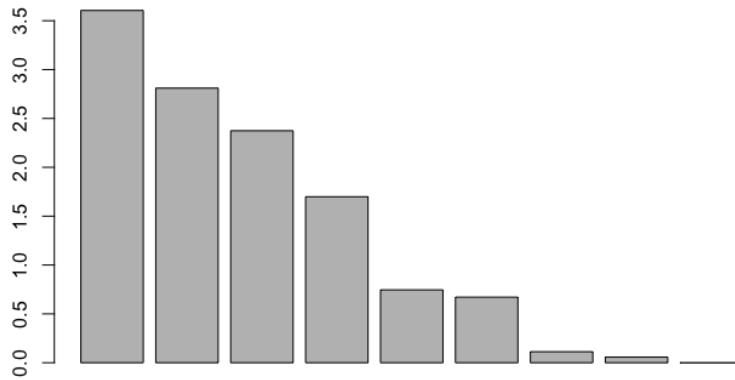


FIGURE 3.3 – Éboullis des valeurs propres

```

#>           .How many cluster?
3
bloc.analysis$clust
#>       1 3 2 1 3 2 1 3 2
summary(med.analysis,opt="avg")
#>           . Estimation  IC.inf  IC.sup P.val
#> 1  ACME.joint  157.0005 153.7084 160.1633  0
#> 2  PM.joint    0.9618  0.9593  0.9641  0
#> 3  ACME.Block.1 74.4760 72.6473 76.2965  0
#> 4  PM.Block.1  0.4562  0.4417  0.4705  0
#> 5  ACME.Block.2 48.9775 47.0978 51.0037  0
#> 6  PM.Block.2  0.2998  0.2871  0.3142  0
#> 7  ACME.Block.3 33.5469 31.5125 35.4381  0
#> 8  PM.Block.3  0.2054  0.1922  0.32187  0
#> 9  ADE         6.2336  5.8750  6.6025  0
#> 10 Total Effect 163.2341 159.9360 166.3634  0

```

On obtient une bonne estimation significative de chaque effet causal.



### 3.3.3 Sélection de médiateur en grande dimension

#### Présentation de la fonction `mahi`

`mahi` est utilisé pour sélectionner des médiateurs en présence d'un nombre élevé de médiateurs et estimer diverses quantités pour l'analyse de la médiation causale en utilisant `multimediate`, avec les médiateurs sélectionnés.

```
mahi( data, name.exposure, name.outcome, name.mediators, nboot = 100, nlambda
= 50, lambda = NULL, Kmax = NULL, p.adjust.method = "bonferroni", pvalseuil
= 0.05 )
```

- `data` : un `data.frame` avec l'exposition, la variable de réponse et les médiateurs.
- `name.exposure` : une chaîne de caractères indiquant le nom de l'exposition dans les données.
- `name.outcome` : une chaîne de caractères indiquant le nom de la variable de réponse dans les données.
- `name.mediators` : une chaîne de caractères indiquant le nom des médiateurs dans les données.
- `nboot` : nombre de bootstrap pour la stability selection des médiateurs. La valeur par défaut est 100.
- `nlambda` : nombre lambda utilisé lorsque la pénalité lambda n'est pas spécifiée.
- `lambda` : vecteur ou une valeur unique d'utilisation lambda pour le groupe-lasso.
- `Kmax` : le maximum de médiateurs retenu après le classement obtenu avec la stability selection.
- `p.adjust.method` : une chaîne de caractères indiquant le nom de la méthode de correction pour le test multiple. Voir l'aide de `p.adjust.methods`. La méthode par défaut est celle de Bonferroni.
- `pvalseuil` : la  $p$ -valeur pour le test multiple.

`mahi` renvoie un objet de classe `mahi`, une liste qui contient des composants générés pendant l'utilisation de la fonction.

#### Exemple d'utilisation de la fonction `mahi`

On considère les données simulées avec le modèle de l'équation (2.3.6) de la section 2.3.3 avec les valeurs de paramètre de la table 2.1.

Les vrais médiateurs à sélectionner sont les 30 premiers ( $M_1, \dots, M_{30}$ ).

```
mahifit <- mahi(data3, name.exposure="T", name.outcome="Y",
name.mediators=paste0("M", 1:500), nboot=30, nlambda=1, lambda=0.008522981,
Kmax=43, p.adjust.method="bonferroni", pvalseuil=0.05)
```

Ensuite, pour afficher les médiateurs sélectionnés on utilise la commande suivante.

```
mahiselection <- as.vector(which(mahifit$selection != FALSE))
mahiselection
```

```
## [1] M1 M2 M3 M4 M5 M6 M7 M8 M9 M10
```



# Chapitre 4

## Applications à des données réelles

### Sommaire

---

<b>4.1 Effet de l'hormonothérapie substitutive sur le cancer du sein</b>	<b>91</b>
<b>4.2 Effets des tempéraments sur l'anorexie mentale . . . . .</b>	<b>100</b>

---

Nous avons appliqué notre méthode d'analyse de médiation multiple sur des données réelles.

Premièrement, en collaboration avec Laura Baglietto, Marco Fornili du Département de Médecine Clinique de l'Université de Pise et Gianluca Severi de la Faculté de Médecine de l'Université Paris-Saclay, nous avons travaillé sur l'effet de l'hormonothérapie substitutive sur le cancer du sein. Ce travail a permis la rédaction d'un article (Fornili et al. (2020), dont nous présentons une partie des résultats.

Deuxièmement, en collaboration avec Enrica Marzola du Département Neurosciences de l'Université de Turin, nous avons aussi travaillé sur l'effet des tempéraments affectifs sur l'anorexie mentale. Ce travail fait l'objet d'un article en cours de rédaction. Nous présentons ici les résultats qui figureront dans l'article.

### 4.1 Effet de l'hormonothérapie substitutive sur le cancer du sein

Les zones mammaires denses, c'est-à-dire les zones du sein qui apparaissent blanches sur une mammographie, font partie des facteurs de risque les plus forts pour le cancer du sein (CS) chez les femmes ménopausées ou non. Des études longitudinales ont montré que la densité mammaire (DM) diminue avec l'âge, avec un plus grand déclin se produisant à la transition ménopausique (Krishnan et al. (2017), Kelemen et al. (2008), McCormack et al. (2010)). L'hormonothérapie substitutive (THS) peut être prescrite aux femmes pour

équilibrer la diminution des œstrogènes se produisant à la ménopause et a pour effet d'augmenter le niveau de la DM (Cagnacci and Venier (2019)). L'un des effets secondaires les plus préoccupants du THS est le risque accru de cancer lié aux hormones, y compris le cancer du sein et de l'endomètre (muqueuse de l'utérus, qui en recouvre la paroi interne). Une analyse récente regroupée sur plus de 130 000 cas postménopausaux de cancer du sein diagnostiqués dans le cadre de 58 études prospectives a conclu que la THS augmentait le risque de cancer du sein chez les utilisateurs, même pendant les 1 à 4 premières années et était toujours persistant 10 ans après l'arrêt (for Research on Cancer and Weltgesundheits organisation (2012)). La même analyse groupée a rapporté un risque accru associé aux œstrogènes-progestatifs et aux œstrogènes seuls, avec un effet légèrement plus élevé dans le premier cas. Les études observationnelles et d'intervention ont montré que la densité mammographique augmente avec le THS et que l'effet est plus fort pour l'utilisation d'œstrogène-progestatif que pour les œstrogènes seuls (Collaborative Group on Hormonal Factors in Breast Cancer (2019)). Les données probantes suggèrent que la proportion de l'effet du THS sur le risque du cancer du sein médiée par son action sur la densité mammographique varie de 30% à 100% (Byrne et al. (2017); Rice et al. (2016)), mais l'effet du mode d'utilisation sur les mécanismes de médiation n'a pas encore été étudié. Le but de notre travail est d'étudier l'association entre les modes d'utilisation du THS et la DM afin de mieux comprendre leurs effets indépendant et médié sur le risque du cancer du sein. Pour nos analyses, nous avons utilisé les données des femmes ménopausées d'une étude cas-témoins du cancer du sein dans la cohorte française E3N (Rice et al. (2016)).

**Présentation des données** Les données proviennent de 906 femmes âgées de plus de 55 ans lors de la mammographie dans l'étude cas-témoins du CS comprises dans la cohorte française E3N (Maskarinec et al. (2017)); l'âge de 55 ans a été choisi comme indicateur du statut postménopausal. Les détails de l'étude cas-témoins imbriquée sont fournis dans (Rice et al. (2016)). En bref, les femmes ont été appariées par l'adénocarcinome invasif du sein diagnostiqué entre 1990 et 2010 avec une latéralité connue et au moins une mammographie effectuée entre le début de l'étude et le diagnostic. L'appariement a été fait en utilisant une procédure d'échantillonnage de densité aux femmes de la cohorte sans CS à l'âge qu'elles avaient au moment du diagnostic du cas correspondant (âge de référence); les facteurs d'appariement comprenaient également l'année de naissance ( $\pm 3$  ans) et le statut ménopausique au départ. Des informations détaillées sur les caractéristiques démographiques, le mode de vie et les antécédents reproductifs étaient disponibles pour toutes les femmes de la cohorte à partir de questionnaires structurés collectés au départ et tous les 2 à 3 ans pendant le suivi. Après avoir exclu les femmes avec des valeurs manquantes, 453 cas et 453 témoins étaient disponibles pour l'analyse.

Le mode d'utilisation du THS et des contraceptifs oraux pendant le suivi, y compris le statut d'utilisation (*jamais / actuel / ancien*), le type de thérapie (œstrogène seul contre œstrogène-progestatif) ont été calculés à partir des données recueillies au travers des questionnaires.

**Modèles de médiations** Tout d'abord, nous avons créé des modèles dans lesquels l'utilisation de la THS était catégorisée en *jamais*, *actuelle* et *ancienne*. Puis pour tenir compte du mode d'utilisation, la durée pour les utilisateurs actuels et anciens et le temps depuis la dernière utilisation pour les anciens utilisateurs ont été dichotomisées en fonction de leur médianes.

La DM a été quantifiée à partir de l'image cranio-caudale du sein homolatéral à la tumeur la plus proche et antérieure à l'âge de référence. Les films mammographiques ont été numérisés avec un numériseur de films haute densité Array 2905 (Array Corporation Europe, Roden, Pays-Bas) avec une résolution de 300 PPI et ont été redimensionnés pour la lecture de densité avec une taille maximale proportionnelle de  $800 \times 400$  pixels. Un lecteur unique (GM), qui ne connaissait pas le statut cas-témoins, a évalué la superficie totale du sein et la zone dense (DA) par lots de 200 mammographies à l'aide d'une technique assistée par ordinateur (Cumulus, Sunnybrook Health Sciences Centre, University of Toronto, Toronto, Canada). La densité mammographique en pourcentage (PDM) a été calculée comme le rapport de la DA à la surface totale du sein, et la zone non dense (NDA) comme la différence entre la surface totale du sein et la DA. Ces variables continues ont été normalisées en utilisant l'approche de vraisemblance de Box-Cox (Box and Cox (1964)),  $t(M) = \frac{M^\lambda - 1}{\lambda}$ , avec  $\lambda$  égale à 0.38 (respectivement 0.34) pour DA (respectivement NDA) comme nous pouvons le voir sur la figure 4.1.

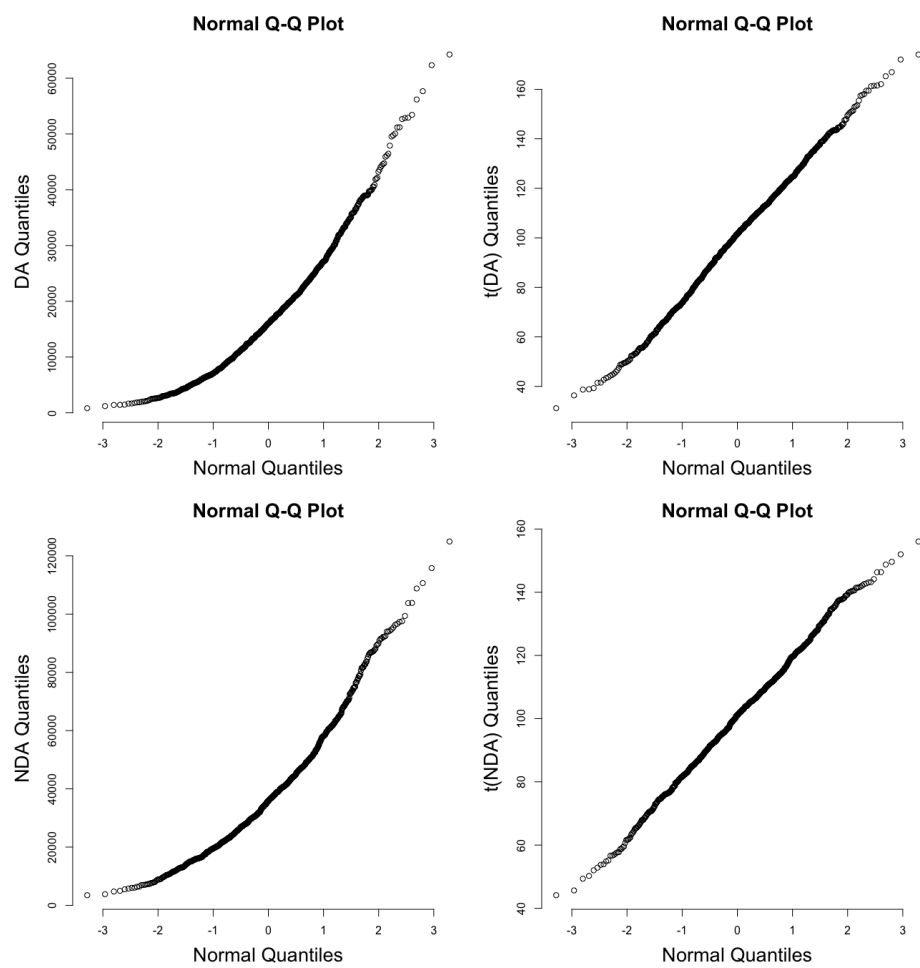


FIGURE 4.1 – Diagrammes Quantile-Quantile de comparaison des distributions des médiateurs avec une loi gaussienne réduite, avant et après la transformation. Pour rappel DA : zone mammaire dense et NDA : zone mammaire non-dense.

#### 4.1. EFFET DE L'HORMONOTHÉRAPIE SUBSTITUTIVE SUR LE CANCER DU SEIN 95

L'effet du THS sur chaque mesure DM a été estimé en ajustant les modèles de régression sur les variables de DM transformées.

Nous avons évalué l'effet des variables de confusions potentielles suivantes : antécédents familiaux du CS ; âge à la ménarche (période des premières menstruations) ; utilisation antérieure de contraceptifs oraux ; parité et allaitement.

L'indice de masse corporel (IMC) étant considéré comme un médiateur de l'effet du THS sur la DM, il n'a pas été inclus parmi les facteurs de confusion potentiels.

La prévalence du CS dans la population réelle de femmes âgées de plus de 55 ans est  $p = 0.08$  (van der Laan (2008)). Nous avons donc une sur-représentation des cas dans notre étude cas-témoins. Les estimateurs obtenus à partir de la régression logistique, en utilisant les données cas-témoins, permettront d'estimer systématiquement les mêmes paramètres qu'une régression logistique utilisant des données de cohorte. Ce résultat bien connu est ce qui justifie l'utilisation de la régression logistique lors de l'analyse des rapports de cotes dans les études cas-témoins pour des effets totaux. Donc lorsque la régression logistique est utilisée, la sur-représentation des cas peut être ignorée. Cependant, lors de l'ajustement du modèle de régression linéaire pour un médiateur en utilisant des données cas-témoins, la sur-représentation des cas ne peut pas être ignorée. On peut tout de même adapter l'approche d'estimation des effets directs et indirects lorsque la prévalence de la variable de réponse est connue. Nous ajustons la régression linéaire du médiateur sur le traitement et les covariables en utilisant les données cas-témoins mais en pondérant chaque cas par  $\frac{p}{2}$  et chaque contrôle par  $\frac{(1-p)}{2}$ . Les coefficients de cette régression pondérée donneront des estimateurs consistants de ceux qu'on aurait obtenus d'une régression linéaire avec des données d'une étude de cohorte de la même population (VanderWeele and Vansteelandt (2010)).

Les analyses de médiation ont été menées pour tous les CS et séparément pour les récepteurs hormonaux positifs (ER positif ou PR positif) et les récepteurs hormonaux négatifs (ER négatif et PR négatif). Dans les deux approches, la proportion médiée par PDM et l'IMC a été calculée sur l'échelle logarithmique ( $\log(\text{OR médié}) / \log(\text{OR total})$ ) (VanderWeele and Vansteelandt (2010)).

Pour estimer l'effet total du THS sur le risque de CS et l'effet médié par la PDM, nous avons adopté deux approches différentes. Premièrement, une approche de médiation simple en considérant comme médiateur la variable PDM. Les odds ratios (OR) pour le CS ont été estimés par un modèle logistique ajusté pour les variables d'appariement (c.-à-d. âge de référence, année de naissance, statut ménopausique au départ) (ajustement partiel). L'effet additionnel des antécédents familiaux de CS, de l'âge de la ménarche, de l'utilisation de contraceptifs oraux, de la parité et de l'allaitement comme facteurs de confusion supplémentaires potentiels a également été évalué (ajustement complet).

Sur la base des preuves issues d'études d'association sur le risque de cancer du sein (Sung et al. (2010), Boyd et al. (2006)), nous pouvons raisonnablement supposer que



l'IMC et la DM sont corrélés de manière non causale, leur corrélation étant probablement due à des traits génétiques communs, comme suggéré par des études jumellées (Sung et al. (2010), Wade et al. (2010)) et d'analyse de randomisation mendélienne (Ooi et al. (2019)). Pour tenir compte de la corrélation entre la PDM et l'IMC dans l'évaluation de leur rôle joint en tant que médiateurs de l'effet du THS sur le risque de CS, nous avons, dans un deuxième temps, appliqué notre méthode d'analyse de médiation multiple.

Nous faisons l'hypothèse implicite que le THS précède les médiateurs et que ceux-ci précèdent le cancer du sein. La Figure 4.2 illustre les hypothèses causales faites pour l'analyse de médiation suivante.

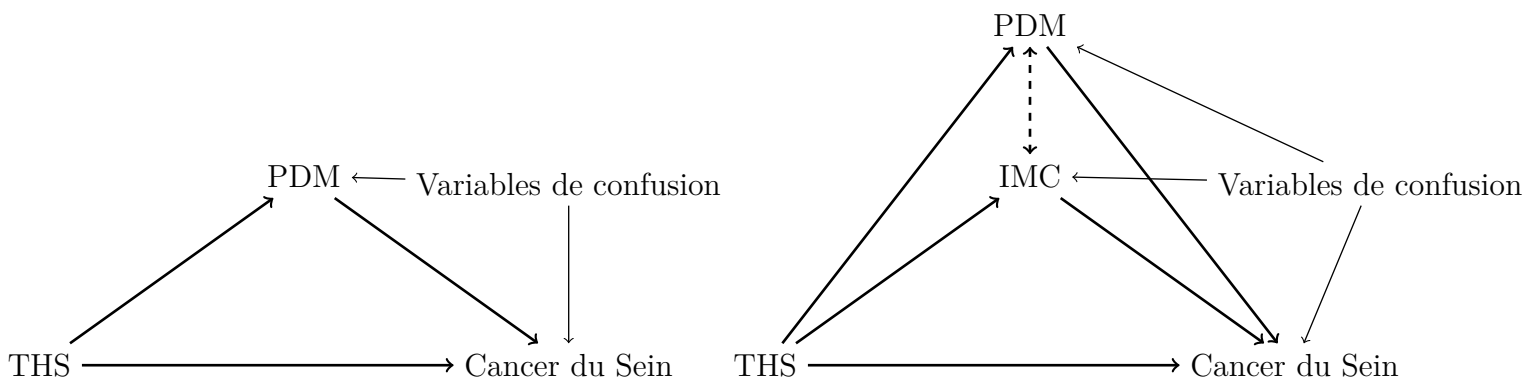


FIGURE 4.2 – Diagrammes causaux de l'hormonothérapie substitutive sur la densité mammographique. Les variables de confusion sont : l'âge de référence, l'année de naissance, le statut ménopausique au début de l'étude, les antécédents familiaux de CS, l'âge de la ménarche, l'utilisation de contraceptifs oraux, la parité et l'allaitement. Pour rappel THS : thérapie hormonal substitutive, PDM : pourcentage de densité mammographique.

**Association entre le THS et la DM.** Le statut d'utilisation du THS à la mammographie était significativement associé à PDM, DA et NDA ( $P < 0,001$ ,  $< 0,001$  et  $0,006$ , respectivement) : pour PDM et DA, les anciennes utilisatrices avaient des niveaux significativement plus élevés que celles qui n'en avait jamais utilisé ( $P = 0,003$  et  $P = 0,005$  respectivement) et des niveaux nettement inférieurs à celles qui en utilise actuellement ( $P = 0,004$  et  $P = 0,003$  respectivement) ; pour NDA, les anciennes et actuelles utilisatrices avaient des niveaux inférieurs à celles qui n'ont jamais utilisé THS, mais les différences n'étaient statistiquement significatives que pour les utilisatrices actuelles par rapport à celles qui n'ont jamais utilisé THS ( $P = 0,001$ ). La Table 4.1 montre les valeurs de DM prédites par type d'utilisation de THS pour une femme âgée de 60 ans. Distinguer les utilisatrices de THS par type d'utilisation n'a pas amélioré la significativité pour aucune des variables DM ( $P = 0,86$ ,  $0,63$  et  $0,67$ , respectivement).

#### 4.1. EFFET DE L'HORMONOTHÉRAPIE SUBSTITUTIVE SUR LE CANCER DU SEIN<sup>97</sup>

Utilisation du THS	N	PMD (%)	DA (cm <sup>2</sup> )	NDA (cm <sup>2</sup> )
		Prédits (IC à 95 %)	Prédits (IC à 95 %)	Prédits (IC à 95 %)
Jamais	189	24 (22 à 26)	24 (22 à 26)	78 (72 à 83)
Tout type de THS				
Ancienne	285	29 (27 à 31)	29 (26 à 31)	71 (66 à 76)
Actuelle	432	33 (31 à 35)	33 (31 à 35)	67 (63 à 70)
Œstrogène				
Ancienne	13	28 (18 à 40)	25 (15 à 36)	62 (43 à 86)
Actuelle	17	35 (27 à 45)	34 (25 à 43)	64 (49 à 81)
Œstrogène plus progestérone				
Ancienne	197	30 (27 à 32)	30 (27 à 33)	72 (67 à 78)
Actuelle	293	33 (31 à 35)	33 (31 à 35)	65 (62 à 70)

TABLE 4.1 – Mesures mammographiques prédites chez les femmes de 60 ans par type d'hormonothérapie ménopausique. Pour rappel THS : thérapie hormonal substitutive, PMD : pourcentage de densité mammographique, NDA : zone mammaire non-dense et DA : zone mammaire dense.

**Analyse de médiation simple.** Considérant que l'effet du THS sur la DM peut être observé jusqu'à 8 ans après l'arrêt du traitement, nous avons effectué des analyses de médiation uniquement sur les utilisatrices actuelles et celles n'ayant jamais pris la THS. Dans le modèle partiellement ajusté, l'OR du CS associé à l'utilisation actuelle ou jamais de THS était de 1,67 (IC à 95%, 1,04 à 2,68) ; lorsque PDM a été ajouté dans le modèle, l'OR est passé à 1,40 (IC à 95%, 0,86 à 2,28), ce qui correspond à un effet médié de 34% sur l'échelle logarithmique. Lorsque l'analyse de médiation a été stratifiée par le statut des récepteurs hormonaux, il est apparu que l'association entre le risque de THS et le CS était principalement due à des cancers du sein positifs aux récepteurs hormonaux : l'OR associé à l'utilisation actuelle du THS était de 1,81 (IC à 95%, 1,05 à 3,10) ; lorsque PDM a été ajouté au modèle, l'OR est devenu 1,46 (IC à 95%, 0,84 à 2,57), ce qui correspond à un effet médié de 36%. Pour les cancers du sein négatifs aux récepteurs hormonaux, l'association entre THS et CS n'était pas significative, que ce soit sans ou avec inclusion de PDM dans le modèle (OR = 0,64, IC à 95% 0,15 à 2,81 et OR = 0,57, IC à 95% 0,12 à 2,70, respectivement). Des résultats similaires ont été obtenus à partir des régressions totalement ajustées, où la proportion de l'effet de la THS médiée par la PDM était de 40% pour tout CS et de 41% pour les CS hormonaux positifs.

**Analyse de médiation de multiple.** La Table 4.2 présente les résultats du modèle de médiation qui expliquent l'effet de médiation joint de la PDM et de l'IMC (coefficient de corrélation de Pearson entre la PDM transformée à racine carrée et l'IMC - 0,39,  $p < 0,001$ ). Dans le modèle partiellement ajusté, l'OR associé à l'utilisation actuelle du THS était de 1,46 (IC à 95%, 0,42 à 2,41); l'effet direct moyen était de 1,37 (0,85 à 2,18); les effets médiés moyens étaient de 1,20 (IC à 95%, 1,06 à 1,41) pour la PDM et de 0,90 (IC à 95%, 0,78 à 0,98) pour l'IMC. Lorsque l'analyse de médiation a été stratifiée sur le statut des récepteurs hormonaux, il est apparu que l'association entre le risque de THS et de CS était entièrement due à des cancers du sein positifs aux récepteurs hormonaux : l'effet total du THS était de 1,58 (IC à 95%, 0,93 à 2,64), résultant d'un effet direct de 1,44 (IC à 95%, 0,85 à 2,34) et d'effets indirects de 1,22 (IC à 95%, 1,06 à 1,52) par PDM et 0,91 (IC à 95%, 0,78 à 0,99) par IMC. Selon ce modèle, la proportion d'effet médié par PDM sur l'échelle logarithmique était de 48% pour tous les CS et de 43% pour les récepteurs hormonaux positifs du CS. Aucun changement important n'a été observé dans les résultats de l'analyse de médiation lorsque les modèles ont été ajustés sur tous les facteurs de confusion potentiels (Table 4.2). Des résultats qualitativement similaires ont été observés pour l'utilisation de la THS codée de façon binaire *ancienne/ jamais* (Table 4.3).

	<b>Tous</b> NCA=196/Nco=196	<b>ER or PR positif</b> NCA=150/Nco=150	<b>ER and PR négatif</b> NCA=22/Nco=22
<b>Modèle partiellement ajusté*</b>			
Effet total	1.46 (0.92 to 2.41)	1.58 (0.93 to 2.64)	0.63 (0.11 to 2.61)
Effet direct	1.37 (0.85 to 2.18)	1.44 (0.85 to 2.34)	0.61 (0.10 to 2.48)
Effet médié joint	1.07 (0.92 to 1.25)	1.10 (0.93 to 1.37)	1.00 (0.48 to 2.52)
Effet médié par PMD	1.20 (1.06 to 1.41)	1.22 (1.06 to 1.52)	1.10 (0.61 to 2.63)
Effet médié par IMC	0.90 (0.78 to 0.98)	0.91 (0.78 to 0.99)	1.00 (0.49 to 1.40)
<b>Modèle totalement ajusté†</b>			
Effet total	1.49 (0.89 to 2.49)	1.53 (0.91 to 2.71)	0.79 (0.14 to 2.92)
Effet direct	1.34 (0.82 to 2.18)	1.36 (0.81 to 2.47)	0.63 (0.13 to 2.63)
Effet médié joint	1.09 (0.94 to 1.31)	1.11 (0.93 to 1.4)	1.18 (0.61 to 2.62)
Effet médié par PMD	1.23 (1.07 to 1.48)	1.26 (1.07 to 1.56)	1.20 (0.84 to 2.71)
Effet médié par IMC	0.89 (0.79 to 0.98)	0.9 (0.75 to 0.99)	1.00 (0.54 to 1.53)

\*Modèle ajusté sur : âge et THS à la mammographie ; état ménopausique au début de l'étude et année de naissance.

†Modèle ajusté sur : âge et THS à la mammographie ; état ménopausique, antécédents familiaux, parité, allaitement, utilisation de contraceptifs oraux ; année de naissance et âge de la ménarche.

TABLE 4.2 – Analyse de médiation évaluant l'effet de l'utilisation (actuelle contre jamais) de THS sur le risque de cancer du sein, globalement et par statut ER et PR en présence de PDM et IMC, médiateurs corrélés. ER : récepteur des œstrogènes ; PR : récepteur de la progestérone ; PDM : pourcentage de densité mammographique ; IMC : indice de masse corporelle ; THS : traitement hormonal de la ménopause.

#### 4.1. EFFET DE L'HORMONOTHÉRAPIE SUBSTITUTIVE SUR LE CANCER DU SEIN<sup>99</sup>

	<b>Tous</b>	<b>ER or PR positif</b>	<b>ER and PR négatif</b>
	<i>N<sub>CA</sub></i> =405/ <i>N<sub>Co</sub></i> =405	<i>N<sub>CA</sub></i> =304/ <i>N<sub>Co</sub></i> =304	<i>N<sub>CA</sub></i> =50/ <i>N<sub>Co</sub></i> =50
<b>Modèle partiellement ajusté*</b>			
Effet total	1.79 (1.26 to 2.58)	1.68 (1.08 to 2.73)	2.06 (0.85 to 5.34)
Effet direct	1.62 (1.15 to 2.30)	1.46 (0.94 to 2.34)	1.98 (0.78 to 5.23)
Effet médié joint	1.09 (1.00 to 1.22)	1.15 (1.01 to 1.35)	1.00 (0.81 to 1.44)
Effet médié par PMD	1.15 (1.06 to 1.27)	1.23 (1.10 to 1.44)	1.08 (0.93 to 1.48)
Effet médié par IMC	0.95 (0.89 to 1.00)	0.94 (0.85 to 1.00)	1.00 (0.78 to 1.07)
<b>Modèle totalement ajusté†</b>			
Effet total	1.79 (1.26 to 2.54)	1.60 (1.03 to 2.43)	2.16 (0.88 to 6.39)
Effet direct	1.62 (1.15 to 2.26)	1.40 (0.91 to 2.13)	1.98 (0.78 to 5.65)
Effet médié joint	1.10 (1.01 to 1.23)	1.13 (1.01 to 1.32)	1.05 (0.85 to 1.55)
Effet médié par PMD	1.15 (1.07 to 1.28)	1.22 (1.09 to 1.41)	1.06 (0.92 to 1.54)
Effet médié par IMC	0.95 (0.89 to 1.00)	0.94 (0.84 to 1.00)	1.00 (0.83 to 1.14)

\*Modèle ajusté sur: âge et THS à la mammographie; état ménopausique au début de l'étude et année de naissance.

†Modèle ajusté sur: âge et THS à la mammographie; état ménopausique, antécédents familiaux, parité, allaitement, utilisation de contraceptifs oraux; année de naissance et âge de la ménarche.

TABLE 4.3 – Analyse de médiation évaluant l'effet de l'utilisation (ancienne contre jamais) de THS sur le risque de cancer du sein, globalement et par statut ER et PR en présence de PDM et IMC, médiateurs corrélés. ER : récepteur des œstrogènes ; PR : récepteur de la progestérone ; PDM : pourcentage de densité mammographique ; IMC : indice de masse corporelle ; THS : traitement hormonal de la ménopause.

**Discussion** L'analyse des données collectées au sein de la cohorte prospective E3N a montré que les femmes ménopausées prenant du THS ont un PDM et un DA plus élevés et un NDA plus faible que celles qui n'ont jamais pris de THS et que cette différence est observée dès la première année d'utilisation du THS. L'arrêt de THS conduit à une diminution de PDM et DA dont les niveaux atteignent ceux de celles qui n'ont jamais utilisé de THS après plus de 8 ans depuis l'arrêt du traitement. Aucune hétérogénéité n'a été observée entre les types de THS : œstrogènes seuls et œstrogènes progestatifs. L'effet de l'utilisation actuelle du THS sur le risque du CS était partiellement direct et partiellement médié par PDM et IMC ; l'effet médié était entièrement limité aux tumeurs positives aux récepteurs hormonaux.

Dans la méta-analyse du Collaborative Group on Hormonal factors in Breast Cancer, le risque accru de CS associé au THS a été observé pour tous les types de THS à l'exception des œstrogènes vaginaux et le risque était plus élevé pour les œstrogènes plus progestatifs que pour les œstrogènes seuls ; de plus, le risque était plus élevé pour les CS positifs aux ER que pour les CS négatifs aux ER. Les analyses menées sur l'ensemble de la cohorte E3N ont identifié une hétérogénéité de l'effet sur le risque de CS du type de préparation de THS et ont fourni des preuves d'un effet différentiel sur le risque par sous-type de CS. Notre analyse actuelle n'a pas montré de preuve d'hétérogénéité de l'effet sur la DM par type de THS, bien que le nombre de femmes prenant uniquement des œstrogènes était trop petit pour fournir une puissance adéquate. En ce qui concerne le sous-type de cancer, en cohérence avec l'observation précédente sur le risque du CS nous avons observé que la PDM médie l'effet du THS sur le risque de CS uniquement pour les CS positifs aux récepteurs hormonaux, mais pas pour les CS négatifs aux récepteurs hormonaux.

## 4.2 Effets des tempéraments sur l'anorexie mentale

L'anorexie mentale (AM) est une maladie complexe et difficile à traiter qui, dans un certain nombre de cas, devient chronique et invalidante. C'est un trouble de cause inconnue qui survient le plus souvent chez les adolescentes. Les individus atteints d'anorexie mentale présentent une résistance égosyntonique<sup>(a)</sup> à l'alimentation conduisant à un poids corporel significativement faible, une peur intense de prendre du poids, des déformations de l'image corporelle et une perte de poids sévère en raison d'une recherche incessante de minceur et d'une alimentation restrictive. Les améliorations dans la compréhension et le traitement de l'anorexie mentale sont d'une importance capitale étant donné son évolution chronique et récurrente, entraînant souvent une morbidité médicale coûteuse et sévère. De surcroît, l'anorexie mentale a le taux de mortalité le plus élevé de toutes les maladies psychiatriques, également parce qu'il n'existe aucun traitement efficace reconnu pour inverser les symptômes.

Le traitement de l'anorexie mentale présente des défis uniques ; il n'existe actuellement

---

(a). intégrée dans la façon d'être de l'individu et difficilement identifiable par l'individu.

aucun médicament approuvé, et peu de psychothérapies sont fondées sur des preuves. En plus de traiter efficacement les troubles de l'humeur et les troubles cognitifs, il est essentiel de rétablir le poids. Les traitements sont difficiles à mettre en place car de nombreux facteurs, dont la personnalité, l'anxiété et la dépression, entrent en jeu.

En collaboration avec la chercheuse en psychiatrie Enrica Marzola, nous nous sommes intéressés à trois types de variables affectant l'anorexie mentale :

Les traits de personnalité (éviter les souffrances, émotivité négative, rigidité, perfectionnisme) sont un aspect clé de la pathogenèse des anorexies mentales, d'après la littérature soutenant leur implication à la fois pour le développement et le maintien de l'anorexie mentale (Atiye et al. (2015); Treasure et al. (2015)).

En revanche, peu d'attention a été accordée jusqu'à présent aux tempéraments affectifs (c'est-à-dire, dysthymiques / dépressifs, hyperthymiques, cyclothymiques, irritables, anxieux) impliqués dans l'anorexie mentale (Marzola et al. (2017); Ramacciotti et al. (2004)). Des travaux antérieurs (Marzola et al. (2017)) se sont concentrés sur les tempéraments affectifs dans l'anorexie mentale, trouvant une association entre les traits dépressifs et anxieux et les comorbidités à part entière. De plus, l'anxiété joue un rôle clé dans l'anorexie mentale pour de multiples raisons : les recherches cliniques et épidémiologiques ont constamment montré que les troubles anxieux sont fréquemment comorbides avec l'anorexie mentale (Godart et al. (2002); Hudson et al. (2007); Kendler (1995); Swinbourne et al. (2012); Walters EE (1995)).

La dépression est l'un des troubles comorbides les plus courants avec l'anorexie mentale (Abbate-Daga et al. (2011); Godart et al. (2007)) potentiellement influençant les performances neurocognitives des patientes (Amianto et al. (2015)). Plus encore que pour l'anxiété, le débat est ouvert quant à savoir si la dépression pourrait être un facteur prédisposant à l'anorexie mentale ou une conséquence de la malnutrition sévère liée à l'anorexie mentale (Mattar et al. (2011)). De plus, l'apparition de la dépression a tendance à précéder l'anorexie mentale (Deep et al. (1995)).

Compte tenu des lacunes actuelles de la littérature, avec notre méthode, nous avons cherché à étudier, les effets de médiation de l'anxiété, des tempéraments affectifs et de la dépression sur la psychopathologie alimentaire (à savoir recherche de la minceur et l'insatisfaction corporelle) et l'IMC chez les patientes atteintes d'anorexie mentale sévère.

**Présentation des données** Le jeu de données initiale était composé de 209 femmes hospitalisées consécutivement dans le service du Centre des troubles de l'alimentation de l'hôpital *Città della Salute e della Scienza* de Turin, à l'Université de Turin, en Italie. Toutes les patientes ont été admises volontairement. Cependant, 9 patientes ont été exclues de cette étude parce qu'elles ne répondaient pas aux critères d'inclusion, 4 candidates n'ont pas fourni de consentement écrit valide et 12 femmes ont renvoyé une évaluation incomplète. Par conséquent, le jeu de données finale contient 184 participantes atteintes d'anorexie mentale.

Les critères d'inclusion pour cette étude étaient les suivants : a) diagnostic complet d'anorexie mentale évalué par un psychiatre expérimenté avec l'entrevue clinique structurée pour le DSM-5 (évaluation des troubles mentaux) lors de l'admission ; b) sexe féminin ; c) aucune dépendance à une substance ; d) aucun trouble du spectre psychotique ; e) pas de comorbidités organiques (par exemple épilepsie). Le comité d'éthique du département de neurosciences de l'Université de Turin, en Italie, a approuvé cette étude et toutes les patientes ont donné leur consentement par écrit.

Sur l'ensemble de l'échantillon, 129 patientes (70,1%) ont été diagnostiquées avec l'anorexie mentale de type R-AM<sup>(b)</sup> et 55 (29,9%) avec de type B-PAM<sup>(c)</sup>. Les patientes étaient toutes de race blanche et leur âge moyen était de  $23,7 \pm 8,7$  ans, l'IMC moyen à l'admission était de  $14,4 \pm 1,8$  et la durée moyenne de la maladie était de  $6 \pm 7,8$  ans. Soixante patientes (39,2%) ont dû être hospitalisées dans le passé et 102 (66,7%) prenaient des médicaments lors de leur admission. La durée moyenne d'hospitalisation était de  $36,6 \pm 18,1$  jours et 72 personnes (47,4%) ont déclaré leur IMC comme étant le plus bas de leur vie à l'admission. Concernant la comorbidité, 61 patientes (40,1%) avaient un trouble psychiatrique comorbide (c'est-à-dire un diagnostic formel de dépression ou d'anxiété).

**Modèles de médiation** Dans nos modèles de médiation on considère les variables suivantes :

En *exposition* on considère le tempérament avec 5 variables : le tempérament dysthymique/dépressif<sup>(d)</sup>, le tempérament cyclothymique<sup>(e)</sup>, le tempérament hyperthymique/maniaque<sup>(f)</sup>, le tempérament irritable<sup>(g)</sup> et le tempérament anxieux<sup>(h)</sup>. Ces variables ont été obtenues à l'aide de l'auto-questionnaire TEMPS-A (Pompili et al. (2008)). C'est un questionnaire à 110 questions (109 pour les hommes) fermées. Il évalue les tempéraments dysthymique (questions 1–22), cyclothymique (questions 23–42), hyperthymique (questions 43–63), irritable (questions 64–84) et anxieux (questions 85–110).

En tant que *variables médiatrices* on considère les 3 suivantes : l'anxiété en tant que trait de personnalité (anxiété T), l'anxiété en tant qu'état émotionnel (anxiété S) et la dépression clinique. Les variables anxiété T et anxiété S ont été obtenues avec l'Inventaire d'anxiété État-Trait STAI (Spielberger (1983)), un instrument psychométrique largement

---

(b). perte de poids uniquement en suivant un régime et en faisant de l'exercice sans frénésie alimentaire ni purge.

(c). perte de poids uniquement en suivant un régime et en faisant de l'exercice, mais avec frénésie alimentaire et/ou purge de façon périodique.

(d). Quelqu'un qui a tendance à être légèrement déprimée et a une grande sensibilité au stress et à ce qui est émotionnel. Vues de l'extérieur, ces personnes sont souvent sombres et rarement joyeuses.

(e). Quelqu'un qui a des fluctuations d'humeur et d'énergie. Elle oscille entre le pôle maniaque et le pôle dépressif.

(f). Quelqu'un d'actif, plein de projets, toujours de bonne humeur, confiant en lui, très peu fatigué.

(g). Quelqu'un qui se plaint souvent, ronchonne, supporte difficilement la contrariété, la frustration.

(h). Quelqu'un qui a une vision du monde marquée par l'anxiété, prévoyant toujours le pire, de manière consciente ou non.

utilisé qui permet de distinguer deux dimensions différentes de l'anxiété : l'anxiété en tant que trait de personnalité (sentiments d'appréhension, de tension, de nervosité et d'inquiétude que le sujet ressent habituellement) et l'anxiété en tant qu'état émotionnel lié à une situation particulière (sentiments d'appréhension, de tension, de nervosité et d'inquiétude que le sujet ressent à un moment précis). Ces deux types d'anxiété sont mesurés par deux échelles d'auto-évaluation différentes de 20 questions (formulaire X pour l'anxiété S et formulaire Y pour l'anxiété T). Pour chaque élément, la gravité des symptômes anxieux est notée de 1 (pas du tout / presque jamais) à 4 (beaucoup / presque toujours) et un score global de 20 à 80 est évalué.

La sévérité de la dépression clinique est mesurée par l'Inventaire de dépression de Beck (BDI) (Beck et al. (1961)), un auto-questionnaire à choix multiples de 13 questions définissant une échelle allant des symptômes dépressifs minimaux (scores de 0 à 4) à la dépression sévère (scores de 16 à 39). Des scores légers (scores de 5 à 7) et modérés (scores de 8 à 15) sont également identifiables.

En *variable de réponse*, on considère la psychopathologie alimentaire à travers 2 variables : le désir intense de minceur et l'insatisfaction corporelle. On considère aussi une troisième variable de réponse avec l'indice de masse corporelle (IMC). Le désir intense de la minceur et l'insatisfaction corporelle sont données par l'Inventaire des troubles de l'alimentation-2 (EDI-2) (Garner (1991)), une mesure d'auto-évaluation psychométriquement robuste (Thiel and Paul (2006)) des troubles de l'alimentation, des comportements et des traits de personnalité communs aux personnes qui reçoivent un diagnostic de trouble de l'alimentation. Onze sous-échelles évaluent les symptômes et les corrélats psychologiques des troubles de l'alimentation ; des scores élevés reflètent une pathologie plus grave.

Nous sommes en présence de 5 expositions potentielles et 3 variables de réponse. Nous avons donc conduit 15 analyses de médiations multiples avec trois médiateurs. Dans la Figure 4.3 on illustre les deux diagrammes causaux étudiés avec nos analyses de médiation. Notons que les analyses sont menées en supposant ces diagrammes causaux mais qu'ils ne peuvent coexister compte tenu du fait que l'IMC est considéré d'une part comme une covariable affectant les médiateurs et d'autre part comme une variable de réponse affectée par les médiateurs et que nous faisons l'hypothèse qu'il n'y a pas de corrélation entre les médiateurs.

**Analyse de médiation de multiple.** Comme indiqué dans la Table 4.4, tous les tempéraments affectifs à l'exception du tempérament maniaque ont eu un effet total significatif sur la psychopathologie alimentaire et l'IMC. Concernant l'effet direct, seuls deux tempéraments affectifs, à savoir dépressif et anxieux, ont montré un impact direct sur le désir intense de minceur et l'IMC. En ce qui concerne les effets médiés, nous avons trouvé un effet médié joint effet significatif pour toutes les expositions et les variables de réponse



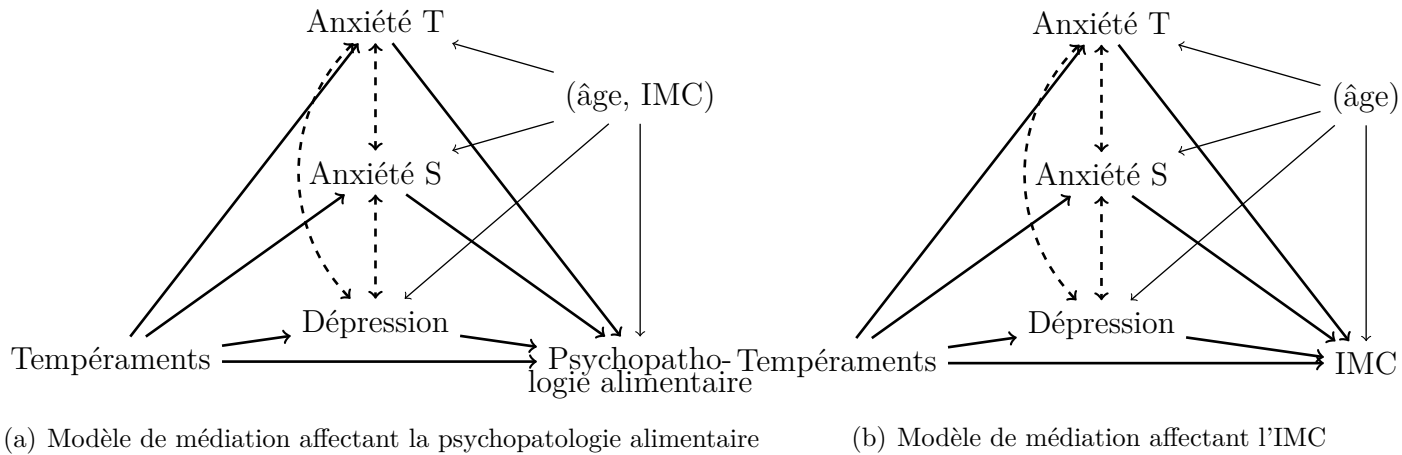


FIGURE 4.3 – Diagrammes causaux des analyses de médiation multiple.

considérés, à l'exception de la relation entre le tempérament dépressif et l'IMC. En considérant les médiateurs de façon individuelle, l'anxiété T n'a pas montré d'effet significatif, tandis que l'anxiété S et la dépression étaient des médiateurs significatifs du désir intense de minceur et d'insatisfaction corporelle. La dépression était le seul médiateur significatif affectant l'IMC pour tous les tempéraments considérés.

**Discussion** Avec cette analyse, nous avons cherché à étudier la relation complexe qui existe entre les tempéraments affectifs, l'anxiété et la dépression sur les aspects fondamentaux de la psychopathologie (c.-à-d., désir intense de minceur et insatisfaction corporelle) et clinique (c.-à-d. IMC) chez des patientes atteintes d'anorexie mentale sévère. Prises ensemble, les données confirment le rôle clé des tempéraments affectifs dans l'anorexie mentale et suggèrent que les tempéraments dépressifs et anxieux ont le plus fort effet direct sur les variables de réponse considérées. Plus particulièrement, l'anxiété d'état et la dépression clinique influencent à la fois le désir intense de minceur et l'insatisfaction corporelle, tandis que l'IMC est affecté indirectement (de façon médiée) uniquement par des symptômes dépressifs. Enfin, selon notre modèle, l'anxiété de trait n'était pas un médiateur significatif. Notons que la dépression et l'anxiété sont si étroitement liées (et donc corrélées) que seule une méthode capable de considérer cette corrélation, telle que la nôtre, pourrait être utile pour étudier leur rôle de médiation.

Ces résultats sont cohérents avec les résultats antérieurs montrant les tempéraments affectifs, principalement anxieux et dépressifs, comme impliqués dans l'anorexie mentale (Marzola et al. (2017); Marzola et al. (2020)). De même, selon les analyses de médiation, les tempéraments anxieux et dépressifs étaient les plus solidement liés à la fois à la psychopathologie de l'alimentation et à l'IMC et cela représente une découverte nouvelle dans le domaine de l'anorexie mentale puisqu'aucune analyse de médiation sur ces questions n'a été menée jusqu'à présent. Fait intéressant, le tempérament maniaque n'a pas

	Désir intense de minceur <sup>§</sup>			Insatisfaction corporelle <sup>§</sup>			IMC*		
	Estimation	95%IC	p	Estimation	95%IC	p	Estimation	95%IC	p
<b>Tempérament</b>									
<b>Dépressif</b>									
Effet total	1.04	.81,1.27	<.001	.68	.45,.92	<.001	.15	.08,.21	<.001
Effets médiés									
Anxiété T	0.06	-.15,.24	.558	.03	-.16,.23	.750	.01	-.04,.08	.640
Anxiété S	.41	.24,.61	<.001	.30	.14,.50	<.001	-.02	-.07,.03	.462
Dépression	.16	-.02,.33	.078	.21	.04,.41	.008	.06	-.003,.12	.062
JOINT	.62	.42,.82	<.001	.54	.33,.76	<.001	.05	-.009,.11	.086
Effet direct	.42	.16,.69	.004	.14	-.16,.40	.338	.09	.009,.18	.032
<b>Cyclothymique</b>									
Effet total	.67	.45,.86	<.001	.40	.21,.61	<.001	.10	.05,.16	<.001
Effets médiés									
Anxiété T	.08	-.02,.18	.160	.03	-.07,.14	.520	.02	-.01,.06	.296
Anxiété S	.28	.16,.43	<.001	.21	.09,.36	<.001	-.02	-.05,.02	.388
Dépression	.12	.01,.23	.024	.15	.03,.28	.010	.04	.002,.09	.044
JOINT	.47	.34,.61	<.001	.39	.26,.54	<.001	.05	.009,.08	.014
Effet direct	.19	-.01,.37	.062	.009	-.19,.22	.938	.06	-.001,.12	.050
<b>Maniaque</b>									
Effet total	-.26	-.50,-.05	.018	-.2	-.04,.03	.098	-.05	-.11,.02	.164
Effets médiés									
Anxiété T	-.08	-.19,-.005	.072	-.03	-.13,.06	.482	-.02	-.06,.01	.202
Anxiété S	-.14	-.27,-.03	.010	-.1	-.2,-.02	.012	.005	-.01,.03	.586
Dépression	-.11	-.23,-.03	.014	-.12	-.25,-.02	.014	-.04	-.08,-.01	.016
JOINT	-.33	-.47,-.2	<.001	-.26	-.38,-.13	.002	-.05	-.09,-.02	<.001
Effet direct	.07	-.13,.25	.466	.06	-.14,.26	.538	.008	-.05,.07	.802
<b>Irritable</b>									
Effet Total	.78	.54,.10	<.001	.55	.32,.80	<.001	.12	.05,.19	<.001
Effets médiés									
Anxiété T	.11	-.03,.25	.124	.04	-.1,.2	.604	.02	-.02,.08	.324
Anxiété S	.37	.21,.57	<.001	.27	.12,.45	<.001	-.02	-.07,.03	.460
Dépression	.16	.02,.31	.026	.19	.05,.36	.006	.06	.002,.11	.036
JOINT	.65	.47,.84	<.001	.51	.35,.7	<.001	.06	.01,.11	.008
Effet direct	.13	-.11,.36	.282	.04	-.21,.28	.756	.06	-.02,.14	.134
<b>Anxieux</b>									
Effet total	.72	.56,.88	<.001	.55	.37,.72	<.001	.11	.06,.15	<.001
Effets médiés									
Anxiété T	.04	-.1,.18	.606	-.008	-.17,.16	.890	.005	-.04,.05	.814
Anxiété S	.29	.16,.42	<.001	.21	.08,.35	<.001	-.02	.05,.01	.346
Dépression	.12	.01,.23	.044	.14	.01,.27	.034	.04	.004,.08	.020
JOINT	.45	.31,.6	<.001	.34	.19,.49	<.001	.03	-.01,.07	.158
Effet direct	.28	.08,.46	.004	.20	-.01,.42	.058	.08	.02,.13	.014

<sup>§</sup> Modèle ajusté sur l'IMC et l'âge.

\* Modèle ajusté sur l'âge.

TABLE 4.4 – Analyse de médiation évaluant les effets des tempéraments affectifs sur la psychopathologie alimentaire et sur l'IMC à travers l'anxiété de trait (anxiété T), l'anxiété d'état (anxiété S) et la dépression clinique.

montré d'effet direct sur les variables de réponse considérées. Cela pourrait être dû au fait que, selon des travaux antérieurs (Marzola et al. (2017)), le tempérament maniaque aurait été sous-représenté dans la population atteinte d'anorexie mentale, remettant ainsi potentiellement en question sa pertinence dans les troubles de la physiopathologie.

Il est intéressant de noter que le modèle proposé a permis de quantifier une notion clinique de tous les jours, car l'anxiété T et la dépression ont été jugées significativement médiatrices de l'effet sur le désir intense de minceur, l'insatisfaction corporelle et l'IMC. De plus, dans le cadre du débat scientifique complexe sur l'anxiété dans l'anorexie mentale, nos résultats sont conformes à la littérature antérieure suggérant que la malnutrition sévère exacerbe les aspects psychopathologiques fondamentaux de l'alimentation des patientes hospitalisées atteintes d'anorexie mentale (Kaye et al. (2009)). Effectivement, l'anxiété a considérablement contribué au désir de minceur, à l'insatisfaction corporelle et à l'IMC.

Lors des analyses de médiation sur l'IMC - un aspect critique de l'anorexie, principalement lorsqu'une hospitalisation d'urgence est nécessaire pour un cas extrême (c'est-à-dire un IMC <15, comme dans notre échantillon) - un rôle majeur des symptômes dépressifs comme médiateur est apparu. En effet, la dépression médie les effets de tous les tempéraments à l'exception du tempérament dépressif. Par conséquent, il est possible que le tempérament dépressif joue un rôle clé dans son effet direct sur l'IMC, en contournant en quelque sorte la symptomatologie dépressive des patientes. Cependant, nos résultats sont cohérents avec des études proposant la dépression comme ayant un impact sur l'insatisfaction corporelle (Paans et al. (2018)) bien qu'une analyse de médiation faisait à ce jour défaut.

Il est à noter que l'anxiété T ne s'est pas révélée être un médiateur pour le désir intense de minceur, l'insatisfaction corporelle et l'IMC. D'une part, il s'agit d'un résultat inattendu, étant donné la pertinence de l'anxiété dans l'anorexie mentale (Swinbourne et al. (2012); Zerwas et al. (2013); Kezelman et al. (2015)), mais d'autre part, il pourrait être également vrai que - parmi les traits - la contribution du tempérament anxieux au modèle proposé pourrait être suffisamment robuste pour «couvrir» celle de l'anxiété de trait. Compte tenu de la nouveauté de ce résultat, des études supplémentaires sont nécessaires pour confirmer ces découvertes intrigantes.

Si ces résultats ainsi que les DAG causaux sous-jacents à nos analyses sont confirmés, des implications cliniques intéressantes pourraient être soulevées, d'après Enrica Marzola et son équipe. En effet, nos données suggèrent que - lors du traitement de personnes souffrant d'anorexie mentale dans une phase sévère de leur trouble - une attention particulière devrait être accordée aux composantes anxieuses et dépressives de leur psychopathologie. Compte tenu de la stabilité des traits de caractère, les interventions thérapeutiques peuvent vouloir s'attaquer aux composants anxieux et dépressifs du traitement. Par exemple, bien qu'aucun médicament ne puisse inverser les principaux symptômes de l'anorexie mentale, la pharmacothérapie pourrait être utile à cet égard (Marzola et al. (2015); Lloyd et al. (2019)). De plus, étant donné la gravité de certaines patientes qui ont tendance à développer un trouble sévère et durable, des stratégies de neuromodula-

tion pourraient également être intéressantes, car les techniques non invasives sont connues pour améliorer les symptômes spécifiquement dépressifs. Malgré les résultats très intéressants que nous avons trouvé avec notre méthode d'analyse de médiation multiple et le recrutement de patientes *du monde réel* atteints d'anorexie mentale sévère et les implications cliniques potentielles, cette étude contient également des limites : les données ne contiennent que des patientes hospitalisées, ce qui pourrait entraver la généralisation des données, le nombre de patientes pourrait être plus grand et des auto-évaluations ont été utilisées.



# Conclusion

## Médiateurs liés de façon non-causale

Le Chapitre 1 aborde le problème de l'estimation des effets directs et indirects, y compris les effets indirects à travers les médiateurs individuels, dans le cadre d'une analyse de médiation multiple avec des médiateurs liés de manière non causale. Les travaux théoriques de Shpitser et de ses coauteurs ont prouvé qu'en présence de variables latentes, toutes les quantités de médiation ne sont pas identifiées (Shpitser (2013), Shpitser and Sherman (2018)). En particulier, en présence d'une cause commune latente entre les médiateurs, les effets indirects à travers les médiateurs individuels ne peuvent pas être exprimés uniquement en fonction des données observables. En revanche, une pratique courante en médiation multiple consiste à effectuer plusieurs analyses de médiation simples, une pour chaque médiateur, malgré l'introduction d'un biais.

La plupart des approches de l'analyse de médiation reposent sur des hypothèses fortes telles que l'Ignorabilité Séquentielle (Imai et al. (2010a), Hafeman and VanderWeele (2011)), et plusieurs auteurs ont tenté de résoudre le problème de l'estimation par différentes techniques.

Dans le cadre de la médiation multiple avec des médiateurs liés de manière non causale, nous définissons un ensemble d'hypothèses, appelé SIMMA, dans lequel nous exprimons l'effet direct et l'effet indirect joint en tant que fonctions de variables observées et l'effet indirect par l'intermédiaire de médiateurs individuels en termes de variables à la fois observées et contrefactuelles. Couplée à un choix de modèle et à l'algorithme quasi-bayésien développé par Imai et al. (2010a), cette dernière formule donne une méthode d'estimation des effets indirects individuels. Notons que nous sommes limités aux modèles avec l'hypothèse supplémentaire que la corrélation entre médiateurs contrefactuels est la même quel que soit le traitement qui les régit. Le développement de méthodes pour faire face à la situation dans laquelle cette hypothèse supplémentaire est violée est laissé à des travaux futurs, ainsi que le développement d'une analyse de sensibilité pour évaluer la robustesse aux écarts par rapport à SIMMA.

Nous avons appliqué notre programme R pour valider empiriquement la méthode proposée. Cette étude de simulation montre que notre méthode fournit une estimation non

biaisée de l'effet direct, alors que, comme prévu, les estimations obtenues en exécutant des analyses de médiation simples avec un médiateur à la fois sont biaisées, même dans le cas de médiateurs indépendants. De plus, lorsque les médiateurs partagent une cause commune non observée, nous montrons que notre analyse multiple fournit des estimations des effets directs à travers des médiateurs individuels qui sont moins biaisées que celles obtenues à partir d'analyses simples avec un médiateur à la fois. La raison de cette amélioration, est que notre méthode, en considérant la loi commune des médiateurs conditionnellement au traitement et la loi de la variable de réponse conditionnellement à tous les médiateurs, prend automatiquement en compte l'influence de la cause commune non observée  $U$  sur les médiateurs et la variable de réponse. Au contraire, faire une analyse simple avec un médiateur à la fois n'est pas approprié dans ce contexte car  $U$  confond la relation entre chaque médiateur et la variable de réponse. De plus, nous montrons empiriquement que, contrairement aux analyses simples répétées, l'algorithme quasi-bayésien proposé fournit des intervalles de confiance avec la propriété de couverture attendue.

Les analyses simples répétées des médiateurs individuels sont encore une approche populaire malgré une littérature croissante mettant en garde contre ses limites. En effet, la présence d'une cause commune non observée pour les médiateurs n'est pas la seule situation dans laquelle une telle approche pose problème. VanderWeele and Vansteelandt (2014) a observé que, même lorsque les médiateurs sont liés de manière non causale, il n'est pas possible de décomposer l'effet indirect conjoint en somme des effets indirects individuels lorsque leur effet sur la variable de réponse est caractérisé par une interaction dans l'échelle additive, une situation que nous avons exclue dans nos résultats théoriques. Dans cette situation, Taguri et al. (2015) a fourni une décomposition à trois voies de l'effet indirect conjoint en effets indirects naturels individuels et en un effet interactif. Fait intéressant, les hypothèses requises pour montrer l'identifiabilité de tous les termes de cette décomposition sont les mêmes que les nôtres, à la seule différence importante que les médiateurs potentiels sont supposés conditionnellement indépendants compte tenu de toutes les covariables observées. Plus récemment, Bellavia and Valeri (2017) ont fourni une décomposition de l'effet total dans la situation plus générale avec à la fois des interactions médiateur-médiateur et médiateurs-variable de réponse.

Une autre situation dans laquelle la répétition d'analyses simples est la mauvaise approche de l'analyse médiation multiple est celui où les médiateurs sont ordonnés de manière causale comme sur la figure 6(c). Dans cette situation, considérer le vecteur des variables intermédiaires comme un médiateur et effectuer une analyse simple permettra d'estimer correctement l'effet indirect conjoint et l'effet direct. Cependant, l'ancien effet indirect conjoint n'est pas égal à la somme de l'effet indirect individuel, chacun estimé avec une analyse simple, car certains chemins sont comptés deux fois et l'effet médié par  $W$  est biaisé par  $M$  qui agit comme un facteur de confusion post-traitement de la relation  $W - Y$ . Plus généralement, à moins que des conditions fortes ne soient vérifiées, il n'est pas possible d'identifier tous les chemins spécifiques (Avin et al. (2005), Daniel et al. (2015)). VanderWeele and Vansteelandt (2014) ont introduit une approche séquentielle

pour identifier l'effet indirect conjoint, l'effet direct, l'effet médié par  $M$  et l'effet médié par  $W$  mais pas par  $M$ . Les différentes étapes de cette stratégie peuvent être implémentées en utilisant `medflex`, un package R récemment introduit basé sur le *Natural Effect Model* et des méthodes d'imputation ou de pondération (Steen et al. (2017)). Une approche alternative basée sur des équations structurelles linéaires avec des coefficients variables a été discutée par Imai and Yamamoto (2013) et implémentée dans le package `mediation`. Nguyen et al. (2015) ont présenté une méthode basée sur l'approche de pondération du rapport de cotes inverse (IOWR) introduite par Tchetgen Tchetgen (2013). Cette méthode est très flexible car elle intègre des modèles linéaires généralisés, des modèles de régression quantile et de survie pour la variable de sortie et de multiples médiateurs continus ou catégoriels, mais elle ne permet pas d'estimer l'effet indirect via des médiateurs individuels, mais uniquement l'effet indirect joint.

Nous concluons ce bref aperçu de la littérature autour de la médiation multiple en soulignant que notre cadre traite des effets directs et indirects *naturels*. Vansteelandt and Daniel (2017) ont récemment introduit des effets indirects directs et spécifiques au chemin d'intérêt, dits *interventionnels* qui s'ajoutent à l'effet total et sont identifiables même lorsque les médiateurs partagent des causes communes non mesurées ou que la dépendance causale entre les médiateurs est inconnue.

Un travail futur pourrait inclure des méthodes d'analyses de sensibilité pour les effets d'intérêts.

## Grande dimension

Le Chapitre 2 aborde les questions de classification et de sélection de médiateurs qui se posent quand le nombre de médiateurs augmente.

Nous avons développé la classification de médiateurs par rapport à leur corrélation qui peut être causale ou non. Cette méthode basée sur la classification spectrale permet d'aboutir à l'analyse de médiation par bloc permettant de réduire le nombre de paramètres à considérer en présence d'un nombre relativement moyen de médiateurs corrélés.

Nous avons également développé une nouvelle méthode de sélection de médiateurs en grande dimension. Pour cela nous avons adapté et utilisé la méthode Group-LASSO suivie d'un test multiple. Une des particularités de notre méthode est qu'elle se base sur les effets médiés à chaque étape pour la sélection. De plus notre méthode a été développée dans le but de sélectionner des médiateurs communs à plusieurs traitements affectant une variable de réponse. Nous avons illustré à travers des simulations que nous obtenons de meilleurs résultats que les méthodes récentes HIMA, HDMA et CMF.

Une piste à explorer pourrait consister en la combinaison de nos méthodes de classification et de sélection de médiateurs. On pourrait ainsi sélectionner des médiateurs corrélés. De plus en présence de plusieurs traitements, la combinaison de nos deux méthodes permettrait de déterminer des groupes de médiateurs communs à l'ensemble des



traitements et qui sont corrélés de façon causale ou non.

## Package R

Le Chapitre 3 présente notre package R `multimediate` où l'ensemble de nos méthodes (analyse de médiation en présence de médiateurs corrélés de façon non-causale, classification de médiateurs, sélection de médiateurs en grande dimension) a été implémenté. Actuellement, notre package permet de travailler avec des modèles paramétriques avec des médiateurs et des variables de réponse continues, binaires, catégorielles ordonnées. Ce package a été publié sur Github et nous travaillons à l'extension du cadre actuel à d'autres types de modèles (par exemple pour des variables de survie et catégorielles non-ordonnées).

## Application à des données réelles

Dans le Chapitre 4, nous avons mené, dans un premier temps, une analyse de médiation multiple sur un ensemble de données réelles d'une grande cohorte pour évaluer l'effet du traitement hormonal substitutif (THS) sur le risque de cancer du sein à travers deux médiateurs non séquentiels, à savoir le pourcentage de zones mammaires denses et l'indice de masse corporelle. Nous avons aussi présenté les effets à l'échelle des OR. Les résultats que nous avons obtenus sont cohérents avec la littérature existante sur les relations causales entre les différentes variables. Au cours de la première année d'utilisation, la densité mammaire était plus élevée chez les utilisatrices actuelles que chez les femmes qui n'ont jamais utilisé de THS ; chez les anciennes utilisatrices, les niveaux ont atteint ceux des femmes qui n'ont jamais utilisé de THS après plus de 8 ans depuis l'arrêt de la thérapie.

Nous avons réalisé une analyse de médiation simple avec la densité mammaire comme médiateur et une analyse de médiation multiple avec la densité mammaire et l'IMC comme médiateurs. De l'analyse simple à l'analyse multiple la proportion médiée par la densité mammaire, passe de 34% à 48%. L'analyse de médiation multiple a permis de montrer que la densité mammaire médie en réalité jusqu'à 50% de l'effet du THS sur le risque de cancer du sein. Nos résultats, s'ils sont confirmés par des études longitudinales indépendantes, indiqueraient que le THS doit être prescrit avec prudence, en particulier chez les femmes ayant une DM élevée et suggèrent que la surveillance de la densité mammaire pendant la THS pourrait être une stratégie utile dans les situations où la prescription de THS est appropriée.

Dans un deuxième temps nous avons réalisé une analyse de médiation multiple pour évaluer les effets des tempérament affectifs sur l'anorexie mentale. Les modèles proposés de médiation par l'anxiété et par la dépression entre les tempéraments affectifs et la psychopathologie alimentaire et l'IMC dans l'anorexie mentale sont basés sur la pratique clinique quotidienne d'un point de vue psychiatrique. En fait, les cliniciens et les

chercheurs expérimentés qui travaillent avec des personnes souffrant d'anorexie mentale savent bien comment les caractéristiques prémorbides (c.-à-d. la structure du tempérament) et les états cliniques (c.-à-d. l'anxiété et la dépression) peuvent avoir un impact sur les symptômes de base de l'anorexie mentale. Cependant, à ce jour, ce concept n'avait pas été mesuré. Par conséquent, s'ils sont confirmées par d'autres études, nos résultats contribuent au débat scientifique en cours visant à faire la lumière sur cette maladie complexe, fournissant également des implications cliniques innovantes.

Dans ces deux travaux, nous avons appliqué notre méthode d'analyse de médiation en présence de plusieurs médiateurs corrélés. Mais il serait intéressant, dans nos applications futures, d'utiliser notre méthode de sélection de médiateur en grande dimension en présence de plusieurs traitements.



# Annexes

## Sommaire

---

<b>A</b>	<b>Médiateurs liés de façon non-causale . . . . .</b>	<b>115</b>
A.1	Lien entre $\delta^Z$ et $\sum_k \delta^k$ . . . . .	115
A.2	Hypothèses . . . . .	116
A.3	Preuve du Théorème 1.4.1 . . . . .	117
A.4	Preuve du Corollaire 1.5.1 . . . . .	120
A.5	Preuve du Corollaire 1.6.1 . . . . .	125
A.6	Modèles de simulation . . . . .	132
A.7	Résultats complémentaires . . . . .	135
<b>B</b>	<b>Grande Dimension . . . . .</b>	<b>142</b>
B.1	Propriétés du graphe laplacien non normalisé . . . . .	142
B.2	MAHI en présence de plusieurs traitements . . . . .	142
B.3	Méthode HIMA . . . . .	144
B.4	Méthode CMF . . . . .	147
<b>C</b>	<b>Package R . . . . .</b>	<b>149</b>
C.1	P-valeur . . . . .	149
C.2	Médiateurs catégoriels ordonnés . . . . .	150
C.3	Médiateurs linéaire et catégoriels ordonnés . . . . .	152

---

## A Médiateurs liés de façon non-causale

### A.1 Lien entre $\delta^Z$ et $\sum_k \delta^k$

Même si intuitivement, il semblerait raisonnable de penser que l'effet indirect via le  $k$ ème médiateur  $\delta^k$  est la différence entre l'effet joint  $\delta^Z$  et l'effet indirect de tous les autres médiateurs  $\eta^k$ , nous montrons que ce n'est pas vrai en général.

Nous voulons exprimer  $\delta^Z$  en fonction de  $\sum_{k=1}^K \delta^k$ . Pour ce faire, on part de  $\delta^k$  :

$$\begin{aligned}
\delta^k(t) &= E[Y(t, M^k(1), W^k(t)) - Y(t, M^k(0), W^k(t))] \\
&= \begin{cases} E[Y(1, Z(1)) - Y(1, M^k(0), W^k(1))] & \text{si } t = 1 \\ E[Y(0, M^k(1), W^k(0)) - Y(0, Z(0))] & \text{si } t = 0 \end{cases} \\
&= \begin{cases} E[\tau + Y(0, Z(0)) - Y(1, M^k(0), W^k(1))] & \text{si } t = 1 \\ E[Y(0, M^k(1), W^k(0)) + \tau - Y(1, Z(1))] & \text{si } t = 0 \end{cases} \\
&= \begin{cases} E[\tau + Y(1, Z(0)) - \zeta(0) - Y(1, M^k(0), W^k(1))] & \text{si } t = 1 \\ E[Y(0, M^k(1), W^k(0)) + \tau - \zeta(1) - Y(0, Z(1))] & \text{si } t = 0 \end{cases} \\
&= \begin{cases} E[\delta^Z(1) - Y(1, M^k(0), W^k(1)) + Y(1, Z(0))] & \text{si } t = 1 \quad (i) \\ E[\delta^Z(0) - Y(0, Z(1)) + Y(0, M^k(1), W^k(0))] & \text{si } t = 0 \end{cases} \\
&= \delta^Z(t) - \eta^k(t).
\end{aligned}$$

$\eta^k$  peut être interprété comme l'effet indirect par tous les médiateurs sauf le  $k$ ième, lorsque le traitement est fixé à  $t$  et que le  $k$ ième médiateur est fixé à la valeur qu'il aurait sous le traitement  $1-t$ . En sommant sur les  $K$  médiateurs, nous avons :

$$\begin{aligned}
\sum_{k=1}^K \delta^k(t) &= \sum_{k=1}^K (\delta^Z(t) - \eta^k(t)) \\
&= K\delta^Z(t) - \sum_{k=1}^K \eta^k(t)
\end{aligned}$$

Ainsi l'effet indirect joint peut être réécrit comme :

$$\delta^Z(t) = \frac{\sum_{k=1}^K (\delta^k(t) + \eta^k(t))}{K}.$$

## A.2 Hypothèses

Selon Imai and Yamamoto (2013), les Hypothèses d'Ignorabilité Séquentielle dans la situation de plusieurs médiateurs sans relation causale sont :

---

(i). En effet  $\tau = \delta^Z(t) + \zeta(1-t)$ .

$$\{Y(t, m, w), M(t'), W(t'')\} \perp\!\!\!\perp T|X = x, \quad (\text{A.1})$$

$$Y(t', m, W(t')) \perp\!\!\!\perp M(t)|T = t, X = x, \quad (\text{A.2})$$

$$Y(t', M(t'), w) \perp\!\!\!\perp W(t)|T = t, X = x, \quad (\text{A.3})$$

où  $\mathbb{P}(T = t|X = x) > 0$  et  $\mathbb{P}(M = m, W = w|T = t, X = x) > 0$  pour tout  $x, t, t', m, w$ .

Nous remplaçons les hypothèses (A.2) et (A.3) avec les hypothèses suivantes :

$$Y(t', m, w) \perp\!\!\!\perp (M(t), W(t))|T = t, X = x \quad (\text{A.4})$$

$$Y(t, m, w) \perp\!\!\!\perp (M(t'), W(t))|T = t, X = x \quad (\text{A.5})$$

de façon à obtenir les **Sequential Ignorability for Multiple Mediators Assumption (SIMMA)** :

$$(Y(t, m, w), M(t'), W(t'')) \perp\!\!\!\perp T|X = x, \quad (\text{A.1})$$

$$Y(t', m, w) \perp\!\!\!\perp (M(t), W(t))|T = t, X = x \quad (\text{A.4})$$

$$Y(t, m, w) \perp\!\!\!\perp (M(t'), W(t))|T = t, X = x \quad (\text{A.5})$$

Il est important de souligner que (A.4) n'implique pas (A.2) et (A.3) : une conséquence de (A.4) est que  $Y(t', m, w) \perp\!\!\!\perp M(t)|T = t, X = x$  pour toutes les valeurs de  $t', m, w$  et  $x$  mais cela n'implique pas que  $Y(t', m, W(t')) \perp\!\!\!\perp M(t)|T = t, X = x$  car  $Y(t', m, w)$  ne dépend que des causes résiduelles de  $Y$  après avoir défini  $T = t', M = m$  et  $W = w$  tandis que  $Y(t', m, W(t'))$  dépend également des causes résiduelles de  $W$  après avoir mis  $T$  à  $t'$ .

### A.3 Preuve du Théorème 1.4.1

#### Effet médié joint, effet direct et effet total

Afin de démontrer le théorème 1.4.1 pour l'effet indirect joint  $\delta^Z$ , l'effet direct  $\zeta$  et l'effet total  $\tau$ , nous commençons par réécrire les définitions en termes de contrefactuels :

$$\begin{aligned} \delta^Z(t) &= \mathbb{E}[Y(t, Z(1))] - \mathbb{E}[Y(t, Z(0))] \\ &= \int \mathbb{E}[Y(t, Z(1))|X = x] - \mathbb{E}[Y(t, Z(0))|X = x] dF_X(x) \\ &= \int \mathbb{E}[Y(t, M(1), W(1))|X = x] - \mathbb{E}[Y(t, M(0), W(0))|X = x] dF_X(x) \end{aligned}$$

$$\begin{aligned} \zeta(t, \dots, t) &= \mathbb{E}[Y(1, Z(t))] - \mathbb{E}[Y(0, Z(t))] \\ &= \int \mathbb{E}[Y(1, Z(t))|X = x] - \mathbb{E}[Y(0, Z(t))|X = x] dF_X(x). \\ &= \int \mathbb{E}[Y(1, M(t), W(t))|X = x] - \mathbb{E}[Y(0, M(t), W(t))|X = x] dF_X(x) \end{aligned}$$

$$\begin{aligned}
\tau &= \mathbb{E}[Y(1, Z(1))] - \mathbb{E}[Y(0, Z(0))] \\
&= \int \mathbb{E}[Y(1, Z(1)) | X = x] - \mathbb{E}[Y(0, Z(0)) | X = x] dF_X(x) \\
&= \int \mathbb{E}[Y(1, M(1), W(1)) | X = x] - \mathbb{E}[Y(0, M(0), W(0)) | X = x] dF_X(x)
\end{aligned}$$

Il suffit alors de démontrer que :

$$\mathbb{E}[Y(t, M(t'), W(t')) | X = x] = \int_{\mathbb{R}^K} \mathbb{E}[Y | T = t, M = m, W = w, X = x] dF_{(M, W) | T=t', X=x}(m, w)$$

Il s'ensuit alors que :

$$\mathbb{E}[Y(t, Z(t')) | X = x] = \int_{\mathbb{R}^K} \mathbb{E}[Y | T = t, Z = z, X = x] dF_{Z | T=t', X=x}(z)$$

On a

$$\begin{aligned}
\mathbb{E}[Y(t, M(t'), W(t')) | X = x] &= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, M(t'), W(t')) | M(t') = m, W(t') = w, X = x] \\
&\quad dF_{(M(t'), W(t')) | X=x}(m, w) \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w) | M(t') = m, W(t') = w, X = x] \\
&\quad dF_{(M(t'), W(t')) | X=x}(m, w) \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w) | T = t', M(t') = m, W(t') = w, X = x] \\
&\quad dF_{(M(t'), W(t')) | X=x}(m, w) && \text{(j)} \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w) | T = t', X = x] dF_{(M(t'), W(t')) | X=x}(m, w) && \text{(k)} \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w) | T = t, X = x] dF_{(M(t'), W(t')) | T=t', X=x}(m, w) && \text{(l)} \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w) | T = t, M(t) = m, W(t) = w, X = x] \\
&\quad dF_{(M, W) | T=t', X=x}(m, w) && \text{(m)} \\
&= \int_{\mathbb{R}^K} \mathbb{E}[Y | T = t, M = m, W = w, X = x] dF_{(M, W) | T=t', X=x}(m, w).
\end{aligned}$$

Notons que dans cette preuve nous n'avons utilisé que les hypothèses (A.1) et (A.4).

---

(j). Par (A.1) et la propriété de l'union faible :  $Y(t, m, w) \perp\!\!\!\perp T | M(t'), W(t), X = x$ .

(k). Par (A.4).

(l). Par (A.1).

(m). Par (A.4) avec  $t' = t$ .

**Effet médié par un médiateur d'intérêt**

Il découle de la définition de l'effet indirect que :

$$\begin{aligned}\delta(t) &= \mathbb{E}[Y(t, M(1), W(t))] - \mathbb{E}[Y(t, M(0), W(t))] \\ &= \int \mathbb{E}[Y(t, M(1), W(t))|X = x] - \mathbb{E}[Y(t, M(0), W(t))|X = x] dF_X(x).\end{aligned}$$

Il est alors suffisant de démontrer que :

$$\mathbb{E}[Y(t, M(t'), W(t))|X = x] = \int_{\mathbb{R}^K} \mathbb{E}[Y|T = t, M = m, W = w, X = x] dF_{(M(t'), W(t))|X=x}(m, w).$$

On a :

$$\begin{aligned}& \mathbb{E}[Y(t, (M(t'), W(t)))|X = x] \\ &= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, M(t'), W(t))|M(t') = m, W(t) = w, X = x] dF_{(M(t'), W(t))|X=x}(m, w) \\ &= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w)|T = t, M(t') = m, W(t) = w, X = x] dF_{(M(t'), W(t))|X=x}(m, w) \quad (\text{n}) \\ &= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w)|T = t, X = x] dF_{(M(t'), W(t))|X=x}(m, w) \quad (\text{o}) \\ &= \int_{\mathbb{R}^K} \mathbb{E}[Y(t, m, w)|T = t, M(t) = m, W(t) = w, X = x] dF_{(M(t'), W(t))|X=x}(m, w) \quad (\text{p}) \\ &= \int_{\mathbb{R}^K} \mathbb{E}[Y|T = t, M = m, W = w, X = x] dF_{(M(t'), W(t))|X=x}(m, w).\end{aligned}$$

Notons que dans cette preuve nous avons utilisé l'ensemble des hypothèses SIMMA. Dans le cas où  $M$  et  $W$  sont indépendants, on a :

$$\begin{aligned}dF_{(M(t'), W(t))|X=x}(m, w) &= f_{(M(t'), W(t))|X=x}(m, w) dm dw \\ &= f_{M(t')|X=x}(m) dm f_{W(t)|X=x}(w) dw \\ &= f_{M|T=t', X=x}(m) dm f_{W|T=t, X=x}(w) dw\end{aligned}$$

---

(n). Par (A.1) et la propriété de l'union faible.

(o). Par (A.5).

(p). Par (A.4) avec  $t = t'$ .



de plus :

$$\begin{aligned}
\delta(t) &= \int \int_{\mathbb{R}^K} \mathbb{E}[Y|T=t, M=m, W=w, X=x] \\
&\quad \{f_{M|T=1, X=x}(m)dm f_{W|T=t, X=x}(w)dw - f_{M|T=0, X=x}(m)dm f_{W|T=t, X=x}(w)dw\} \\
&\quad dF_X(x) \\
&= \int \int \int_{\mathbb{R}^{K-1}} \mathbb{E}[Y|T=t, M=m, W=w, X=x] \\
&\quad f_{W|T=t, X=x}(w)dw \{f_{M|T=1, X=x}(m) - f_{M|T=0, X=x}(m)\} dm dF_X(x) \\
&= \int \int \mathbb{E}[Y|T=t, M=m, X=x] \{f_{M|T=1, X=x}(m) - f_{M|T=0, X=x}(m)\} dm dF_X(x) \\
&= \int \int \mathbb{E}[Y|T=t, M=m, X=x] \{dF_{M|T=1, X=x}(m) - dF_{M|T=0, X=x}(m)\} dF_X(x).
\end{aligned}$$

## A.4 Preuve du Corollaire 1.5.1

### Identifiabilité des coefficients

Les coefficients des équations (2.2.1) et (2.2.2) sont identifiés sous les hypothèses du théorème 1.4.1. En effet en réécrivant les équations (2.2.1) et (2.2.2) avec des contrefactuels nous avons :

$$Z(T) = \alpha_2^\Gamma + \beta_2^\Gamma T + \xi_2^\Gamma X + \Upsilon_2(T) \quad (\text{A.6})$$

$$Y(T, Z(T)) = \alpha_3 + \beta_3 T + \gamma^\Gamma Z(T) + \xi_3^\Gamma X + \epsilon_3(T, Z(T)) \quad (\text{A.7})$$

où  $\mathbb{E}[\Upsilon_2(t)] = \mathbb{E}[\epsilon_3(t, z)] = 0$ .

Les hypothèses (A.4) et (A.5) impliquent les indépendances suivantes :

$$\begin{aligned}
\Upsilon_2(t) &\perp\!\!\!\perp T|X=x \\
\epsilon_3(t, z) &\perp\!\!\!\perp T|X=x.
\end{aligned}$$

L'hypothèse (A.5) implique l'indépendance suivante :

$$\epsilon_3(t, z) \perp\!\!\!\perp Z(t)|T=t, X=x.$$

On a

$$\begin{aligned}
\mathbb{E}[Z(T)|T=t, X=x] &= \alpha_2^\Gamma + \beta_2^\Gamma t + \xi_2^\Gamma x + \mathbb{E}[\Upsilon_2(T)|T=t, X=x] \\
&= \alpha_2^\Gamma + \beta_2^\Gamma t + \xi_2^\Gamma x + \mathbb{E}[\Upsilon_2(t)|T=t, X=x] \\
&= \alpha_2^\Gamma + \beta_2^\Gamma t + \xi_2^\Gamma x + \mathbb{E}[\Upsilon_2(t)] \\
&= \alpha_2^\Gamma + \beta_2^\Gamma t + \xi_2^\Gamma x.
\end{aligned}$$

Les coefficients de l'équation (2.2.1) sont alors identifiés.

De plus

$$\begin{aligned}
\mathbb{E}[Y(T, Z(T)) | Z = z, T = t, X = x] &= \alpha_3 + \beta_3 t + \gamma^\Gamma z + \xi_3^\Gamma x \\
&\quad + \mathbb{E}[\epsilon_3(T, Z(T)) | Z = z, T = t, X = x] \\
&= \alpha_3 + \beta_3 t + \gamma^\Gamma z + \xi_3^\Gamma x \\
&\quad + \mathbb{E}[\epsilon_3(t, z) | Z = z, T = t, X = x] \\
&= \alpha_3 + \beta_3 t + \gamma^\Gamma z + \xi_3^\Gamma x \\
&\quad + \mathbb{E}[\epsilon_3(t, z)] \\
&= \alpha_3 + \beta_3 t + \gamma^\Gamma z + \xi_3^\Gamma x.
\end{aligned}$$

Les coefficients de l'équation (2.2.2) sont alors identifiés.

### Effet médié par un médiateur d'intérêt

Selon le Théorème 1.4.1 on a :

$$\begin{aligned}
\delta^1(t) &= \int \int_{\mathbb{R}^K} \mathbb{E}[Y | M = m, W = w, T = t, X = x] \\
&\quad \{dF_{(M(1), W(t)) | X=x}(m, w) - dF_{(M(0), W(t)) | X=x}(m, w)\} dF_X(x) \\
&= \int \int_{\mathbb{R}^K} \left( \alpha_3 + \beta_3 t + \sum_{j=1}^K \gamma_j m^j + \xi_3^\Gamma x \right) \\
&\quad \{dF_{(M(1), W(t)) | X=x}(m, w) - dF_{(M(0), W(t)) | X=x}(m, w)\} dF_X(x) \\
&= \int \int_{\mathbb{R}^K} \left( \alpha_3 + \beta_3 t + \xi_3^\Gamma x \right) \{dF_{(M(1), W(t)) | X=x}(m, w) - dF_{(M(0), W(t)) | X=x}(m, w)\} \\
&\quad + \int \int_{\mathbb{R}^K} \sum_{j=1}^K \gamma_j m^j \{dF_{(M(1), W(t)) | X=x}(m, w) - dF_{(M(0), W(t)) | X=x}(m, w)\} dF_X(x). \\
&= \int \int_{\mathbb{R}^K} \sum_{j=1}^K \gamma_j m^j \{dF_{(M(1), W(t)) | X=x}(m, w) - dF_{(M(0), W(t)) | X=x}(m, w)\} dF_X(x) \\
&= \int \int_{\mathbb{R}^K} \sum_{j=1}^K \gamma_j m^j dF_{(M(1), W(t)) | X=x}(m, w) \\
&\quad - \int \int_{\mathbb{R}^K} \sum_{k=1}^K \gamma_j m^j dF_{(M(0), W(t)) | X=x}(m, w) dF_X(x).
\end{aligned}$$

En faisant les substitutions suivantes :  $m^1 = \alpha_2^1 + \beta_2^1 t + \xi_2^{\Gamma 1} x + e_2^1$  et  $m^j = \alpha_2^j + \beta_2^j t + \xi_2^{\Gamma j} x + e_2^j, \forall j \in [2, K]$ , on a :

$$((M(t'), W(t)) | X = x) \sim \mathcal{N}(\mu_{(t',t)}, \Sigma_2) \text{ avec } \mu_{(t',t)} = \begin{pmatrix} \alpha_2^1 + \beta_2^1 t' + \xi_2^{\Gamma 1} x \\ \alpha_2^2 + \beta_2^2 t + \xi_2^{\Gamma 2} x \\ \alpha_2^3 + \beta_2^3 t + \xi_2^{\Gamma 3} x \\ \vdots \\ \alpha_2^K + \beta_2^K t + \xi_2^{\Gamma K} x \end{pmatrix} \text{ et où il dé-}$$

coule de l'hypothèse (1.4.2) que  $\Sigma_2$  est la matrice de covariance de  $e_2^j$ .

Il s'ensuit alors que :

$$\begin{aligned} f_{(M(t'), W(t)) | X=x}(m, w) &= \frac{1}{\sqrt{2\pi} |\Sigma_2|^{1/2}} \exp\left(\frac{1}{2} \left( \begin{pmatrix} m \\ w \end{pmatrix} - \mu_{(t',t)} \right) \Sigma_2^{-1} \left( \begin{pmatrix} m \\ w \end{pmatrix} - \mu_{(t',t)} \right)\right) \\ &= \frac{1}{\sqrt{2\pi} |\Sigma_2|^{1/2}} \exp\left(\frac{1}{2} \begin{pmatrix} e_2^1 \\ \vdots \\ e_2^K \end{pmatrix} \Sigma_2^{-1} \begin{pmatrix} e_2^1 \\ \vdots \\ e_2^K \end{pmatrix}\right) \\ &= f_{\Upsilon} \begin{pmatrix} e_2^1 \\ \vdots \\ e_2^K \end{pmatrix}. \end{aligned}$$

En remplaçant la densité jointe des médiateurs contrefactuels dans l'équation précédente de l'effet indirect  $\delta^1(t)$ , on a :

$$\begin{aligned} \delta^1(t) &= \int \int_{\mathbb{R}^K} \gamma_1 (\alpha_2^1 + \beta_2^1 \times 1 + \xi_2^{\Gamma 1} x + e_2^1) + \sum_{j=2}^K \gamma_j (\alpha_2^j + \beta_2^j t + \xi_2^{\Gamma j} x + e_2^j) \\ &\quad dF_{\Upsilon} \begin{pmatrix} e_2^1 \\ \vdots \\ e_2^K \end{pmatrix} \\ &\quad - \int \int_{\mathbb{R}^K} \gamma_1 (\alpha_2^1 + \beta_2^1 \times 0 + \xi_2^{\Gamma 1} x + e_2^1) + \sum_{j=2}^K \gamma_j (\alpha_2^j + \beta_2^j t + \xi_2^{\Gamma j} x + e_2^j) \\ &\quad dF_{\Upsilon} \begin{pmatrix} e_2^1 \\ \vdots \\ e_2^K \end{pmatrix} dF_X(x) \\ &= \int \int_{\mathbb{R}^K} \gamma_1 \beta_2^1 dF_{\Upsilon} \begin{pmatrix} e_2^1 \\ \vdots \\ e_2^K \end{pmatrix} dF_X(x) \\ &= \gamma_1 \beta_2^1. \end{aligned}$$

On peut alors conclure que de façon générale pour  $k \in [1, K]$ ,  $\delta^k(t) = \gamma_k \beta_2^k$ .

**Effet direct**

$$\begin{aligned}
\zeta(t) &= \int \int \mathbb{E}(Y|Z = z, T = 1, X = x) - \mathbb{E}(Y|Z = z, T = 0, X = x) \\
&\quad dF_{Z|T=t, X=x}(z) dF_X(x) \\
&= \int \int \{\alpha_3 + \beta_3 + \gamma^\Gamma z + \xi_3^\Gamma x - \alpha_3 - \gamma^\Gamma z - \xi_3^\Gamma x\} dF_{Z|T=t}(z) dF_X(x) \\
&= \int \int \beta_3 dF_{Z|T=t}(z) dF_X(x) \\
&= \beta_3.
\end{aligned}$$

où  $\gamma^\Gamma = (\gamma_1, \dots, \gamma_K)^\Gamma$ .

**Effet médié joint**

Notons que

$$\begin{aligned}
\eta^1(t) &= \int \int_{\mathbb{R}^K} \mathbb{E}[Y|M = m, W = w, T = t, X = x] \\
&\quad \{dF_{(M(1-t), W(1))|X=x}(m, w) - dF_{(M(1-t), W(0))|X=x}(m, w)\} dF_X(x) \\
&= \int \int_{\mathbb{R}^K} \left( \alpha_3 + \beta_3 t + \sum_{j=1}^K \gamma_j m^j + \xi_3^\Gamma x \right) \\
&\quad \{dF_{(M(1-t), W(1))|X=x}(m, w) - dF_{(M(1-t), W(0))|X=x}(m, w)\} dF_X(x) \\
&= \int \int_{\mathbb{R}^K} \sum_{j=1}^K \gamma_j m^j dF_{(M(1-t), W(1))|X=x}(m, w) \\
&\quad - \int \int_{\mathbb{R}^K} \sum_{j=1}^K \gamma_j m^j dF_{(M(1-t), W(0))|X=x}(m, w) dF_X(x)
\end{aligned}$$

En faisant les substitutions suivantes  $m^1 = \alpha_2^1 + \beta_2^1(1-t) + \xi_2^{\Gamma 1}x + e_2^1$  et  $m^j = \alpha_2^j + \beta_2^j t' + \xi_2^{\Gamma j}x + e_2^j, \forall j \in [2, K]$ , on a :

$$\begin{aligned}
\eta^1(t) &= \int \int_{\mathbb{R}^K} \gamma_1 \left( \alpha_2^1 + \beta_2^1(1-t) + \xi_2^{\Gamma^1} x + e_2^1 \right) \\
&\quad + \sum_{j=2}^K \gamma_j \left( \alpha_2^j + \beta_2^j \times 1 + \xi_2^{\Gamma^j} x + e_2^j \right) dF_{\Upsilon} \left( \begin{array}{c} e_2^1 \\ \vdots \\ e_2^K \end{array} \right) \\
&\quad - \int \int_{\mathbb{R}^K} \gamma_1 \left( \alpha_2^1 + \beta_2^1(1-t) + \xi_2^{\Gamma^1} x + e_2^1 \right) \\
&\quad + \sum_{j=2}^K \gamma_j \left( \alpha_2^j + \beta_2^j \times 0 + \xi_2^{\Gamma^j} x + e_2^j \right) dF_{\Upsilon} \left( \begin{array}{c} e_2^1 \\ \vdots \\ e_2^K \end{array} \right) dF_X(x),
\end{aligned}$$

Il s'ensuit que

$$\begin{aligned}
\eta^1(t) &= \int \int_{\mathbb{R}^K} \sum_{j=2}^K \gamma_j \beta_2^j dF_{\Upsilon} \left( \begin{array}{c} e_2^1 \\ \vdots \\ e_2^K \end{array} \right) dF_X(x) \\
&= \sum_{j=2}^K \gamma_j \beta_2^j \\
&= \sum_{j=2}^K \delta^j(t).
\end{aligned}$$

On peut donc conclure que de façon générale pour  $k \in [1, K]$ ,  $\eta^k(t) = \sum_{j=1, j \neq k}^K \delta^j(t)$ .

On a donc :

$$\begin{aligned}
\delta^Z(t) &= \frac{\sum_{k=1}^K \left( \delta^k(t) + \eta^k(t) \right)}{K} \\
&= \frac{\sum_{k=1}^K \left( \delta^k(t) + \sum_{j=1, j \neq k}^K \delta^j(t) \right)}{K} \\
&= \frac{K \sum_{k=1}^K \delta^k(t)}{K} \\
&= \sum_{k=1}^K \delta^k(t).
\end{aligned}$$

### A.5 Preuve du Corollaire 1.6.1

En injectant (2.2.3) dans (2.2.4), on a :

$$Y^* = (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2) + (\beta_3 + \sum_{k=1}^K \gamma_k \beta_2^k) T + (\epsilon_3 + \sum_{k=1}^K \gamma_k \epsilon_2^k).$$

On redéfinit le terme d'erreur de l'équation obtenue par  $U = \epsilon_3 + \varepsilon$ , avec  $\varepsilon = \sum_{k=1}^K \gamma_k \epsilon_2^k$ .

Les termes d'erreur des médiateurs forment un vecteur gaussien  $\Upsilon = \begin{pmatrix} \epsilon_2^1 \\ \vdots \\ \epsilon_2^K \end{pmatrix}$ . Cela

implique que  $\varepsilon$  suit une loi gaussienne.

Avant de passer à la preuve du Corollaire 1.6.1, nous devons déterminer les fonctions de distributions de  $U$ , notée  $F_U$ , pour les modélisations probit et logit de la variable de réponse.

**Modélisation Probit.** Dans le cas d'une modélisation probit,  $\epsilon_3 \sim \mathcal{N}(0, \sigma_3^2)$ , et, par les hypothèses (A.4) et (A.5),  $\epsilon_3 \perp \Upsilon$ . On en déduit alors que  $U$  suit une loi gaussienne et que la fonction de distribution de  $U$  est :

$$f_U(z) = \frac{\exp\left(\frac{z^2}{2(\sigma_3^2 + \sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j))}\right)}{\sqrt{2\pi} \sqrt{\sigma_3^2 + \sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}$$

$$F_U(z) = \Phi\left(\frac{z \cdot pdf}{\sqrt{\sigma_3^2 + \sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}\right)$$

Notons que  $1 - F_U(-z) = F_U(z)$  par symétrie de la fonction de densité.

**Modélisation logit.** Dans le cas d'une modélisation logit,  $\epsilon_3 \sim \mathcal{L}(0,1)$  et  $f_{\epsilon_3}(z) = \frac{\exp(-z)}{(1 + \exp(-z))^2}$ . Pour déterminer la fonction de densité de  $U$ , considérons la fonction générique continue  $G$  avec support compact. La densité de  $U$  est alors la fonction  $f_U(u)$  telle que  $E[G(U)] = \int G(u)f_U(u)du$ . Selon (A.5),  $\epsilon_3 \perp \Upsilon$ , on a donc  $\epsilon_3 \perp \varepsilon$ .

$$\begin{aligned} \mathbb{E}[G(U)] &= \mathbb{E}[G(\varepsilon + \epsilon_3)] \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} G(e + e_3) f_{\varepsilon}(e) f_{\epsilon_3}(e_3) de de_3. \end{aligned}$$

On fait les substitutions suivantes :

$$\begin{aligned} z &= e + e_3 \\ e &= z - e_3 \\ de &= dz. \end{aligned} \tag{A.8}$$

On a alors :

$$\begin{aligned} E[G(U)] &= \int_{\mathbb{R}} \int_{\mathbb{R}} G(z) f_{\varepsilon}(z - e_3) f_{\epsilon_3}(e_3) dz de_3 \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} G(z) \frac{\exp\left(\frac{(z - e_3)^2}{2\left(\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)\right)}\right)}{\sqrt{2\pi} \sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} f_{\epsilon_3}(e_3) dz de_3 \\ &= \int_{\mathbb{R}} G(z) \frac{1}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \int_{\mathbb{R}} f_{\mathcal{N}(0,1)}\left(\frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}\right) f_{\epsilon_3}(e_3) de_3 dz. \end{aligned}$$

On en déduit la fonction de densité de  $U$  :

$$f_U(u) = \frac{1}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \int_{\mathbb{R}} f_{\mathcal{N}(0,1)} \left( \frac{u - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) f_{\epsilon_3}(e_3) de_3.$$

La fonction de distribution de  $U$  est alors :

$$\begin{aligned} F_U(z) &= \int_{-\infty}^{z.pdf} f_U(u) du \\ &= \int_{-\infty}^{z.pdf} \frac{1}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \int_{\mathbb{R}} f_{\mathcal{N}(0,1)} \left( \frac{u - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) f_{\epsilon_3}(e_3) de_3 du \\ &= \int_{\mathbb{R}} \frac{1}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \int_{-\infty}^{z.pdf} f_{\mathcal{N}(0,1)} \left( \frac{u - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) du f_{\epsilon_3}(e_3) de_3. \end{aligned}$$

On fait les substitutions suivantes :

$$\begin{aligned} t &= \frac{u - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \\ du &= dt \sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)} \end{aligned}$$

Les bornes de l'intégrale deviennent donc :

$u$	$-\infty$	$z$
$t$	$-\infty$	$\frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}$



On pose  $b = \frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}$ , on a alors :

$$\begin{aligned}
 F_U(z) &= \int_{\mathbb{R}} \frac{1}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \int_{-\infty}^b f_{\mathcal{N}(0,1)}(t) \sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)} dt f_{\epsilon_3}(e_3) de_3 \\
 &= \int_{\mathbb{R}} \int_{-\infty}^b f_{\mathcal{N}(0,1)}(t) dt f_{\epsilon_3}(e_3) de_3 \\
 &= \int_{\mathbb{R}} \Phi(b) f_{\epsilon_3}(e_3) de_3 \\
 &= \int_{\mathbb{R}} \Phi\left(\frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}\right) f_{\epsilon_3}(e_3) de_3 \\
 &= \int_{\mathbb{R}} \Phi\left(\frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}\right) \frac{\exp(-e_3)}{(1 + \exp(-e_3))^2} de_3.
 \end{aligned}$$

Comme on a

$$\begin{aligned}
 f_{\epsilon_3}(e_3) &= \frac{\exp(-e_3)}{(1 + \exp(-e_3))^2} \\
 &= \frac{\exp(e_3)}{(1 + \exp(e_3))^2},
 \end{aligned}$$

On déduit que

$$F_U(z) = \int_{\mathbb{R}} \Phi\left(\frac{z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}}\right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3.$$

De plus, tout comme dans la modélisation probit on peut prouver que  $1 - F_U(-z) = F_U(z)$ .

$$\begin{aligned}
1 - F_U(-z) &= 1 - \int_{\mathbb{R}} \Phi \left( \frac{-z - e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \\
&= 1 - \int_{\mathbb{R}} \Phi \left( -\frac{z + e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \\
&= 1 - \int_{\mathbb{R}} \left[ 1 - \Phi \left( \frac{z + e_3}{\sqrt{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)}} \right) \right] \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \\
&= 1 - \int_{\mathbb{R}} \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \\
&\quad + \int_{\mathbb{R}} \Phi \left( \frac{z + e_3}{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)} \right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \\
&= 1 - 1 + \int_{\mathbb{R}} \Phi \left( \frac{z + e_3}{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)} \right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3
\end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbb{R}} \Phi \left( \frac{z + e_3}{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)} \right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \\
&= \int_{\mathbb{R}} \Phi \left( \frac{z - e_3}{\sum_{k=1}^K \sum_{j=1}^K \gamma_k \gamma_j \text{cov}(\epsilon_2^k, \epsilon_2^j)} \right) \frac{\exp(e_3)}{(1 + \exp(e_3))^2} de_3 \text{ (q)} \\
&= F_U(z).
\end{aligned}$$

Nous pouvons à présent donner la preuve du Corollaire 1.6.1. Pour faciliter le développement de la preuve on pose :

$$\begin{aligned}
\delta^Z(t) &= \int A_{t1} - A_{t0} dF_X(x) \\
\zeta(t) &= \int B_{1t'} - B_{0t'} dF_X(x),
\end{aligned}$$

où :

$$\begin{aligned}
A_{tt'} &= \int_{\mathbb{R}^K} \mathbb{E}[Y|T = t, M = m, W = w, X = x] dF_{(M(t'), W(t'))|X=x}(m, w) \\
B_{tt'} &= \int_{\mathbb{R}^K} \mathbb{E}[Y|Z = z, T = t, X = x] dF_{Z|T=t', X=x}(z).
\end{aligned}$$

Pour l'effet indirect  $\delta(t)$  on a :

$$\begin{aligned}
A_{tt'} &= \int_{\mathbb{R}^K} \mathbb{E}[Y|T = t, M = m, W = w, X = x] dF_{(M(t'), W(t'))|X=x}(m, w) \\
&= \int_{\mathbb{R}^K} \mathbb{E}[\mathbb{1}_{\{Y^* > 0\}}|T = t, M = m, W = w, X = x] dF_{(M(t'), W(t'))|X=x}(m, w) \\
&= \int_{\mathbb{R}^K} \mathbb{P}[Y^* > 0|T = t, M = m, W = w, X = x] dF_{(M(t'), W(t'))|X=x}(m, w).
\end{aligned}$$

On remplace  $Y^*$  par son expression et on obtient :

$$A_{tt'} = \int_{\mathbb{R}^K} \mathbb{P}[(\alpha_3 + \beta_3 t + \sum_{k=2}^K \gamma_k m^k + \gamma m + \xi_3^\Gamma x + \epsilon_3 > 0] dF_{(M(t'), W(t'))|X=x}(m, w).$$

---

(q). Ici on substitut  $e_3$  par  $-e_3$ .

À l'aide des substitutions  $m^1 = \alpha_2^1 + \beta_2^1 t' + \xi_2^{\Gamma 1} x + e_2^1$  et  $m^k = \alpha_2^k + \beta_2^k t + \xi_2^{\Gamma k} x + e_2^k \forall k \geq 2$ , on a :

$$\begin{aligned}
A_{tt'} &= \int_{\mathbb{R}^K} \mathbb{P} \left[ (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + (\epsilon_3 + \sum_{k=1}^K \gamma_k e_2^k) > 0 \right] \\
&\quad dF_{\Upsilon} \left( \begin{array}{c} e_2^1 \\ \vdots \\ e_2^K \end{array} \right) \\
&= \int_{\mathbb{R}} \mathbb{P} \left[ (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + (\epsilon_3 + e) > 0 \right] dF_\epsilon(e) \\
&= \int_{\mathbb{R}} \mathbb{E} \left[ \mathbb{1}_{\left\{ (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + (\epsilon_3 + e) > 0 \right\}} \right] dF_\epsilon(e) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_{\left\{ (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + (e_3 + e) > 0 \right\}} dF_{e_3}(e_3) dF_\epsilon(e) \\
&= \int_{\mathbb{R}} \mathbb{1}_{\left\{ (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + u > 0 \right\}} dF_U(u) \\
&= \mathbb{E} \left[ \mathbb{1}_{\left\{ (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + U > 0 \right\}} \right] \\
&= \mathbb{P} \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x + U > 0 \right) \\
&= \mathbb{P} \left( U > -(\alpha_3 + \sum_{j=1}^J \gamma_j \alpha_2^j) - (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t - \gamma \beta_2^1 t' - (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right) \\
&= 1 - \mathbb{P} \left( U \leq -(\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) - (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t - \gamma \beta_2^1 t' - (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right) \\
&= 1 - F_U \left( -(\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) - (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t - \gamma \beta_2^1 t' - (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right)
\end{aligned}$$

Comme  $1 - F_U(-z) = F_U(z)$ , la dernière égalité devient :

$$A_{tt'} = F_U \left( (\alpha_3 + \sum_{k=1}^K \gamma_k \alpha_2^k) + (\beta_3 + \sum_{k=2}^K \gamma_k \beta_2^k) t + \gamma \beta_2^1 t' + (\xi_3^\Gamma + \sum_{k=1}^K \gamma_k \xi_2^{\Gamma k}) x \right)$$

En injectant  $A_{tt'}$ , donné dans cette dernière égalité, dans  $\delta(t) = \int A_{t1} - A_{t0} dF_X(x)$ , on obtient le résultat du corollaire. De façon analogue la preuve est obtenue pour les autres effets causaux de médiation.

## A.6 Modèles de simulation

Nous donnons ici les modèles utilisés pour l'étude de simulation dans la sous-section 1.7.3.

### Modèle 1 : Variable de réponse continue et médiateurs continus

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$
- La loi jointe des deux médiateurs contrefactuels est

$$\begin{pmatrix} M^1(1) \\ M^1(0) \\ M^2(1) \\ M^2(0) \end{pmatrix} \sim \mathcal{N} \left( \mu = \begin{pmatrix} 1 + 4 \times 1 \\ 1 + 4 \times 0 \\ 2 + 6 \times 1 \\ 2 + 6 \times 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 1 & \rho & \rho \\ 1 & 1 & \rho & \rho \\ \rho & \rho & 1 & 1 \\ \rho & \rho & 1 & 1 \end{pmatrix} \right)$$

- La variable de réponse contrefactuel suit une loi normale :

$$Y(t, M^1(t'), M^2(t'')) \sim \mathcal{N}(1 + 10t + 5M^1(t') + 4M^2(t''), 1).$$

Dans la Table 4.5, nous montrons que les vrais effets causaux du modèle 1.

$\delta^Z$	$\delta^1$	$\delta^2$	$\zeta$	$\tau$
44	20	24	10	54

TABLE 4.5 – Variable de réponse continu (modèle 1) : Valeurs des effets causaux obtenus avec modèle 1.

### Modèle 2 : Variable de réponse binaire (logit) avec des médiateurs continus

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$
- La loi jointe des deux médiateurs contrefactuels est :

$$\begin{pmatrix} M^1(1) \\ M^1(0) \\ M^2(1) \\ M^2(0) \end{pmatrix} \sim \mathcal{N} \left( \mu = \begin{pmatrix} 0.1 + 0.6 \times 1 \\ 0.1 + 0.6 \times 0 \\ 0.2 + 0.8 \times 1 \\ 0.2 + 0.8 \times 0 \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 1 & \rho & \rho \\ 1 & 1 & \rho & \rho \\ \rho & \rho & 1 & 1 \\ \rho & \rho & 1 & 1 \end{pmatrix} \right)$$

- La variable de réponse contrefactuelle suit une loi logit :

$$Y(t, M^1(t'), M^2(t'')) \sim B \left( \frac{1}{1 + \exp(-2 + 0.4t + 0.6M^1(t') + 0.8M^2(t''))} \right).$$

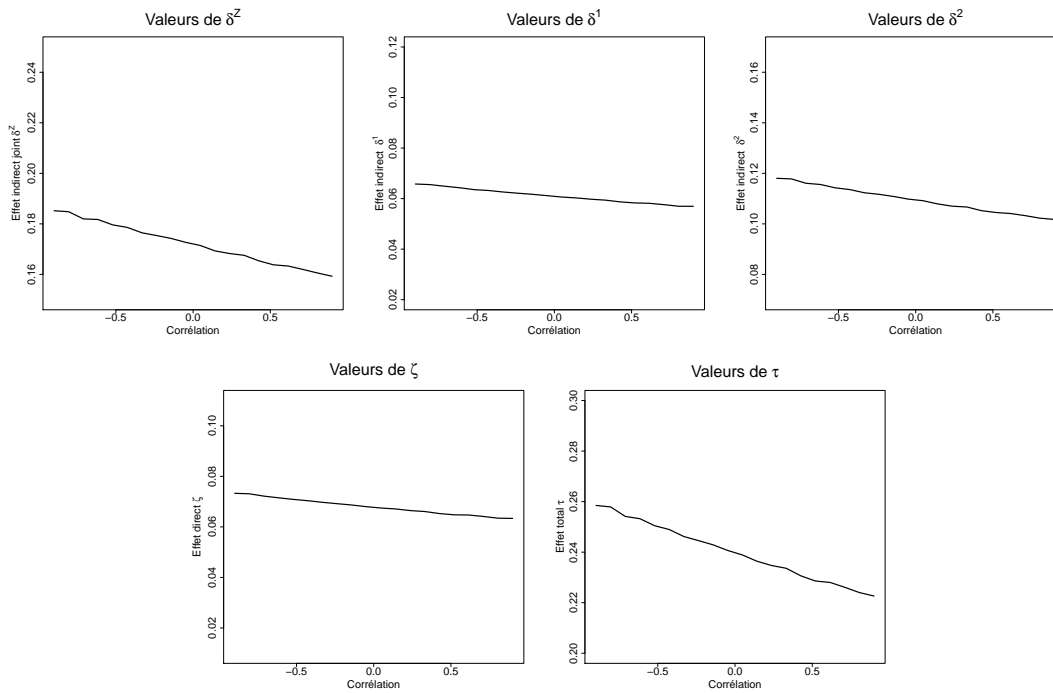


FIGURE 4.4 – Variable de réponse binaire (modèle 2) : variation des effets causaux due à la corrélation.

Avec ce choix de paramètres, 30% des observations échantillonnées sont des cas. Comme nous pouvons le voir dans le Corollaire 1.6.1, avec un résultat binaire, les effets causaux sont liés à la covariance des médiateurs. La Figure 4.4 montre comment les vraies valeurs causales changent lorsque la corrélation change.

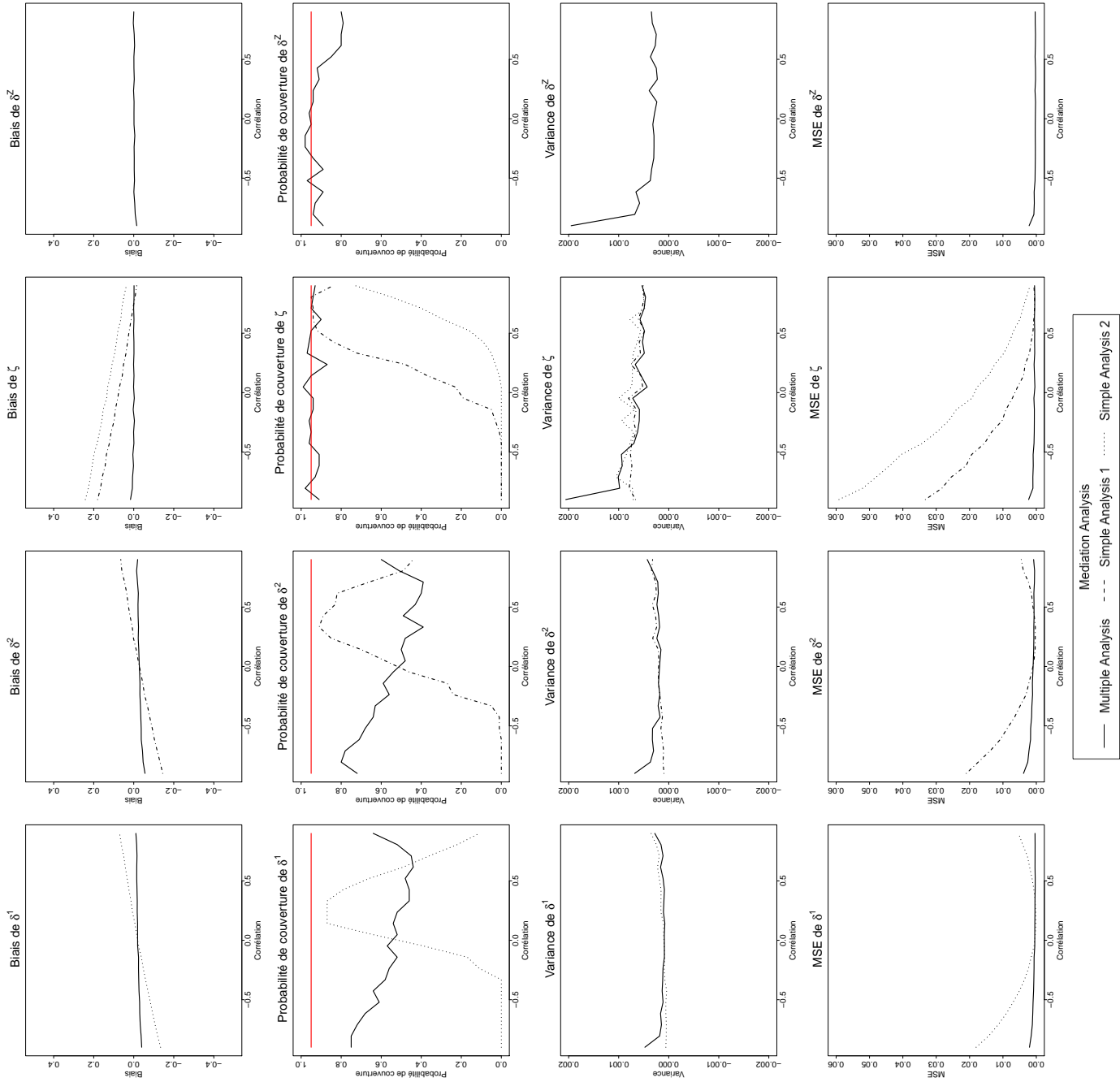


FIGURE 4.5 – Modèle 2 (variable de réponse binaire) : biais, probabilité de couverture, variance et MSE des estimateurs de l'effet de médiation lorsque la corrélation entre les médiateurs varie. Ces résultats ont été obtenus avec 200 simulations. Chaque simulation est un jeu de données de taille 1000.

## A.7 Résultats complémentaires

L'une des hypothèses nécessaires à nos résultats est que les corrélations entre les médiateurs potentiels sont les mêmes quel que soit le traitement régissant les médiateurs :  $cor(M^i(t), M^j(t'))|T, X) = \rho_{ij}$  pour tout  $t, t'$ . Dans les deux modèles de génération de données suivants pour les médiateurs, nous ne considérons pas cette hypothèse.  $M^1(t)$  et  $M^2(t)$  suivent une distribution normale bivariée avec une covariance  $\Sigma$  pour chaque  $t = 0, 1$  séparément. Alors, clairement les données simulées ont  $cor(M^1(1), M^2(0)) = 0$ . En analysant les données générées avec les modèles 1 et 2, nous vérifions la robustesse de notre méthode par rapport à notre hypothèse. Notez que nous avons choisi la même valeur de paramètre que le modèle utilisé dans la sous-section 1.7.3.

Les résultats pour le biais et la probabilité de couverture sont montrés dans les figures 4.7 et 4.5. Cette Figure montre clairement que notre approche permet une estimation non biaisée, contrairement aux analyses simples, tant pour les effets directs qu'indirects. Les interprétations de ces résultats sont très proches de celles de l'analyse avec le modèle respectant notre hypothèse dans l'article.

### Modèle 1b : Variable de réponse continue et médiateurs continus

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$ .
- La loi jointe des deux médiateurs contrefactuels est

$$\begin{pmatrix} M^1(t) \\ M^2(t) \end{pmatrix} \sim \mathcal{N}\left(\mu = \begin{pmatrix} 1 + 4t \\ 2 + 6t \end{pmatrix}, \Sigma\right).$$

- La variable de réponse contrefactuelle suit une loi normale

$$Y(t, M^1(t'), M^2(t'')) \sim \mathcal{N}(1 + 10t + 5M^1(t') + 4M^2(t''), 1).$$

Dans la Table 4.6, nous montrons les effets causaux obtenus avec le modèle 1b.

$\delta^Z$	$\delta^1$	$\delta^2$	$\zeta$	$\tau$
44	20	24	10	54

TABLE 4.6 – Variable de réponse continu (modèle 1b) : Valeurs des effets causaux obtenus avec le modèle 1b.

### Modèle 2b : Variable de réponse binaire (logit) avec des médiateurs continus

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$ .



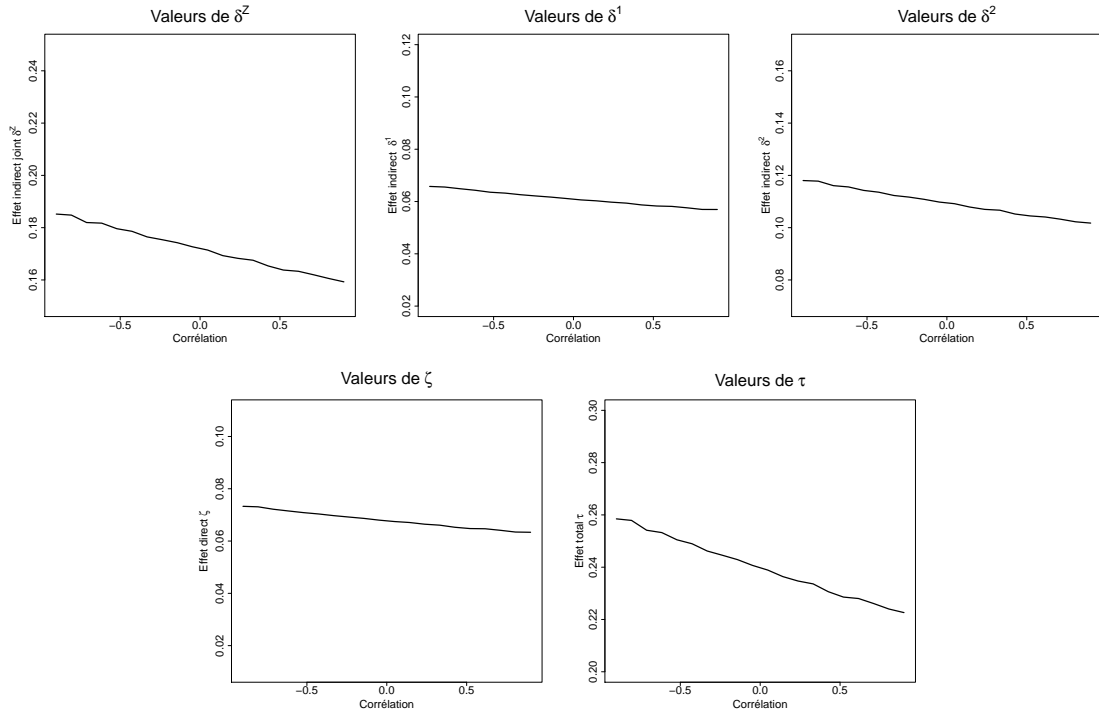


FIGURE 4.6 – Variable de réponse binaire (modèle 2b) : variation des effets causaux due à la corrélation.

— La loi jointe des deux médiateurs contrefactuels est

$$\begin{pmatrix} M^1(t) \\ M^2(t) \end{pmatrix} \sim \mathcal{N}\left(\mu = \begin{pmatrix} 0.1 + 0.6t \\ 0.2 + 0.8t \end{pmatrix}, \Sigma\right).$$

— La variable de réponse contrefactuelle suit une loi logit :

$$Y(t, M^1(t'), M^2(t'')) \sim B\left(\frac{1}{1 + \exp(-2 + 0.4t + 0.6M^1(t') + 0.8M^2(t''))}\right).$$

Avec ce choix de paramètres, 30% des observations échantillonnées sont des cas. Comme nous pouvons le voir dans le Corollaire 1.6.1, avec un résultat binaire, les effets causaux sont liés à la covariance des médiateurs. La Figure 4.6 montre comment les vraies valeurs causales changent lorsque la corrélation change.

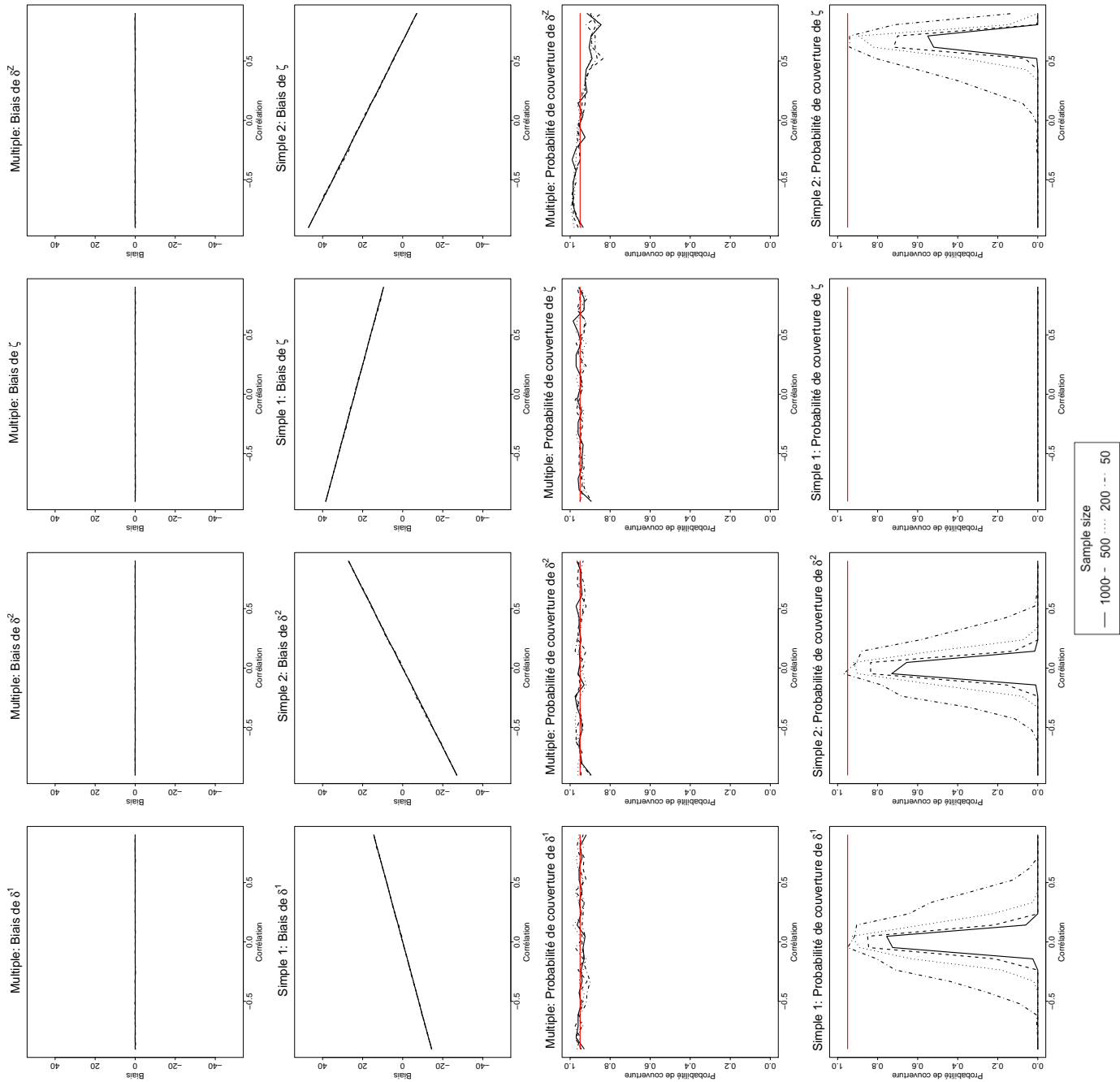


FIGURE 4.7 – Modèle 1b (variable de réponse continue) : biais, probabilité de couverture, variance et MSE des estimateurs de l'effet de médiation lorsque la corrélation entre les médiateurs varie. Ces résultats ont été obtenus avec 200 simulations. Chaque simulation est un jeu de données de taille 1000.

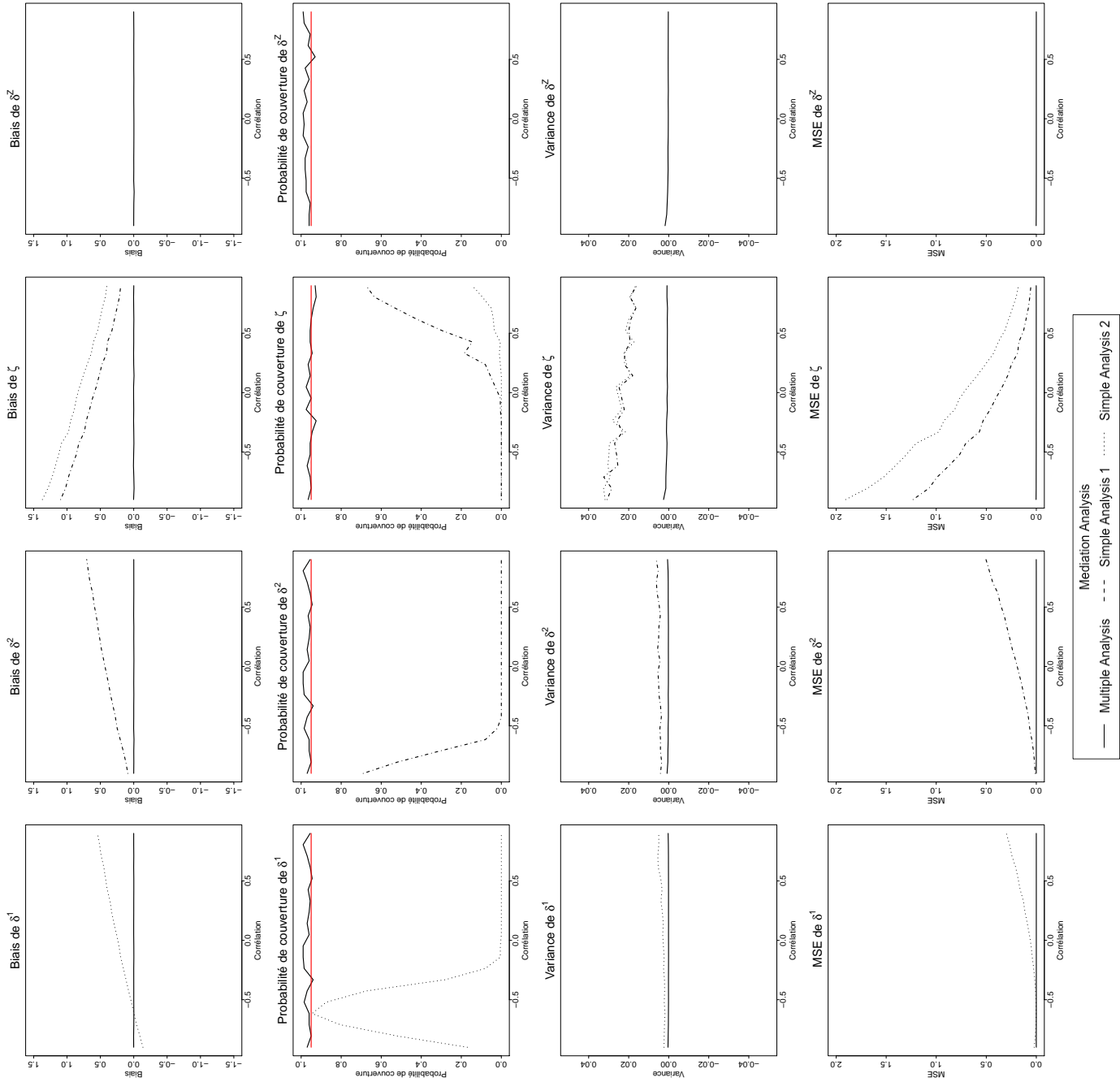


FIGURE 4.8 – Modèle 2b (variable de réponse binaire) : biais, probabilité de couverture, variance et MSE des estimateurs de l'effet de médiation lorsque la corrélation entre les médiateurs varie. Ces résultats ont été obtenus avec 200 simulations. Chaque simulation est en un jeu de données de taille 1000.

**Modèle 3 : Variable de réponse continue et médiateurs continus**

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$ .
- La loi jointe des deux médiateurs contrefactuels est

$$\begin{pmatrix} M^1(t) \\ M^2(t) \end{pmatrix} \sim \mathcal{N}\left(\mu = \begin{pmatrix} \frac{1}{2} + \frac{3}{2}t \\ \frac{2}{2+6t} \end{pmatrix}, \Sigma\right).$$

- La variable de réponse contrefactuelle suit une loi normale

$$Y(t, M^1(t'), M^2(t'')) \sim \mathcal{N}(4 + 35t + 2M^1(t') + 3M^2(t''), 1).$$

**Modèle 4 : Variable de réponse binaire (probit) avec des médiateurs continus**

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$ .
- La loi jointe des deux médiateurs contrefactuels est

$$\begin{pmatrix} M^1(t) \\ M^2(t) \end{pmatrix} \sim \mathcal{N}\left(\mu = \begin{pmatrix} 0.2 + 0.7t \\ 0.4 + 0.7t \end{pmatrix}, \Sigma\right).$$

- La variable de réponse contrefactuelle suit une loi probit :

$$Y(t, M^1(t'), M^2(t'')) \sim \mathcal{B}\left(\Phi\left(-0.5 + 0.8t + 0.7M^1(t') + 0.7M^2(t'')\right)\right).$$

**Modèle 5 : Variable de réponse binaire (logit) avec des médiateurs continus**

- $T$  suit une loi de Bernoulli  $\mathcal{B}(0.3)$ .
- La loi jointe des deux médiateurs contrefactuels est

$$\begin{pmatrix} M^1(t) \\ M^2(t) \end{pmatrix} \sim \mathcal{N}\left(\mu = \begin{pmatrix} 0.2 + 0.7t \\ 0.4 + 0.7t \end{pmatrix}, \Sigma\right).$$

- La variable de réponse contrefactuelle suit une loi logit :

$$Y(t, M^1(t'), M^2(t'')) \sim B\left(\frac{1}{1 + \exp(0.5 - 0.8t - 0.7M^1(t') - 0.7M^2(t''))}\right).$$

**Médiateurs indépendants et corrélés**

Pour tous les modèles on considère deux situations :

- modèle I quand  $\Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  (médiateurs indépendants)
- modèle C quand  $\Sigma = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}$  (médiateurs corrélés).

Par exemple le Modèle 4 C correspond au cas où la variable de réponse est binaire (probit) avec des médiateurs continus.

La Table 4.7 résume les estimations fournies par notre analyse de médiation multiple et par deux analyses de médiation simples, une pour chaque médiateur, pour les données simulées selon les six modèles. Nous pouvons voir que lorsque les médiateurs sont indépendants, les deux analyses simples et notre analyse multiple estiment correctement les effets indirects individuels, l'effet total et la proportion médiée. Comme prévu, les estimations de l'effet direct fournies par les deux analyses simples sont éloignées de la valeur réelle qui n'appartient même pas aux intervalles de confiance. Au contraire, notre analyse multiple fournit une estimation précise et précise de l'effet direct. Lors de l'analyse de données avec des médiateurs corrélés générés par le modèle 3C, une analyse simple fournit des estimations erronées des effets directs et indirects. Les proportions médiatisées sont largement surestimées. Contrairement au modèle 3I, où les médiateurs sont indépendants, la somme des effets indirects estimés par les deux analyses simples ne correspond pas à l'effet indirect conjoint, montrant ainsi l'une des limites de l'application de multiples analyses simples en parallèle. Au contraire, notre analyse multiple fournit des estimations exactes et précises de l'effet indirect conjoint lorsque les médiateurs sont corrélés. Pour les données simulées à partir du modèle 4C et du modèle 5C, des analyses simples fournissent des estimations erronées des effets indirects et une estimation exacte et précise de l'effet direct. Notre méthode produit des estimations exactes et précises de tous les effets.

	Effects	$\delta^Z$	$PM^Z$	$\delta^1$	$PM^1$	$\delta^2$	$PM^2$	$\zeta$	$\tau$
Mod 3.1	Valeur	27	0.43	3	0.05	24	0.39	35	62
	Analyse simple $M^1$			2.68 [1.98;3.52]	0.04 [0.03;0.06]	23.69 [21.81;25.52]	0.38 [0.36;0.41]	59.22 [58.05;60.34]	61.9 [60.82;63.00]
	Analyse simple $M^2$		0.43 [0.41;0.44]	2.78 [2.26;3.27]	0.05 [0.04;0.05]	23.85 [22.7;24.97]	0.38 [0.37;0.40]	38.2 [36.64;39.84]	61.89 [60.86;62.98]
Mod 3.C	Analyse multiple	27	0.43	3	0.05	24	0.39	35	62
	Valeur			8.30 [6.95;9.72]	0.13 [0.11;0.15]	34.83 [33.21;36.5]	0.56 [0.55;0.58]	53.6 [53.04;54.24]	61.9 [60.39;63.36]
	Analyse simple $M^1$		0.44 [0.42;0.46]	2.94 [2.35;3.58]	0.05 [0.04;0.06]	24.13 [22.33;25.95]	0.39 [0.36;0.42]	27.06 [26.16;27.99]	61.9 [60.45;63.32]
Mod 4.1	Analyse simple $M^2$	0.23	0.55	0.11	0.27	0.11	0.27	0.18	0.41
	Valeur			0.09 [0.05;0.15]	0.23 [0.11;0.39]	0.13 [0.07;0.19]	0.31 [0.17;0.49]	0.3 [0.19;0.4]	0.4 [0.28;0.49]
	Analyse simple $M^1$		0.59 [0.39;0.82]	0.09 [0.04;0.14]	0.23 [0.11;0.42]	0.13 [0.08;0.18]	0.33 [0.2;0.54]	0.28 [0.16;0.4]	0.41 [0.29;0.51]
Mod 4.C	Analyse multiple	20	0.55	0.09	0.24	0.09	0.24	0.16	0.36
	Valeur			0.19 [0.11;0.27]	0.55 [0.35;0.79]	0.2 [0.12;0.27]	0.54 [0.33;0.75]	0.16 [0.06;0.26]	0.35 [0.23;0.46]
	Analyse simple $M^1$		0.56 [0.4;0.8]	0.08 [0.02;0.16]	0.21 [0.04;0.46]	0.12 [0.05;0.2]	0.32 [0.12;0.65]	0.18 [0.07;0.28]	0.38 [0.25;0.48]
Mod 5.1	Analyse simple $M^2$	0.18	0.55	0.09	0.27	0.09	0.27	0.15	0.32
	Valeur			0.11 [0.05;0.17]	0.32 [0.16;0.56]	0.08 [0.03;0.13]	0.23 [0.1;0.44]	0.23 [0.11;0.35]	0.33 [0.21;0.46]
	Analyse simple $M^1$		0.57 [0.35;0.93]	0.1 [0.05;0.17]	0.31 [0.15;0.61]	0.08 [0.03;0.13]	0.23 [0.09;0.48]	0.25 [0.11;0.37]	0.32 [0.19;0.44]
Mod 5.C	Analyse multiple	0.16	0.55	0.07	0.25	0.07	0.25	0.13	0.30
	Valeur			0.15 [0.1;0.2]	0.51 [0.34;0.75]	0.14 [0.09;0.18]	0.48 [0.31;0.7]	0.14 [0.05;0.23]	0.29 [0.19;0.38]
	Analyse simple $M^1$		0.53 [0.37;0.77]	0.08 [0.03;0.15]	0.28 [0.1;0.59]	0.07 [0.01;0.12]	0.23 [0.04;0.48]	0.15 [0.06;0.24]	0.29 [0.2;0.38]
Mod 5.C	Analyse simple $M^2$			0.08 [0.03;0.15]	0.28 [0.1;0.59]	0.07 [0.01;0.12]	0.23 [0.04;0.48]	0.14 [0.05;0.22]	0.29 [0.2;0.38]
	Analyse multiple								

TABLE 4.7 – Comparaison entre les analyses de médiation simple un médiateur à la fois et notre méthode d'analyse de médiation multiple; estimations et intervalles de confiance à 95% des effets de médiation (en colonnes) pour chaque modèle de simulation (en blocs de quatre lignes). Les lignes *Valeur* contiennent les vraies valeurs, calculées à partir d'un grand nombre de contrefactuels simulés. Les lignes *Analyse simple  $M^k$*  et *Analyse multiple* donnent respectivement les estimations de l'analyse simple faite pour le médiateur  $k$  et de notre méthode prenant en compte tous les médiateurs.

## B Grande Dimension

### B.1 Propriétés du graphe laplacien non normalisé

La matrice  $L$  a la forme suivante :

$$\begin{pmatrix} L_1 & & & \\ & L_2 & & \\ & & \ddots & \\ & & & L_K \end{pmatrix}$$

où chaque sous matrice  $L_k$  est une matrice du sous-graphe laplacien non normalisé correspondant à chaque bloc. Chaque sous matrice  $L_k$  a donc les propriétés suivantes :

- symétrique
- semi définie positive
- sa plus petite valeur propre est 0 à laquelle est associé le vecteur propre  $\mathbb{1} \in \mathbb{R}^{P_k}$
- Pour tout vecteur  $f$  on a :

$$f'Lf = \frac{1}{2} \sum_{i,j=1}^Q w_{ij}(f_i - f_j)^2$$

Plaçons nous dans le cas où tous les médiateurs appartiennent au même bloc. Supposons que  $f$  soit le vecteur propre associé à la valeur propre 0, on a :

$$\begin{aligned} Lf &= 0 \\ f'Lf &= 0 \\ \frac{1}{2} \sum_{i,j=1}^Q w_{ij}(f_i - f_j)^2 &= 0 \end{aligned}$$

$w_{ij}$  étant strictement positif, cette dernière égalité n'est possible que si  $f_i = f_j$  pour tout  $i$  et  $j$ . Un vecteur propre avec des valeurs constantes est donc indicatrice que tous les éléments sont connectés. Le nombre d'élément connecté dans le graphe est la dimension du noyau de  $L_k$ . Ainsi 0 est de multiplicité  $K$  le nombre de bloc.

### B.2 MAHI en présence de plusieurs traitements

Un des avantages de notre méthode de sélection est la prise en compte de plusieurs traitements, c'est-à-dire lorsque  $P > 1$ . Notre méthode permet de sélectionner les médiateurs communs de plusieurs traitements identifiés sur la variable de réponse.

Pour cela nous réorganisons nos données de la façon suivante :

—  $\mathbf{z}$  le vecteur de toutes les observations, de taille  $(K+1)I$  :  $\mathbf{z} = \begin{pmatrix} \mathbf{M}_1 \\ \vdots \\ \mathbf{M}_K \\ \mathbf{Y} \end{pmatrix}$ ,

avec  $\mathbf{M}_k = \begin{pmatrix} M_{k,1} \\ \vdots \\ M_{k,I} \end{pmatrix}$  le vecteur des observations du  $k$ -ième médiateur et  $\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_I \end{pmatrix}$

le vecteur des observations de la variable de réponse.

—  $X$  la matrice des variables explicatives de taille  $(K+1)I \times (P+1)K + P$  définie

$$\text{par : } X = \begin{pmatrix} \mathbf{T} & \mathbf{0}_{I \times P} & \dots & \mathbf{0}_{I \times P} & \mathbf{0}_{I \times K} & \mathbf{0}_{I \times P} \\ \mathbf{0}_{I \times P} & \mathbf{T} & \dots & \mathbf{0}_{I \times P} & & \\ \vdots & & & \vdots & \vdots & \vdots \\ \mathbf{0}_{I \times P} & \mathbf{0}_{I \times P} & \dots & \mathbf{T} & \mathbf{M} & \mathbf{0}_{I \times P} \end{pmatrix},$$

avec  $\mathbf{T} = \begin{pmatrix} T_{1,1} & T_{2,1} & \dots & T_{P,1} \\ \vdots & \vdots & \dots & \vdots \\ T_{1,I} & T_{2,I} & \dots & T_{P,I} \end{pmatrix}$  le vecteur des observations du traitement,  $\mathbf{0}_I$  le vecteur

nul de taille  $I$ ,  $\mathbf{0}_{I \times K}$  la matrice nulle de taille  $I \times K$ ,  $\mathbf{0}_{I \times P}$  la matrice nulle de taille  $I \times P$  et  $\mathbf{M}$  la matrice de terme général  $M_{k,i}$ .

En remarquant que  $\mathbf{z} - X\theta = \begin{pmatrix} \mathbf{M}_1 - \sum_p^P \alpha_{p,1} \mathbf{T}_p \\ \vdots \\ \mathbf{M}_K - \sum_p^P \alpha_{p,K} \mathbf{T}_p \\ \mathbf{Y} - \sum_k^K \alpha_k \mathbf{M}_k - \sum_p^P \alpha_p \mathbf{T}_p \end{pmatrix}$  on a l'égalité suivante :

$$\sum_k^K \|\mathbf{M}_k - \sum_p^P \alpha_{p,k} \mathbf{T}_p\|^2 + \|\mathbf{Y} - \sum_k^K \beta_k \mathbf{M}_k - \sum_p^P \alpha_p \mathbf{T}_p\|^2 = \|\mathbf{z} - X\theta\|^2$$

On peut donc réécrire notre problème (2.3.2) et appliquer la méthode de Group-LASSO :

$$\hat{\theta}_\lambda^{GL} = \underset{\theta \in \mathbb{R}^{(P+1)K+P}}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\mathbf{z} - X\theta\|_2^2 + \lambda \sum_{1 \leq k \leq K} \sqrt{P+1} \|\theta_k\|_2 \right\}, \quad (\text{B.1})$$

avec  $\theta_k = (\alpha_k, \beta_k)$

On peut alors étendre l'algorithme présenté dans la section 2.3.1

**Algorithme.** *L'algorithme MAHI se déroule en deux étapes :*

1. *Stability selection.* À partir des données on va générer un nombre  $n_{boot}$  de données bootstrap. Pour chaque données bootstrap on va utiliser la méthode Group-LASSO telle que définit dans l'équation (2.3.4). On peut alors compter  $N_k$ , le nombre de fois où le médiateur  $M_k$  a été sélectionné sur les  $n_{boot}$  fois. On peut alors ranger



les médiateurs par rapport à  $N_k$ . On choisit alors les  $K_{\max}$  premiers médiateurs, où  $K_{\max}$  est un nombre choisi arbitrairement.

2. *Test multiple.* En considérant un traitement à la fois on utilise notre fonction  $R$  multimediate, pour réaliser une analyse de médiation multiple avec les  $K_{\max}$  médiateurs. La fonction multimediate nous permet d'avoir une  $p$ -valeur pour chaque effet médié. La méthode utilisée pour calculer cette  $p$ -valeur est présentée dans le Chapitre 3. On a donc un vecteur de  $p$ -valeur de longueur  $K_{\max}$ , noté  $P_{\text{vect}}$  auquel on applique la correction de Bonferroni :

$$P_{\text{corr},k} = \min(P_{km} \times K_{\max}, 1),$$

où  $P_{km}$  est la  $p$ -valeur du  $k$ -ième élément de  $P_{\text{vect}}$

Si  $P_{\text{corr},k} < 0,05$ , nous pouvons conclure qu'il existe un effet de médiation significatif passant par  $M_k$ .

On répète cette étape pour chaque traitement et on garde les médiateurs significatifs pour tous les traitements identifiés.

### B.3 Méthode HIMA

Zhang et al. (2016) propose une approche HIMA (HIgh dimensional Mediation Analysis), qui repose sur des régressions pénalisées et utilise une fonction de pénalité de type LASSO appelée pénalité concave minimax (MCP) (Zhang (2010)).

**Algorithme.** L'algorithme HIMA se déroule en trois étapes :

1. La première étape consiste à réduire le nombre initial de médiateurs potentiels d'un très grand à un nombre raisonnable (c'est-à-dire plus petit que la taille de l'échantillon) en utilisant le sure independence screening (SIS) décrite dans Fan and Lv (2008). La procédure de sure independence screening est la suivante : soit  $\omega \in \mathbb{R}^K$  le vecteur défini par

$$\omega = M^\Gamma Y,$$

où chaque colonne de la matrice  $M$  est normalisée, c'est-à-dire que pour tout  $k = 1, \dots, K$  la colonne  $M_k$  a pour moyenne 0 et variance 1.

$\bar{Y}$  et  $\bar{M}_k$  désignent les moyennes respectives de  $Y$  et  $M_k$ , et  $sd(Y)$  et  $sd(M_k)$  désignent les écarts-types respectifs de  $Y$  et  $M_k$ , alors nous pouvons réécrire pour tout  $k$ ,

$$\begin{aligned} \omega_k &= (M^\Gamma Y)_k \\ &= \sum_{i=1}^n (M^\Gamma)_{ki} Y_i = \sum_{i=1}^n M_{ik} Y_i \\ &= \sum_{i=1}^n (M_{ik} - \bar{M}_k) (Y_i - \bar{Y}) + \sum_{i=1}^n M_{ik} \bar{Y} \\ &= \text{cov}(M_k, Y) + n \bar{Y} \bar{M}_k \\ &= \text{corr}(M_k, Y) sd(M_k) sd(Y) = \text{corr}(M_k, Y) sd(Y), \end{aligned}$$

où  $\omega_k$  est le  $k$ -ième élément de  $\omega$ .

On voit alors que  $\omega$  est le vecteur des corrélations de médiateurs avec  $Y$ , réajustées par l'écart-type de  $Y$ . Les  $K$  éléments de  $\omega$  sont ensuite ordonnés par ordre décroissant. On obtient ainsi un rang pour chaque médiateurs selon leur corrélations avec  $Y$  et on élimine ceux qui ont une corrélation faible avec  $Y$ . On se retrouve avec l'ensemble de médiateurs  $\mathcal{I}$  définit comme l'ensemble  $\{1 \leq k \leq K : |\omega_k| \text{ est parmi les } d = \lceil 2n/\log(n) \rceil \text{ premiers médiateurs}\} = \{1 \leq k \leq K : M_k \text{ est parmi les } d = \lceil 2n/\log(n) \rceil \text{ plus fort médiateurs affectant } Y\}$ .

Toutes les variables importantes survivent à ce premier filtre avec une probabilité qui tend vers 1. La procédure est dite indépendante puisque chaque variable  $M_k$  est utilisée indépendamment comme prédicteur pour décider de son importance pour prédire la variable de réponse  $Y$ .

2. La seconde étape d'HIMA consiste à sélectionner les médiateurs en utilisant la pénalité concave minimax (MCP). Les estimations  $\hat{\beta}_k$ ,  $k \in \mathcal{I}$  sont calculées en minimisant le critère de pénalisation MCP (Zhang (2010)),

$$\mathcal{Q}^{ols} = \sum_{i=1}^n (Y_i - \mu_0 - \alpha_0 T_i - \sum_{k \in \mathcal{I}} \beta_k M_{ik})^2 + \sum_{k \in \mathcal{I}} p_{\lambda, \delta}(\beta_k), \quad (\text{B.2})$$

où  $p_{\lambda, \delta}(\cdot)$  est la fonction MCP définit par :

$$p_{\lambda, \delta}(\beta_k) = \lambda \left[ |\beta_k| - \frac{|\beta_k|^2}{2\delta\lambda} \right] \mathbb{1}_{\{0 \leq |\beta_k| < \delta\lambda\}} + \frac{\lambda^2 \delta}{2} \mathbb{1}_{\{|\beta_k| \geq \delta\lambda\}},$$

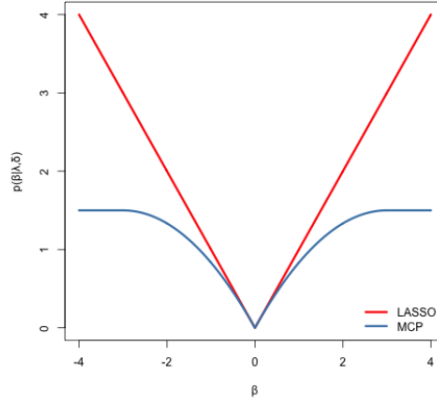
où  $\lambda > 0$  est le paramètre de régularisation, et  $\delta > 0$  détermine la concavité du MCP. Plus précisément, la pénalité concave minimax commence par appliquer le même taux de pénalisation que le LASSO, puis relâche progressivement jusqu'à un taux constant à mesure que la valeur absolue du coefficient  $\beta_k$  augmente. La Figure 4 affiche les fonctions LASSO et MCP pour les valeurs fixes de  $\lambda$  et  $\delta$ .

Pour une séquence de  $\lambda$  donnée, l'algorithme choisit le  $\lambda$  qui minimise le critère d'information bayésienne (BIC) du modèle de régression correspondant à l'équation (B.2). Pour rappel, le critère BIC est définit par :

$$BIC = \ln(n)k - 2\ln(\hat{L})$$

où  $n$  est le nombre d'observation,  $k = |\mathcal{I}|$  est le nombre de variable (c'est-à-dire le nombre de médiateur) estimé par le modèle, et  $\hat{L}$  est la valeur maximisée de la fonction de vraisemblance du modèle.

3. La troisième étape consiste à effectuer des test multiples pour les effets médiés. Soit  $\mathcal{S} = \{1 \leq k \leq K : \hat{\beta}_k \neq 0\}$  l'ensemble de médiateur obtenu par la procédure de pénalisation MCP. La correction de Bonferroni pour un test multiple est défini par :

FIGURE 4.9 – LASSO et MCP en fonction de  $\beta$  pour  $\lambda = 1$  et  $\delta = 3$ 

$$P_{corr,k} = \max(P_{corr,1k}, P_{corr,2k})$$

où les  $p$ -valeurs  $P_{corr,1k}$  et  $P_{corr,2k}$  sont définies par :

$$P_{corr,1k} = \min\{P_{raw,1k} \cdot |\mathcal{S}|, 1\}$$

avec  $P_{raw,1k}$  la  $p$ -valeur pour tester l'hypothèse nulle  $H_0 : \beta_k = 0$  donné par

$$P_{raw,1k} = 2 \left\{ 1 - \Phi \left( \frac{|\widehat{\beta}_k|}{\widehat{\sigma}_{1k}} \right) \right\},$$

De façon analogue on a :

$$P_{corr,2k} = \min\{P_{raw,2k} \cdot |\mathcal{S}|, 1\}$$

avec  $P_{raw,2k}$  la  $p$ -valeur pour tester l'hypothèse nulle  $H_0 : \alpha_k = 0$  donné par

$$P_{raw,2k} = 2 \left\{ 1 - \Phi \left( \frac{|\widehat{\alpha}_k|}{\widehat{\sigma}_{2k}} \right) \right\},$$

où  $k \in \mathcal{S}$ ,  $\Phi$  est la fonction de distribution d'une loi normal  $\mathcal{N}(0,1)$ , et  $\widehat{\sigma}_{1k}$  et  $\widehat{\sigma}_{2k}$  sont les erreurs standards de  $\widehat{\beta}_k$  et  $\widehat{\alpha}_k$ . Notons que les  $\widehat{\beta}_k$  sont obtenus à l'étape 2 (estimateurs des moindres carrés avec la pénalité MCP), et que les  $\widehat{\alpha}_k$  sont les estimateurs du moindre carrés pour  $\alpha_k$ ,  $k = 1, \dots, K$ . De plus,  $\widehat{\sigma}_{1k}$  et  $\widehat{\sigma}_{2k}$  peuvent être obtenus à partir de la propriété oracle de MCP (Zhang (2010), Fan and Li (2001))

Si  $P_{corr,k} < 0,05$ , nous pouvons conclure qu'il existe un effet de médiation significatif pour  $M_k$  sur le outcome  $Y$ ,  $k \in \mathcal{S}$ . C'est-à-dire que l'hypothèse nulle d'absence d'effet de médiation avec  $M_k$  n'est rejetée que si  $\alpha_k$  et  $\beta_k$  sont tous deux significatifs.

L'ensemble de la procédure est implémenté dans le package R `hima`.

## B.4 Méthode CMF

van Kesteren and Oberski (2019) propose une autre approche qui permet de faire de l'analyse de médiation en grande dimension, appelée la méthode du *coordinate-wise mediation filter* (CMF, filtre de médiation par coordonnées). L'implémentation CMF se compose de deux composants : un algorithme interne qui effectue la sélection des médiateurs par descente de coordonnées à l'aide d'une fonction de décision  $\mathcal{D}$ , et un algorithme externe qui effectue plusieurs lancement aléatoires de l'algorithme interne et l'agrégation ultérieure des multiples sorties. Avant de présenter les boucles internes et externes de l'algorithme CMF, nous introduisons d'abord le test Sobel qui sera utilisé comme fonction de décision  $\mathcal{D}$ .

**Le test de Sobel** Le test de Sobel permet de tester la significativité de l'effet de médiation (Sobel (1986)) dans le cadre du modèle (2.3.1) avec  $P = 1$ . Plus précisément, on teste l'hypothèse de nullité suivante :

$$H_0 : \alpha_k \beta_k = 0 \text{ vs } H_1 : \alpha_k \beta_k \neq 0,$$

utilisant le test statistique suivant :

$$t_{Sobel} = \frac{|\hat{\alpha}_k \hat{\beta}_k|}{\sqrt{\hat{\alpha}_k^2 \hat{\sigma}_\beta^2 + \hat{\beta}_k^2 \hat{\sigma}_\alpha^2}}$$

où  $\hat{\sigma}_\alpha^2$  et  $\hat{\sigma}_\beta^2$  sont les estimations respectives des variances de  $\hat{\alpha}_k$  et  $\hat{\beta}_k$ . Le dénominateur est l'erreur-type estimée de l'effet indirect  $\hat{\alpha}_k \hat{\beta}_k$ .

Le test statistique  $t_{Sobel}$  peut alors être comparée à une distribution normale standard  $\mathcal{N}(0,1)$  afin de déterminer sa significativité. Ainsi, la p-valeur suivante peut être obtenue :

$$P_{Sobel} = 2 \left\{ 1 - \Phi \left( \frac{|\hat{\alpha}_k \hat{\beta}_k|}{\sqrt{\hat{\alpha}_k^2 \hat{\sigma}_\beta^2 + \hat{\beta}_k^2 \hat{\sigma}_\alpha^2}} \right) \right\}$$

Notez que le test de Sobel est utilisé comme fonction de décision  $\mathcal{D}$ , mais cette dernière pourrait être toute autre fonction de décision spécifiée par l'utilisateur. Il est à noter que le produit des coefficients  $\alpha_k \beta_k$  est une estimation de l'effet indirect médié par  $M_k$  dans le cas du modèle (2.3.1) avec  $P = 1$ .

**Algorithme Interne** L'objectif de l'algorithme interne est le suivant. Dans un premier temps, un vecteur de décision  $P$  de longueur  $K$  (c'est-à-dire le nombre total de médiateurs) contenant la sélection actuelle de médiateurs est initialisé au hasard avec des valeurs 0 et 1. Puis à chaque étape de l'algorithme interne,  $\sqrt{K}$  médiateurs sont sélectionnés au hasard parmi tous les médiateurs et le vecteur de décision est mis à jour comme suit. Pour chaque médiateur  $M_k$  parmi les  $\sqrt{K}$  sélectionnés, une matrice de données  $M_*$ , qui contient tous les médiateurs actuellement sélectionnés (ce sont des médiateurs de valeur correspondante 1 dans le vecteur de décision  $P$ ) à l'exclusion de la variable  $M_k$  considérée, est créée. Ensuite, les résidus de l'exposition  $T$  et de la variable de réponse  $Y$  par rapport à  $M_*$  sont calculés, c'est-à-dire :

$$\begin{aligned} R_T &= T - M_*(M_*^\Gamma M_*)^{-1} M_*^\Gamma T \\ R_Y &= Y - M_*(M_*^\Gamma M_*)^{-1} M_*^\Gamma Y \end{aligned} \quad (\text{B.3})$$

Enfin la décision binaire pour savoir si  $M_k$  est sélectionné par rapport à la matrice  $M_*$ , est prise en utilisant le test de Sobel. Si  $P_{Sobel} < 0.1$  alors l'hypothèse nulle est rejetée, ainsi l'effet indirect du médiateur  $M_k$  est significatif à 10%, et la  $k$ -ième composante du vecteur de décision  $P$  est mis à la valeur 1. De même dans le cas où  $P_{Sobel} \leq 0.1$ , la  $k$ -ième composante du vecteur  $P$  est mise à 0. Notons que c'est à cette étape que  $\alpha_k$  et  $\beta_k$  sont estimés.

Le paramètre  $\alpha_k$  est estimé en résolvant le système d'équation :

$$R_T^\Gamma R_T \alpha_k = R_T^\Gamma M_k$$

De plus, le paramètre  $\beta_k$  est estimé en résolvant le système d'équation suivant :

$$[R_T M_k]^\Gamma [R_T M_k] \beta_k = [R_T M_k]^\Gamma R_Y,$$

où  $[R_T M_k]$  est la matrice résultant de la concaténation des vecteurs  $R_T$  et  $M_k$ .

**Algorithme Externe** La valeur du vecteur de décision résultant de l'algorithme interne dépend dans une certaine mesure des valeurs initiales, en raison de la nature discrète de ses coordonnées. Par conséquent, l'algorithme est intégré dans une boucle externe qui effectue un nombre prédéfini de boucle. Lorsque toutes les boucles sont terminées, un taux de sélection  $S_k$  est calculé pour chaque médiateur  $M_k$ . Il est défini comme :

$$S_k = \frac{\sum_{i=1}^N \mathbb{1}_{\{M_k \text{ est sélectionné à la boucle } i\}}}{N}$$

où nous avons noté  $N$  le nombre de boucles effectuées par la boucle externe. Le taux de sélection  $S_k$  représente la proportion de fois où le médiateur potentiel  $M_k$  a été sélectionné par l'algorithme interne. Ces proportions, ou probabilités de sélection empiriques, conduisent naturellement à un classement de médiateur. Ce classement est ensuite à nouveau dichotomisé en utilisant une valeur seuil qui doit être spécifiée par l'utilisateur. Tous

les médiateurs dont le taux de sélection est strictement supérieur à la valeur seuil sont classés comme vrais médiateurs et ceux dont le taux de sélection est inférieur ou égal au seuil sont classés comme faux médiateurs. Il n'y a pas de procédure proposée pour le choix de la coupure dans l'article qui présente la méthode CMF (van Kesteren and Oberski (2019)). Les auteurs de van Kesteren and Oberski (2019) ont arbitrairement fixé le seuil à 0,075 dans leur exemple d'application à des données réelles. L'algorithme interne s'arrête soit lorsque le vecteur décision ne change pas entre deux étapes soit lorsque nombre d'itération maximal est atteint.

L'ensemble de la procédure CMF est implémentée dans le package R `cmfilter`, disponible sur <https://github.com/vankesteren/cmfilter>.

## C Package R

### C.1 P-valeur

L'algorithme quasi-bayésien présenté dans la section 3.2.3 permet d'obtenir la distribution empirique pour les effet causaux d'intérêt. Dans cette section on montre que, à l'aide de cette distribution, on peut aussi tester la significativité de ces effets. En particulier la  $p$ -valeur peut être calculée en analogie avec la construction de  $p$ -valeurs par bootstrap.

Notons  $\delta^*$  l'échantillon  $(\delta^{(1)}, \dots, \delta^{(J)})$  obtenu à l'issue de notre algorithme et qui nous donne une estimation de la distribution de  $\delta$ . Un estimateur ponctuel de  $\delta$  est donné par la moyenne  $\bar{\delta}^*$ . On souhaite tester les hypothèses  $H_0 : \delta = 0$  contre  $H_1 : \delta \neq 0$  à l'aide de  $\delta^*$ . Pour cela nous avons besoin de connaître la distribution de  $\delta^*$  sous  $H_0$ . Sous l'hypothèse d'équivariance translationnelle, cette distribution est donnée simplement en translatant  $\delta^*$  de façon à que la moyenne de l'échantillon se trouve au centre, c'est à dire en prenant  $\delta^* - \bar{\delta}^*$ . On peut définir la  $p$ -valeur de ce test par

$$\mathbb{P}(|\delta^* - \bar{\delta}^*| \geq |\bar{\delta}^*|) = 2 \min(\mathbb{P}(\delta^* \geq 0), \mathbb{P}(\delta^* \leq 0))$$

Cette définition est justifiée par le même argument justifiant le calcul des  $p$ -valeurs par bootstrap. En particulier, on peut définir la  $p$ -valeur par  $\mathbb{P}(|\delta^* - \bar{\delta}^*| \geq |\delta_0^*|)$  où  $\delta_0^*$  est la réalisation de  $\delta^*$  obtenue à partir de notre algorithme.  $\delta_0^*$  correspond donc à la valeur de la statistique de référence calculée à partir des données originales dans la construction classique de la  $p$ -valeur par bootstrap. Puisque  $\delta_0^* \approx \bar{\delta}^*$  on obtient la définition proposée ci-dessus.

Une estimation de cette  $p$ -valeur est donc donnée par

$$\frac{2}{J} \min \left( \sum_{j=1}^J \mathbf{1}_{\{\delta^* \geq 0\}}, \sum_{j=1}^J \mathbf{1}_{\{\delta^* < 0\}} \right).$$

Cette  $p$ -valeur est valable pour l'ensemble des quantités d'intérêt.

L'objectif dans les sections suivantes est d'estimer la matrice de variance covariance  $\Sigma$  qui a permis de générer les médiateurs. On notera  $M_i$  les médiateurs que l'on peut observer. Et  $M_i^*$  les médiateurs non-accessibles dans les données observables mais qui sont utilisés pour générer les  $M_i$ .

## C.2 Médiateurs catégoriels ordonnés

$$\left( M_1^*, M_2^* \mid T = t, C = c \right) \sim \mathcal{N} \left( \mu = \begin{pmatrix} \alpha_1 + \beta_1 t + \gamma_1 c \\ \alpha_2 + \beta_2 t + \gamma_2 c \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \right)$$

On obtient les médiateurs catégoriels en appliquant la discrétisation suivante :

$$\begin{aligned} M_1 &= \sum_{q=1}^Q q \mathbb{1}_{\{a_q < M_1^* \leq a_{q+1}\}} \\ M_2 &= \sum_{p=1}^P p \mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}} \end{aligned}$$

avec  $a_1 = a'_1 = -\infty$  et  $a_{Q+1} = a'_{P+1} = \infty$ .

Pour estimer  $\Sigma$  il suffit d'estimer  $\rho = Cov(M_1^*, M_2^* \mid T = t, C = c)$  qui n'est pas observable. Mais on peut calculer  $Cov(M_1, M_2 \mid T = t, C = c)$  sur les données et l'exprimer en fonction de  $\rho$ , en effet on a :

$$\begin{aligned}
Cov(M_1, M_2|T = t, C = c) &= \mathbb{E}[M_1 M_2|T = t, C = c] - \mathbb{E}[M_1|T = t, C = c]\mathbb{E}[M_2|T = t, C = c] \\
&= \sum_{q=1}^Q \sum_{p=1}^P qp \mathbb{P}(\mathbb{1}_{\{a_q < M_1^* \leq a_{q+1}\}}, \mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}}) \\
&\quad - \sum_{q=1}^Q q \mathbb{P}(\mathbb{1}_{\{a_q < M_1^* \leq a_{q+1}\}}) \times \sum_{p=1}^P p \mathbb{P}(\mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}}) \\
&= \sum_{q=1}^Q \sum_{p=1}^P qp \mathbb{P}(\mathbb{1}_{\{a_q < \alpha_1 + \beta_1 t + \gamma_1 c + \epsilon_1 \leq a_{q+1}\}}, \mathbb{1}_{\{a_p < \alpha_2 + \beta_2 t + \gamma_2 c + \epsilon_2 \leq a_{p+1}\}}) \\
&\quad - \sum_{q=1}^Q q \mathbb{P}(\mathbb{1}_{\{a_q < \alpha_1 + \beta_1 t + \gamma_1 c + \epsilon_1 \leq a_{q+1}\}}) \times \sum_{p=1}^P p \mathbb{P}(\mathbb{1}_{\{a_p < \alpha_1 + \beta_1 t + \gamma_1 c + \epsilon_1 \leq a_{p+1}\}}) \\
&= \sum_{q=1}^Q \sum_{p=1}^P qp \mathbb{P}(\mathbb{1}_{\{BI_q < \epsilon_1 \leq BS_q\}}, \mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}) \\
&\quad - \sum_{q=1}^Q q \mathbb{P}(\mathbb{1}_{\{BI_q < \epsilon_1 \leq BS_q\}}) \times \sum_{p=1}^P p \mathbb{P}(\mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}) \\
&= \sum_{q=1}^Q \sum_{p=1}^P qp \Phi_{(0, \Sigma)}(\{BI_q, BS_q\}, \{BI_p, BS_p\}) \\
&\quad - \sum_{q=1}^Q q \left( \Phi_{(0,1)}(BS_q) - \Phi_{(0,1)}(BI_q) \right) \times \sum_{p=1}^P p \left( \Phi_{(0,1)}(BS_p) - \Phi_{(0,1)}(BI_p) \right)
\end{aligned}$$

La différence des deux termes de cette égalité est nulle, on peut approcher la valeur de chaque  $\rho$ , en approchant la valeur qui annule la fonction :

$$\begin{aligned}
f(\rho) &= Cov(M_1, M_2|T = t, C = c) - \sum_{q=1}^Q \sum_{p=1}^P qp \Phi_{(0, \Sigma)}(BS_q, BI_q, BS_p, BI_p) \\
&\quad + \sum_{q=1}^Q q \left( \Phi_{(0,1)}(BS_q) - \Phi_{(0,1)}(BI_q) \right) \times \sum_{p=1}^P p \left( \Phi_{(0,1)}(BS_p) - \Phi_{(0,1)}(BI_p) \right)
\end{aligned}$$

avec  $\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ ,  $BI_q = a_q - (\alpha_1 + \beta_1 t + \gamma_1 c)$ ,  $BS_q = a_{q+1} - (\alpha_1 + \beta_1 t + \gamma_1 c)$ ,  $BI_p = a_p - (\alpha_2 + \beta_2 t + \gamma_2 c)$  et  $BS_p = a_{p+1} - (\alpha_2 + \beta_2 t + \gamma_2 c)$ .



### C.3 Médiateurs linéaire et catégoriels ordonnés

$$(M_1, M_2^* | T = t, C = c) \sim \mathcal{N}\left(\mu = \begin{pmatrix} \alpha_1 + \beta_1 t + \gamma_1 c \\ \alpha_2 + \beta_2 t + \gamma_2 c \end{pmatrix}, \Sigma = \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

$$M_2 = \sum_{p=1}^P p \mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}}$$

On peut estimer  $\sigma_1^2$  en calculant la variance des résidus de la régression linéaire de  $M_1$ .

On ne peut pas observer  $\rho = Cov(M_1, M_2^* | T = t, C = c)$  mais on peut calculer  $Cov(M_1, M_2 | T = t, C = c)$  sur les données et l'exprimer en fonction de  $\rho$ , en effet on a :

$$\begin{aligned} Cov(M_1, M_2 | T = t, C = c) &= \mathbb{E}[M_1 M_2 | T = t, C = c] - \mathbb{E}[M_1 | T = t, C = c] \mathbb{E}[M_2 | T = t, C = c] \\ &= \sum_{p=1}^P p \mathbb{E}[M_1 \mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}} | T = t, C = c] \\ &\quad - \mathbb{E}[M_1 | T = t, C = c] \sum_{p=1}^P p \mathbb{E}[\mathbb{1}_{\{a_p < M_2^* \leq a_{p+1}\}} | T = t, C = c] \\ &= \sum_{p=1}^P p \mathbb{E}[(\alpha_1 + \beta_1 t + \gamma_1 c + \epsilon_1) \mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &\quad - \mathbb{E}[(\alpha_1 + \beta_1 t + \gamma_1 c + \epsilon_1)] \sum_{p=1}^P p \mathbb{E}[\mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &= \sum_{p=1}^P p \mathbb{E}[(\alpha_1 + \beta_1 t + \gamma_1 c + \epsilon_1) \mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &\quad - (\alpha_1 + \beta_1 t + \gamma_1 c) \sum_{p=1}^P p \mathbb{E}[\mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &= \sum_{p=1}^P p (\alpha_1 + \beta_1 t + \gamma_1 c) \mathbb{E}[\mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] + \sum_{p=1}^P p \mathbb{E}[\epsilon_1 \mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &\quad - (\alpha_1 + \beta_1 t + \gamma_1 c) \sum_{p=1}^P p \mathbb{E}[\mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &= \sum_{p=1}^P p \mathbb{E}[\epsilon_1 \mathbb{1}_{\{BI_p < \epsilon_2 \leq BS_p\}}] \\ &= \sum_{p=1}^P p \int_{\mathbb{R}} \int_{BI_p}^{BS_p} e_1 f_{(0, \Sigma)}(e_1, e_2) de_2 de_1 \end{aligned} \tag{C.1}$$

Pour développer le deuxième terme de l'égalité rappelons que  $\Sigma = \begin{pmatrix} \sigma_1^2 & \rho \\ \rho & 1 \end{pmatrix}$ ,  $|\Sigma| = \sigma_1^2 - \rho^2$  et  $\Sigma^{-1} = \frac{1}{\sigma_1^2 - \rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & \sigma_1^2 \end{pmatrix}$ . Ainsi on peut exprimer la fonction  $f_{(0,\Sigma)}(e_1, e_2)$ , on obtient alors :

$$\begin{aligned}
& \int_{\mathbb{R}} \int_{BI_p}^{BS_p} e_1 f_{(0,\Sigma)}(e_1, e_2) de_2 de_1 \\
= & \int_{\mathbb{R}} \int_{BI_p}^{BS_p} e_1 \frac{1}{2\pi\sqrt{(\sigma_1^2 - \rho^2)}} \exp\left(-\frac{e_1^2 - 2e_1e_2\rho + e_2^2\sigma_1^2}{2(\sigma_1^2 - \rho^2)}\right) de_2 de_1 \\
= & \int_{\mathbb{R}} \int_{BI_p}^{BS_p} e_1 \frac{1}{2\pi\sqrt{(\sigma_1^2 - \rho^2)}} \exp\left(-\frac{(e_1 - \rho e_2)^2 - (\rho e_2)^2 + \sigma_1^2 e_2^2}{2(\sigma_1^2 - \rho^2)}\right) de_2 de_1 \\
= & \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(\rho e_2)^2 - \sigma_1^2 e_2^2}{2(\sigma_1^2 - \rho^2)}\right) \int_{\mathbb{R}} e_1 \frac{1}{\sqrt{2\pi(\sigma_1^2 - \rho^2)}} \exp\left(-\frac{(e_1 - \rho e_2)^2}{2(\sigma_1^2 - \rho^2)}\right) de_1 de_2
\end{aligned}$$

En observant que la loi conditionnelle  $\epsilon_1 | \epsilon_2 = e_2$  est la loi normale  $\mathcal{N}(\rho e_2, \sigma_1^2 - \rho^2)$ , on peut écrire :

$$\begin{aligned}
\int_{\mathbb{R}} \int_{BI_p}^{BS_p} e_1 f_{(0,\Sigma)}(e_1, e_2) de_2 de_1 &= \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(\rho e_2)^2 - \sigma_1^2 e_2^2}{2(\sigma_1^2 - \rho^2)}\right) \int_{\mathbb{R}} e_1 f_{\epsilon_1 | \epsilon_2 = e_2}(e_1) de_1 de_2 \\
&= \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(\rho e_2)^2 - \sigma_1^2 e_2^2}{2(\sigma_1^2 - \rho^2)}\right) \int_{\mathbb{R}} e_1 f_{(\rho e_2, \sigma_1^2 - \rho^2)}(e_1) de_1 de_2 \\
&= \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(\rho e_2)^2 - \sigma_1^2 e_2^2}{2(\sigma_1^2 - \rho^2)}\right) \mathbb{E}[\epsilon_1 | \epsilon_2 = e_2] de_2 \\
&= \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(\frac{(\rho e_2)^2 - \sigma_1^2 e_2^2}{2(\sigma_1^2 - \rho^2)}\right) \rho e_2 de_2 \\
&= \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{e_2^2(\sigma_1^2 - \rho^2)}{2(\sigma_1^2 - \rho^2)}\right) \rho e_2 de_2 \\
&= \int_{BI_p}^{BS_p} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{e_2^2}{2}\right) \rho e_2 de_2 \\
&= \rho \int_{BI_p}^{BS_p} e_2 f_{(0,1)}(e_2) de_2 \tag{C.2}
\end{aligned}$$

D'après (C.1) et (C.2) on a :

$$\rho = \frac{Cov(M_1, M_2 | T = t, C = c)}{\sum_{p=1}^P p \int_{BI_p}^{BS_p} e_2 f_{(0,1)}(e_2) de_2}$$

avec  $BI_p = a_p - (\alpha_2 + \beta_2 t + \gamma_2 c)$  et  $BS_p = a_{p+1} - (\alpha_2 + \beta_2 t + \gamma_2 c)$ .

# Bibliographie

## Bibliographie

- Abbate-Daga, G., Buzzichelli, S., Amianto, F., Rocca, G., Marzola, E., McClintock, S. M., and Fassino, S. (2011). Cognitive flexibility in verbal and nonverbal domains and decision making in anorexia nervosa patients : a pilot study. *BMC Psychiatry*, 11(1) :162.
- Amianto, F., Ottone, L., Abbate Daga, G., and Fassino, S. (2015). Binge-eating disorder diagnosis and treatment : a recap in front of DSM-5. *BMC Psychiatry*, 15(1) :70.
- Atiye, M., Miettunen, J., and Raevuori-Helkamaa, A. (2015). A Meta-Analysis of Temperament in Eating Disorders : Temperament in Eating Disorders. *European Eating Disorders Review*, 23(2) :89–99.
- Avin, C., Shpitser, I., and Pearl, J. (2005). Identifiability of path-specific effects. In *IJCAI International Joint Conference on Artificial Intelligence*, pages 357–363.
- Baron, R. M. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research : conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6) :1173–1182.
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., and Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry*, 4(6) :561–571.
- Bellavia, A. and Valeri, L. (2017). Decomposition of the total effect in the presence of multiple mediators and interactions. *American journal of epidemiology*, 187(6) :1311–1318.
- Boca, S. M. (2013). "r package". <https://academic.oup.com/bioinformatics/article/30/2/214/218964>.
- Boca, S. M., Sinha, R., Cross, A. J., Moore, S. C., and Sampson, J. N. (2014). Testing multiple biological mediators simultaneously. *Bioinformatics*, 30(2) :214–220.

- Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2) :211–252.
- Boyd, N. F., Martin, L. J., Sun, L., Guo, H., Chiarelli, A., Hislop, G., Yaffe, M., and Minkin, S. (2006). Body Size, Mammographic Density, and Breast Cancer Risk. *Cancer Epidemiology Biomarkers & Prevention*, 15(11) :2086–2092.
- Byrne, C., Ursin, G., Martin, C. F., Peck, J. D., Cole, E. B., Zeng, D., Kim, E., Yaffe, M. D., Boyd, N. F., Heiss, G., McTiernan, A., Chlebowski, R. T., Lane, D. S., Manson, J. E., Wactawski-Wende, J., and Pisano, E. D. (2017). Mammographic Density Change With Estrogen and Progestin Therapy and Breast Cancer Risk. *JNCI : Journal of the National Cancer Institute*, 109(9).
- Cagnacci, A. and Venier, M. (2019). The Controversial History of Hormone Replacement Therapy. *Medicina*, 55(9) :602.
- Collaborative Group on Hormonal Factors in Breast Cancer, C. (2019). Type and timing of menopausal hormone therapy and breast cancer risk : individual participant meta-analysis of the worldwide epidemiological evidence. *The Lancet*, 394(10204) :1159–1168.
- Daniel, R. M., De Stavola, B. L., Cousens, S. N., and Vansteelandt, S. (2015). Causal mediation analysis with multiple mediators. *Biometrics*, 71(1) :1–14.
- Deep, A. L., Nagy, L. M., Weltzin, T. E., Rao, R., and Kaye, W. H. (1995). Premorbid onset of psychopathology in long-term recovered anorexia nervosa. *International Journal of Eating Disorders*, 17(3) :291–297.
- Fan, J. and Li, R. (2001). Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties. *Journal of the American Statistical Association*, 96(456) :1348–1360.
- Fan, J. and Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 70(5) :849–911.
- for Research on Cancer, I. A. and Weltgesundheits organisation, I., editors (2012). *IARC monographs on the evaluation of carcinogenic risks to humans, volume 100 A, pharmaceuticals : this publication represents the views and expert opinions of an IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, which met in Lyon, 14 - 21 October 2008*. IARC, Lyon. Meeting Name : Meeting. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans OCLC : 934857422.
- Forastiere, L., Mealli, F., and VanderWeele, T. J. (2016). Identification and Estimation of Causal Mechanisms in Clustered Encouragement Designs : Disentangling Bed Nets Using Bayesian Principal Stratification. *Journal of the American Statistical Association*, 111(514) :510–525.

- Fornili, M., Perduca, V., Fournier, A., Jérolon, A., Boutron-Ruault, M. C., Maskarinec, G., Severi, G., and Baglietto, L. (2020). Association between menopausal hormone therapy, mammographic density and breast cancer risk : results from the E3N cohort study. preprint, In Review.
- Gao, Y. (2019). " r package hdma". <https://github.com/YuzhaoGao/High-dimensional-mediation-analysis-R>.
- Gao, Y., Yang, H., Fang, R., Zhang, Y., Goode, E. L., and Cui, Y. (2019). Testing Mediation Effects in High-Dimensional Epigenetic Studies. *Frontiers in Genetics*, 10 :1195.
- Garner, D. (1991). *Eating Disorder Inventory2 : Professional manual*.
- Godart, N., Perdereau, F., Rein, Z., Berthoz, S., Wallier, J., Jeammet, P., and Flament, M. (2007). Comorbidity studies of eating disorders and mood disorders. Critical review of the literature. *Journal of Affective Disorders*, 97(1-3) :37–49.
- Godart, N. T., Flament, M. F., Perdereau, F., and Jeammet, P. (2002). Comorbidity between eating disorders and anxiety disorders : A review. *International Journal of Eating Disorders*, 32(3) :253–270.
- Grimonprez, Q., Celisse, A., and Marot, G. (2015). Sélection de groupes de variables corrélées par classification ascendante hiérarchique et group-lasso. In *47èmes Journées de Statistique*, Lille, France.
- Hafeman, D. M. and VanderWeele, T. J. (2011). Alternative Assumptions for the Identification of Direct and Indirect Effects :. *Epidemiology*, 22(6) :753–764.
- Hudson, J. I., Hiripi, E., Pope, H. G., and Kessler, R. C. (2007). The Prevalence and Correlates of Eating Disorders in the National Comorbidity Survey Replication. *Biological Psychiatry*, 61(3) :348–358.
- Imai, K., Keele, L., and Tingley, D. (2010a). A general approach to causal mediation analysis. *Psychological Methods*, 15(4) :309–334.
- Imai, K., Keele, L., and Yamamoto, T. (2010b). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, 25(1) :51–71.
- Imai, K. and Yamamoto, T. (2013). Identification and Sensitivity Analysis for Multiple Causal Mechanisms : Revisiting Evidence from Framing Experiments. *Political Analysis*, 21(02) :141–171.
- James, L., Mulaik, S., and Brett, J. M. (1982). *Causal Analysis : Assumptions, Models, and Data*.

- Jérolon, A., Baglietto, L., Birmelé, E., Alarcon, F., and Perduca, V. (2020). Causal mediation analysis in presence of multiple mediators uncausally related. *The International Journal of Biostatistics*, 0(0).
- Kaye, W. H., Fudge, J. L., and Paulus, M. (2009). New insights into symptoms and neurocircuit function of anorexia nervosa. *Nature Reviews Neuroscience*, 10(8) :573–584.
- Kelemen, L. E., Pankratz, V. S., Sellers, T. A., Brandt, K. R., Wang, A., Janney, C., Fredericksen, Z. S., Cerhan, J. R., and Vachon, C. M. (2008). Age-specific Trends in Mammographic Density : The Minnesota Breast Cancer Family Study. *American Journal of Epidemiology*, 167(9) :1027–1036.
- Kendler, K. S. (1995). The Structure of the Genetic and Environmental Risk Factors for Six Major Psychiatric Disorders in Women : Phobia, Generalized Anxiety Disorder, Panic Disorder, Bulimia, Major Depression, and Alcoholism. *Archives of General Psychiatry*, 52(5) :374.
- Kezelman, S., Touyz, S., Hunt, C., and Rhodes, P. (2015). Does anxiety improve during weight restoration in anorexia nervosa ? a systematic review. *Journal of eating disorders*, 3(1) :7.
- Kim, C., Daniels, M. J., Hogan, J. W., Choirat, C., and Zigler, C. M. (2019). Bayesian methods for multiple mediators : Relating principal stratification and causal mediation in the analysis of power plant emission controls. *The Annals of Applied Statistics*, 13(3) :1927–1956.
- Krishnan, K., Baglietto, L., Stone, J., Simpson, J. A., Severi, G., Evans, C. F., MacInnis, R. J., Giles, G. G., Apicella, C., and Hopper, J. L. (2017). Longitudinal Study of Mammographic Density Measures That Predict Breast Cancer Risk. *Cancer Epidemiology Biomarkers & Prevention*, 26(4) :651–660.
- Lange, T., Rasmussen, M., and Thygesen, L. C. (2014). Assessing Natural Direct and Indirect Effects Through Multiple Pathways. *American Journal of Epidemiology*, 179(4) :513–518.
- Lange, T., Vansteelandt, S., and Bekaert, M. (2012). A simple unified approach for estimating natural direct and indirect effects. *American journal of epidemiology*, 176(3) :190–195.
- Lloyd, E. C., Haase, A. M., Foster, C. E., and Verplanken, B. (2019). A systematic review of studies probing longitudinal associations between anxiety and anorexia nervosa. *Psychiatry research*, 276 :175–185.

- Loh, W. W., Moerkerke, B., Loeys, T., and Vansteelandt, S. (2020). Non-linear Mediation Analysis with High-dimensional Mediators whose Causal Structure is Unknown. *arXiv :2001.07147 [stat]*. arXiv : 2001.07147.
- Luo, C., Fa, B., Yan, Y., Wang, Y., Zhou, Y., Zhang, Y., and Yu, Z. (2020). High-dimensional mediation analysis in survival models. *PLOS Computational Biology*, 16(4) :e1007768.
- MacKinnon, D. (2008). An introduction to statistical mediation analysis. page 245.
- Marzola, E., Fassino, S., Amianto, F., and Abbate-Daga, G. (2017). Affective temperaments in anorexia nervosa : The relevance of depressive and anxious traits. *Journal of Affective Disorders*, 218 :23–29.
- Marzola, E., Knatz, S., Murray, S. B., Rockwell, R., Boutelle, K., Eisler, I., and Kaye, W. H. (2015). Short-term intensive family therapy for adolescent eating disorders : 30-month outcome. *European Eating Disorders Review*, 23(3) :210–218.
- Marzola, E., Porliod, A., Panero, M., De-Bacco, C., and Abbate-Daga, G. (2020). Affective temperaments and eating psychopathology in anorexia nervosa : Which role for anxious and depressive traits? *Journal of Affective Disorders*, 266 :374–380.
- Maskarinec, G., Dartois, L., Delalogue, S., Hopper, J., Clavel-Chapelon, F., and Baglietto, L. (2017). Tumor characteristics and family history in relation to mammographic density and breast cancer : The French E3n cohort. *Cancer Epidemiology*, 49 :156–160.
- Mattar, L., Huas, C., Duclos, J., Apfel, A., and Godart, N. (2011). Relationship between malnutrition and depression or anxiety in Anorexia Nervosa : A critical review of the literature. *Journal of Affective Disorders*, 132(3) :311–318.
- McCormack, V. A., Perry, N. M., Vinnicombe, S. J., and dos Santos Silva, I. (2010). Changes and tracking of mammographic density in relation to Pike’s model of breast tissue aging : a UK longitudinal study. *International Journal of Cancer*, pages NA–NA.
- Meinshausen, N. and Bühlmann, P. (2010). Stability selection : Stability Selection. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 72(4) :417–473.
- Nguyen, Q. C., Osypuk, T. L., Schmidt, N. M., Glymour, M. M., and Tchetgen Tchetgen, E. J. (2015). Practical guidance for conducting mediation analysis with multiple mediators using inverse odds ratio weighting. *American journal of epidemiology*, 181(5) :349–356.
- Ooi, B. N. S., Loh, H., Ho, P. J., Milne, R. L., Giles, G., Gao, C., Kraft, P., John, E. M., Swerdlow, A., Brenner, H., Wu, A. H., Haiman, C., Evans, D. G., Zheng, W., Fasching, P. A., Castelao, J. E., Kwong, A., Shen, X., Czene, K., Hall, P., Dunning, A., Easton,



- D., Hartman, M., and Li, J. (2019). The genetic interplay between body mass index, breast size and breast cancer risk : a Mendelian randomization analysis. *International Journal of Epidemiology*, 48(3) :781–794.
- Paans, N. P., Bot, M., van Strien, T., Brouwer, I. A., Visser, M., and Penninx, B. W. (2018). Eating styles in major depressive disorder : Results from a large-scale study. *Journal of psychiatric research*, 97 :38–46.
- Park, Y. (2017). "r package zqtl". <https://github.com/YPARK/zqtl>.
- Park, Y., Sarkar, A. K., He, L., Davila-Velderrain, J., De Jager, P. L., and Kellis, M. (2017). A Bayesian approach to mediation analysis predicts 206 causal target genes in Alzheimer’s disease. preprint, Genetics.
- Pearl, J. (2001). Direct and Indirect Effects. In *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, UAI’01, pages 411–420, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.
- Petersen, M. L., Sinisi, S. E., and van der Laan, M. J. (2006). Estimation of Direct Causal Effects :. *Epidemiology*, 17(3) :276–284.
- Pompili, M., Girardi, P., Tatarelli, R., Iliceto, P., De Pisa, E., Tondo, L., Akiskal, K. K., and Akiskal, H. S. (2008). Temps-a (rome) : psychometric validation of affective temperaments in clinically well subjects in mid-and south italy. *Journal of Affective Disorders*, 107(1-3) :63–75.
- Ramacciotti, C. E., Paoli, R. A., Ciapparelli, A., Marcacci, G., Placidi, G. F., Dell’ Osso, L., and Garfinkel, P. E. (2004). Affective temperament in the eating disorders. *Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity*, 9(2) :114–119.
- Rice, M. S., Bertrand, K. A., VanderWeele, T. J., Rosner, B. A., Liao, X., Adami, H.-O., and Tamimi, R. M. (2016). Mammographic density and breast cancer risk : a mediation analysis. *Breast Cancer Research*, 18(1) :94.
- Robins, J. M. and Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology (Cambridge, Mass.)*, 3(2) :143–155.
- Rubin, D. B. (1980). Randomization Analysis of Experimental Data : The Fisher Randomization Test Comment. *Journal of the American Statistical Association*, 75(371) :591.
- Shpitser, I. (2013). Counterfactual Graphical Models for Longitudinal Mediation Analysis With Unobserved Confounding. *Cognitive Science*, 37(6) :1011–1035.
- Shpitser, I. and Sherman, E. (2018). Identification of personalized effects associated with causal pathways. pages 530–539.

- Sobel, M. E. (1986). Some New Results on Indirect Effects and Their Standard Errors in Covariance Structure Models. *Sociological Methodology*, 16 :159.
- Song, Y., Zhou, X., Zhang, M., Zhao, W., Liu, Y., Kardia, S. L. R., Diez Roux, A. V., Needham, B. L., Smith, J. A., and Mukherjee, B. (2018). Bayesian Shrinkage Estimation of High Dimensional Causal Mediation Effects in Omics Studies. preprint, *Epidemiology*.
- Spielberger, C. D. (1983). State-trait anxiety inventory for adults.
- Steen, J., Loeys, T., Moerkerke, B., and Vansteelandt, S. (2017). Medflex : an r package for flexible mediation analysis using natural effect models. *Journal of Statistical Software*, 76(11).
- Sung, J., Song, Y.-M., Stone, J., Lee, K., and Kim, S.-Y. (2010). Association of Body Size Measurements and Mammographic Density in Korean Women : The Healthy Twin Study. *Cancer Epidemiology Biomarkers & Prevention*, 19(6) :1523–1531.
- Swinbourne, J., Hunt, C., Abbott, M., Russell, J., St Clare, T., and Touyz, S. (2012). The comorbidity between eating disorders and anxiety disorders : Prevalence in an eating disorder sample and anxiety disorder sample. *Australian & New Zealand Journal of Psychiatry*, 46(2) :118–131.
- Taguri, M., Featherstone, J., and Cheng, J. (2015). Causal mediation analysis with multiple causally non-ordered mediators. *Statistical Methods in Medical Research*, 27(1) :3–19.
- Tchetgen Tchetgen, E. J. (2013). Inverse odds ratio-weighted estimation for causal mediation analysis. *Statistics in medicine*, 32(26) :4567–4580.
- Thiel, A. and Paul, T. (2006). Test–retest reliability of the eating disorder inventory 2. *Journal of psychosomatic research*, 61(4) :567–569.
- Tingley, D., Yamamoto, T., Hirose, K., Keele, L., and Imai, K. (2014). mediation : R Package for Causal Mediation Analysis. *Journal of Statistical Software*, 59(5).
- Treasure, J., Stein, D., and Maguire, S. (2015). Has the time come for a staging model to map the course of eating disorders from high risk to severe enduring illness? An examination of the evidence : A staging model for eating disorders. *Early Intervention in Psychiatry*, 9(3) :173–184.
- van der Laan, M. J. (2008). Estimation based on case-control designs with known prevalence probability. *The International Journal of Biostatistics*, 4(1).
- van Kesteren (2013). "r package cmf". <https://github.com/vankesteren/cmfilter>.

- van Kesteren, E.-J. and Oberski, D. L. (2019). Exploratory Mediation Analysis with Many Potential Mediators. *Structural Equation Modeling : A Multidisciplinary Journal*, 26(5) :710–723.
- VanderWeele, T. (2015). *Explanation in causal inference : methods for mediation and interaction*. Oxford University Press.
- VanderWeele, T. and Vansteelandt, S. (2014). Mediation Analysis with Multiple Mediators. *Epidemiologic Methods*, 2(1).
- VanderWeele, T. J. and Tchetgen Tchetgen, E. J. (2017). Mediation analysis with time varying exposures and mediators. *Journal of the Royal Statistical Society : Series B (Statistical Methodology)*, 79(3) :917–938.
- VanderWeele, T. J. and Vansteelandt, S. (2009). Conceptual issues concerning mediation, interventions and composition. *Statistics and its Interface*, 2(4) :457–468.
- VanderWeele, T. J. and Vansteelandt, S. (2010). Odds Ratios for Mediation Analysis for a Dichotomous Outcome. *American Journal of Epidemiology*, 172(12) :1339–1348.
- Vansteelandt, S. and Daniel, R. M. (2017). Interventional effects for mediation analysis with multiple mediators. *Epidemiology (Cambridge, Mass.)*, 28(2) :258.
- von Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4) :395–416.
- Wade, T. D., Zhu, G., and Martin, N. G. (2010). Body Mass Index and Breast Size in Women : Same or Different Genes? *Twin Research and Human Genetics*, 13(5) :450–454.
- Walters EE, K. K. (1995). Anorexia nervosa and anorexic-like syndromes in a population-based female twin sample. *American Journal of Psychiatry*, 152(1) :64–71.
- Wasserman, L. (2010). *All of statistics : a concise course in statistical inference*. Springer texts in statistics. Springer, New York, NY, corrected second printing, 2005 edition. OCLC : 837651382.
- Zerwas, S., Lund, B. C., Von Holle, A., Thornton, L. M., Berrettini, W. H., Brandt, H., Crawford, S., Fichter, M. M., Halmi, K. A., Johnson, C., et al. (2013). Factors associated with recovery from anorexia nervosa. *Journal of psychiatric research*, 47(7) :972–979.
- Zhang, C.-H. (2010). Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2) :894–942.

- Zhang, H., Zheng, Y., Zhang, Z., Gao, T., Joyce, B., Yoon, G., Zhang, W., Schwartz, J., Just, A., Colicino, E., Vokonas, P., Zhao, L., Lv, J., Baccarelli, A., Hou, L., and Liu, L. (2016). Estimating and testing high-dimensional mediation effects in epigenetic studies. *Bioinformatics*, 32(20) :3150–3154.
- Zhao, Y. and Luo, X. (2016). Pathway Lasso : Estimate and Select Sparse Mediation Pathways with High Dimensional Mediators. *arXiv :1603.07749 [stat]*. arXiv : 1603.07749.



# Table des figures

1	Modèle de médiation simple avec un médiateur $M$ . . . . .	14
2	Relation de cause à effet entre le traitement $T$ et la variable de réponse $Y$ , quantifié par l'effet causal moyen $\tau$ . . . . .	15
3	Diagramme causal d'analyse de médiation simple avec un médiateur $M$ . . .	17
4	Diagramme causal de discrimination à l'embauche. . . . .	17
5	Extrait du rapport de l'affaire opposant Carson à Bethlehem Steel Corp. .	18
6	Les trois relations possibles entre deux médiateurs $M$ et $W$ . . . . .	22
7	Corrélation entre les médiateurs induite par $U$ . . . . .	23
8	Modèle de médiation en présence d'un grand nombre de médiateurs potentiels. Les variables $(M^1, \dots, M^L)$ ont un rôle de médiation et ce sont ces variables que l'on cherche à sélectionner. Les variables $(M^{L+1}, \dots, M^K)$ ne sont pas des médiateurs et on cherche à les exclure. . . . .	24
1.1	Diagramme causal de médiation multiple avec des médiateurs éventuellement corrélés. Le vecteur des facteurs de confusion du prétraitement $X$ n'apparaît pas dans la figure. Les lignes en pointillés représentent les possibles corrélations non causales et les traits pleins les relations causales. Une corrélation non causale est possible entre chaque paire de médiateurs mais cela n'est pas montré pour une meilleure lisibilité de la figure. Ainsi, chaque paire de médiateur peut être liés causalement. . . . .	29
1.2	Analyses de médiation multiples et simples, $U$ étant observé. Les données sont simulées selon le modèle en (a). . . . .	39
1.3	Modèle 1 (variable de réponse continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance des estimateurs à effet indirect $\hat{\delta}^1$ calculés comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ . . . . .	43

1.4	Modèle 1 (résultat continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance des estimateurs des effets indirects $\hat{\delta}^2$ calculés comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ . . . . .	44
1.5	Modèle 1 (variable de réponse continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance de l'estimateur de effet direct $\hat{\zeta}$ calculé comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ . . . . .	45
1.6	Modèle 1 (variable de réponse continue) : biais, probabilité de couverture de l'intervalle de confiance, erreur quadratique moyenne (MSE) et variance de l'estimateur de l'effet indirect joint $\hat{\delta}^Z$ calculé comme des moyennes sur les 200 simulations lorsque la corrélation entre les médiateurs varie. La formule de biais utilisée ici est $Bias = \Theta - \mathbb{E}[\hat{\Theta}]$ . . . . .	46
2.1	Médiation par bloc . . . . .	49
2.2	Comparaison du temps d'estimation des paramètres pour l'analyse de médiation multiple et pour l'analyse de médiation par blocs. Le temps est donné en minutes. Les blocs de médiateurs sont constitués de deux médiateurs corrélés de façon non-causale. . . . .	54
2.3	Modèle de médiation en présence de trois expositions, d'un grand nombre de médiateurs potentiels et d'une variable de réponse. Les effets directs des expositions $(T_1, T_2, T_3)$ sont présents dans le modèle mais non explicités sur la figure pour sa lisibilité. . . . .	58
2.4	Comparaison des boxplots des méthodes (SIS, MGLASSO) pour la précision, du rappel, du taux de faux positifs, du taux d'erreurs, du taux de faux négatifs et de la spécificité pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. Suite au rang obtenus par les méthodes, on choisit les 21 premiers médiateurs. . . . .	68
2.5	Comparaison des boxplots des méthodes (SIS, MGLASSO) pour la précision, du rappel, du taux de faux positifs, du taux d'erreurs, du taux de faux négatifs et de la spécificité pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. Suite au rang obtenus par les méthodes, on choisit les 43 premiers médiateurs. . . . .	69
2.6	Comparaison des boxplots des méthodes pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. À la première étape des méthodes HIMA, HDMA et MAHI on retient 21 médiateurs. . . . .	70
2.7	Comparaison des boxplots des méthodes pour 100 jeux de données où il y a 30 vrais médiateurs parmi 500 variables. À la première étape des méthodes HIMA, HDMA et MAHI on retient 43 médiateurs . . . . .	71

3.1 DAG du premier exemple d'utilisation du package. . . . . 81

3.2 DAG du deuxième exemple d'utilisation du package. . . . . 86

3.3 Éboulis des valeurs propres . . . . . 87

4.1 Diagrammes Quantile-Quantile de comparaison des distributions des médiateurs avec une loi gaussienne réduite, avant et après la transformation. Pour rappel DA : zone mammaire dense et NDA : zone mammaire non-dense. 94

4.2 Diagrammes causaux de l'hormonothérapie substitutive sur la densité mammographique. Les variables de confusion sont : l'âge de référence, l'année de naissance, le statut ménopausique au début de l'étude, les antécédents familiaux de CS, l'âge de la ménarche, l'utilisation de contraceptifs oraux, la parité et l'allaitement. Pour rappel THS : thérapie hormonal substitutive, PMD : pourcentage de densité mammographique. . . . . 96

4.3 Diagrammes causaux des analyses de médiation multiple. . . . . 104

4.4 Variable de réponse binaire (modèle 2) : variation des effets causaux due à la corrélation. . . . . 133

4.5 Modèle 2 (variable de réponse binaire) : biais, probabilité de couverture, variance et MSE des estimateurs de l'effet de médiation lorsque la corrélation entre les médiateurs varie. Ces résultats ont été obtenus avec 200 simulations. Chaque simulation est un jeu de données de taille 1000. . . . . 134

4.6 Variable de réponse binaire (modèle 2b) : variation des effets causaux due à la corrélation. . . . . 136

4.7 Modèle 1b (variable de réponse continue) : biais, probabilité de couverture, variance et MSE des estimateurs de l'effet de médiation lorsque la corrélation entre les médiateurs varie. Ces résultats ont été obtenus avec 200 simulations. Chaque simulation est un jeu de données de taille 1000. . . . . 137

4.8 Modèle 2b (variable de réponse binaire) : biais, probabilité de couverture, variance et MSE des estimateurs de l'effet de médiation lorsque la corrélation entre les médiateurs varie. Ces résultats ont été obtenus avec 200 simulations. Chaque simulation est en un jeu de données de taille 1000. . . . . 138

4.9 LASSO et MCP en fonction de  $\beta$  pour  $\lambda = 1$  et  $\delta = 3$  . . . . . 146





# Liste des tableaux

1	Exemple de données observables. On écrit NA lorsque la variable n'est pas observable. . . . .	15
2	Type de patient en fonction des valeurs des deux variables de réponse potentielles. . . . .	16
3	Récapitulatif des contributions de chaque chapitre et leurs valorisations. . .	26
1.1	Données contrefactuelles simulées avec deux médiateurs indépendants. . . .	38
1.2	Données observées simulées avec deux médiateurs indépendants. Les observations ont été extraites de la table 1.1. . . . .	38
1.3	Comparaison des méthodes en ajustant pour $U$ lorsque toutes les variables de la figure 1.2(a) sont observées. Pour rappel $\delta^a$ : effet médié par $a$ , $PM^a$ : proportion médiée par $a$ , $\zeta$ : effet direct et $\tau$ : effet total. . . . .	40
1.4	Comparaison des méthodes en n'ajustant pas pour $U$ : les données sont générées comme dans la figure 1.2(a) mais analysées comme si $U$ n'était pas observé. Pour rappel $\delta^a$ : effet médié par $a$ , $PM^a$ : proportion médiée par $a$ , $\zeta$ : effet direct et $\tau$ : effet total. . . . .	40
2.1	50 premières valeurs des paramètres $\alpha_k$ et $\beta_k$ , les 450 autres étant à 0. . . .	64
2.2	Exemple de table de confusion . . . . .	65
3.1	Exemple de données observées avec des médiateurs linéaires. . . . .	75
3.2	Exemple de données observables avec des médiateurs catégoriels ordonnés. . . . .	76
3.3	Exemple de données observables avec un médiateur linéaire et un médiateur catégoriel. . . . .	77
3.4	Les modèles supportés par notre package R <code>multimediate</code> . . . . .	80
3.5	Valeur des effets causaux pour le premier jeu de donnée. Pour rappel $\delta^a$ : effet médié par $a$ , $\zeta$ : effet direct et $\tau$ : effet total. . . . .	83
3.6	Valeur des effets causaux pour le deuxième jeu de données. Pour rappel $\delta^a$ est l'effet médié par $a$ , $\zeta$ est l'effet direct et $\tau$ est l'effet total. . . . .	85

4.1	Mesures mammographiques prédites chez les femmes de 60 ans par type d'hormonothérapie ménopausique. Pour rappel THS : thérapie hormonal substitutive, PMD : pourcentage de densité mammographique, NDA : zone mammaire non-dense et DA : zone mammaire dense. . . . .	97
4.2	Analyse de médiation évaluant l'effet de l'utilisation (actuelle contre jamais) de THS sur le risque de cancer du sein, globalement et par statut ER et PR en présence de PDM et IMC, médiateurs corrélés. ER : récepteur des œstrogènes; PR : récepteur de la progestérone; PDM : pourcentage de densité mammographique; IMC : indice de masse corporelle; THS : traitement hormonal de la ménopause. . . . .	98
4.3	Analyse de médiation évaluant l'effet de l'utilisation (ancienne contre jamais) de THS sur le risque de cancer du sein, globalement et par statut ER et PR en présence de PDM et IMC, médiateurs corrélés. ER : récepteur des œstrogènes; PR : récepteur de la progestérone; PDM : pourcentage de densité mammographique; IMC : indice de masse corporelle; THS : traitement hormonal de la ménopause. . . . .	99
4.4	Analyse de médiation évaluant les effets des tempéraments affectifs sur la psychopathologie alimentaire et sur l'IMC à travers l'anxiété de trait (anxiété T), l'anxiété d'état (anxiété S) et la dépression clinique. . . . .	105
4.5	Variable de réponse continu (modèle 1) : Valeurs des effets causaux obtenus avec modèle 1. . . . .	132
4.6	Variable de réponse continu (modèle 1b) : Valeurs des effets causaux obtenus avec le modèle 1b. . . . .	135
4.7	Comparaison entre les analyses de médiation simple un médiateur à la fois et notre méthode d'analyse de médiation multiple; estimations et intervalles de confiance à 95% des effets de médiation (en colonnes) pour chaque modèle de simulation (en blocs de quatre lignes). Les lignes <i>Valeur</i> contiennent les vraies valeurs, calculées à partir d'un grand nombre de contrefactuels simulés. Les lignes <i>Analyse simple <math>M^k</math></i> et <i>Analyse multiple</i> donnent respectivement les estimations de l'analyse simple faite pour le médiateur $k$ et de notre méthode prenant en compte tous les médiateurs. . . . .	141