



HAL
open science

Impacts des incitatifs économiques en médecine générale : Analyse des préférences et des motivations des médecins

Jonathan Sicsic

► **To cite this version:**

Jonathan Sicsic. Impacts des incitatifs économiques en médecine générale : Analyse des préférences et des motivations des médecins. Economies et finances. Université Paris Dauphine - Paris IX, 2014. Français. NNT : 2014PA090044 . tel-01130684

HAL Id: tel-01130684

<https://theses.hal.science/tel-01130684>

Submitted on 12 Mar 2015

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.

L'UNIVERSITE PARIS-DAUPHINE n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions doivent être considérées comme propres à leurs auteurs.

Remerciements

Je tiens à exprimer ma profonde reconnaissance et ma gratitude à Madame Carine Franc pour avoir accepté de diriger mon travail de thèse. Sa disponibilité, son enthousiasme et la confiance qu'elle m'a accordée ont été inestimables, et les nombreux conseils qu'elle m'a donnés m'ont permis d'aborder le travail de recherche avec davantage de confiance et de détermination. Ce fut un honneur de bénéficier de son expérience de recherche.

Je souhaite également remercier Monsieur Marc Le Vaillant pour ses précieux conseils et sa disponibilité. Ce fut un plaisir de bénéficier de ses connaissances et de sa pédagogie, dans ces longues heures consacrées à la construction d'un score composite.

J'adresse mes sincères remerciements à Madame Nathalie Pelletier-Fleury pour le soutien qu'elle m'a apporté durant les différentes étapes de cette thèse, et ses conseils quant aux prolongements à donner à mon travail.

Je remercie vivement Madame Katharina Janus et Madame Nora Moumjid de me faire l'honneur d'être rapporteur de ma thèse. J'espère que nous pourrons continuer à échanger autour des questions de préférences et de motivations au travers de collaborations futures.

Je suis également très reconnaissant envers Madame Brigitte Dormont, Monsieur Laurent Letriliart et Madame Christine Peyron pour l'intérêt qu'ils témoignent à mon travail en acceptant d'être membre de mon jury de thèse.

J'exprime toute ma reconnaissance aux membres du groupe PROSPERE, Messieurs Philippe Boismault, Yann Bourgueil, Thomas Cartier, Pascal Clerc, Didier Duhot, Julien Mousqués, Michel Naiditch, Olivier Saint-Lary, Philippe Szidon et Madame Cécile Fournier pour avoir partagé leur réflexion sur l'organisation des soins de premiers recours et pour m'avoir encouragé dans mes recherches.

Ce travail n'aurait pas été aussi enrichissant si je n'avais pas fait la rencontre et partagé le quotidien de Clémence Bussière, Magali Dumontet, Nicolas Krucien, Clémence Perraudin. Je les remercie pour m'avoir apporté leur aide, leurs conseils avisés et leur amitié, ainsi que pour la bonne ambiance qu'ils ont su faire régner au sein du Cermes3. J'ai aussi une

pensée pour tous ceux qui ont fait une apparition plus ou moins brève au Bâtiment E, en particulier Aurélie, Julia, Laurie, Guillaume, Frédérique, et Diane.

Je tiens par ailleurs à remercier tous mes camarades doctorants ou jeunes docteurs pour les agréables moments partagés en leur compagnie et les discussions enrichissantes que nous avons eu : Louis Arnault, Rosalind Bell-Aldeghi, Eleonore Bayen, Damien Bricard, Laureen Dartois, Roméo Fontaine, Lise Gerbaud, Léontine Goldzahl, Mathilde Godard, Robin Hege, Gaëlle Jamet, Sandrine Juin, Elie Lacroix, Béranger Lueza, Mathilde Peron, Romain Sibille, Lexane Weber, Nina Zerrar.

J'adresse aussi mes remerciements à toutes les personnes qui, au gré des séminaires de travail, colloques ou plus simplement au détour d'un couloir, m'ont apporté leurs conseils et leurs encouragements. Je pense en particulier à Thomas Barnay, Julia Bonastre, Eve Caroli, Paul Dourgnon, Veronique Drouglazet, Hélène Huber, Tor Iversen, Florence Jusot, Marc Perronin, Aurélie Pierre, Lise Rochaix, Anne-Laure Samson, Yann Videau et Jérôme Wittwer.

Cette thèse n'aurait pas eu lieu sans le soutien inconditionnel de mes parents. Je les remercie du fond du cœur pour m'avoir apporté leur affection et les nombreux principes qui m'ont permis de mener ce travail jusqu'au bout. Je remercie également Jérémie, Michaël, Marlène, Sarah, Marie-Christine, David, Juliette, Line, Guy, Nicole, Henri, Max, Alex, Ben, Alain, Cathy, et toute ma famille pour m'avoir encouragé dans cette voie. J'exprime également mon affection à mes ami(e)s : Alexis, Charlie, Pauline, Magalie, Guillaume, Charlotte, Julien L, Claire, Cécile, Laura, Pierre-Alain, Nico, Lisa, Anne-Lise, Marion, Elsa, Julien A, Irimi, Hugo, Flavien, Catherine, Marie-Karelle, Anthony, Audrey, Riyad, Pierre, Alexandre, Yann, Gab, Régis, Aurélie, Antho, Marina, pour avoir été à mes côtés. J'en oublie certainement. Je leur adresse à tous ma plus profonde sympathie.

Enfin, je tiens à remercier certains enseignants qui, d'une manière ou une autre, ont été source d'inspiration dans les différentes étapes de cette thèse : C. Teisseidre, E. Buisson, A. Beitone, C. Leroy, F. Beninel, B. Gelein, A. Trognon.

Valorisations scientifiques

Articles acceptés à publication dans le cadre de la thèse

Saint-Lary O, Sicsic J. Impact of a pay for performance programme on French GPs' consultation length. Accepté dans la revue **Health Policy** le 01/09/2014 (révisions mineures).

Sicsic J, Franc C. Obstacles to the uptake of breast, cervical, and colorectal cancer screenings. What remains to be achieved by French national programmes ? Accepté dans la revue **BMC Health Services Research** (*in press*).

Sicsic J, Le Vaillant M, Franc C. Building a composite score of general practitioners' intrinsic motivation : a comparison of methods. **Int J Qual Health Care**, 2014 Apr ; 26(2) :167-73.

Sicsic J, Le Vaillant M, Franc C. Intrinsic and extrinsic motivations in primary care : an explanatory study among French general practitioners. **Health Policy**, 2012 Dec ; 108(2-3) : 140-8.

Autres publications

Saint-Lary O, Bernard E, Sicsic J, Plu I, François-Pursell I, Franc C. Why did most French GPs choose not to join the voluntary national pay-for-performance program ? **Plos One**, 2013 Sep 9,8(9) :e72684.

Table des matières

Remerciements	i
Valorisations scientifiques	iii
Liste des Tableaux.....	x
Liste des Figures.....	xii
Abréviations	xiii
Introduction générale.....	1
Chapitre 1.....	13
Impact des incitatifs économiques sur différents indicateurs de la qualité des soins en médecine générale	13
1. Impact du CAPI sur la durée de consultation des médecins généralistes français	14
1.1. Introduction	14
1.2. La base de données Ecogen	16
1.3. Méthodologie statistique.....	17
1.3.1. Le modèle multi-niveaux de la durée de consultation.....	17
1.3.2. Les hypothèses du modèle.....	18
1.3.3. Les différents modèles multi-niveaux estimés	18
1.4. Statistiques descriptives et modèles univariés.....	19
1.4.1. Distribution de la durée de consultation.....	19
1.4.2. Déterminants de la durée de consultation.....	20
1.5. Résultats des modèles multi-niveaux	23
1.5.1. Modèles tout motif de consultation confondu	23
1.5.2. Modèles par motif de consultation	26
1.6. Discussion.....	29
2. Impact des incitatifs économiques sur le dépistage des cancers.....	32
2.1. Introduction	32

2.2.	Analyse des freins au dépistage des cancers	34
2.2.1.	Contexte	34
2.2.2.	L'enquête ESPS.....	36
2.2.3.	Variables de recours au dépistage des cancers	36
2.2.4.	Déterminants du non recours au dépistage des cancers	37
2.2.5.	Méthodologie d'analyse statistique	38
2.2.6.	Evolution du recours au dépistage des cancers et de leurs déterminants	39
2.2.7.	Analyses multivariées du non recours au dépistage des cancers.....	43
2.2.8.	Discussion	49
2.3.	Impact du CAPI sur le recours au dépistage du cancer du sein.....	53
2.3.1.	Objectif et hypothèse à tester	53
2.3.2.	L'échantillon généraliste des bénéficiaires (EGB).....	55
2.3.3.	Accès à l'EGB et constructions des bases de données	55
2.3.4.	Variables dépendantes et indépendantes	57
2.3.5.	Analyse économétrique	58
2.3.6.	Analyse descriptive du recours au dépistage du cancer du sein.....	60
2.3.7.	Résultats des modèles économétriques	64
2.3.8.	Discussion	66
	Annexes.....	70
	Annexe A. Tables et graphiques additonnels	70
	Chapitre 2.....	74
	Préférences des médecins généralistes pour des dispositifs d'amélioration de leur pratique : application au dépistage des cancers	74
1.	Introduction.....	75
1.1.	Problématique et objectifs	75
1.2.	Revue de la littérature.....	77
1.2.1.	Expériences étrangères	77

1.2.2.	Contexte national.....	78
2.	Construction du questionnaire	79
2.1.	Définition des dispositifs et de leurs niveaux	80
2.1.1.	Entretiens avec les experts	80
2.1.2.	Présentation des dispositifs retenus.....	80
2.1.3.	Faisabilité des dispositifs	82
2.2.	Méthodes de révélation des préférences	84
2.2.1.	Synthèses des méthodes de révélation des préférences.....	84
2.2.2.	Choix d'une méthode	86
2.3.	Design expérimental	87
2.4.	Population d'étude et collecte des données	88
2.5.	Description des données de l'enquête	90
3.	Analyse des données de choix discrets (DCE).....	93
3.1.	Méthodes	93
3.1.1.	Des données au modèle économétrique sous-jacent	93
3.1.2.	Le modèle logit conditionnel « simple » à effets fixes (M1)	94
3.1.3.	Le modèle logit conditionnel à effets fixes avec interactions (M2).....	96
3.1.4.	Estimation des taux marginaux de substitution (TMS).....	97
3.1.5.	Le modèle logit conditionnel à effets aléatoires (M3)	97
3.1.6.	Non réponse et non arbitrage	101
3.2.	Résultats des modèles DCE.....	103
3.2.1.	Statistiques descriptives	103
3.2.2.	Modèles à effets fixes (M1 et M2).....	105
3.2.3.	Estimation des taux marginaux de substitution (TMS).....	108
3.2.4.	Modèles à effets aléatoires (M3).....	109
3.3.	Discussion.....	113
4.	Analyse des données de best-worst scaling (BWS).....	117

4.1. Méthodes	117
4.1.1. Analyse descriptive	117
4.1.2. Analyse économétrique : approche par les paires versus par les attributs	118
4.1.3. Approche par les paires	119
4.1.4. Approche par les attributs	122
4.1.5. Analyse de l'hétérogénéité des préférences	124
4.2. Résultats des modèles BWS	125
4.2.1. Analyses descriptives	125
4.2.2. Modélisations économétriques	126
4.3. Discussion.....	131
5. Conclusion	134
Annexes.....	136
Annexe B- Tableaux et figures additionnels	136
Chapitre 3.....	142
Motivations intrinsèques et extrinsèques en médecine générale: complémentarité ou substituabilité ?.....	142
1. Introduction.....	143
1.1. Problématique et objectifs	143
1.2. Revue de la littérature.....	145
1.2.1. Définitions du concept de motivations intrinsèques	145
1.2.2. Les théories du crowding-out effect.....	147
2. Données utilisées pour la mesure des motivations.....	152
2.1. Description de l'enquête	152
2.2. Sélection des variables de motivations intrinsèques.....	153
3. Construction d'un indicateur composite de la motivation intrinsèque des médecins généralistes	156
3.1. Introduction	156
3.2. Validation des variables de motivations intrinsèques.....	157

3.2.1.	L'analyse des correspondances multiples	157
3.2.2.	Sélection des indicateurs de motivations intrinsèques	158
3.3.	Méthodes pour la construction des ICMI	159
3.3.1.	L'analyse des correspondances multiples	160
3.3.2.	L'analyse factorielle confirmatoire pour indicateurs catégoriels.....	161
3.3.3.	Le modèle logistique à deux paramètres (modèle de Birnbaum).....	162
3.3.4.	Méthodes pour la comparaison des trois indicateurs composites.....	163
3.4.	Estimation et comparaison des trois ICMI	165
3.4.1.	Pondération des indicateurs.....	165
3.4.2.	Distribution des ICMI	166
3.5.	Discussion.....	168
4.	Relation entre motivation intrinsèque et extrinsèque en médecine générale. Une étude exploratoire.....	171
4.1.	Objectifs.....	171
4.2.	Données et méthodes	171
4.2.1.	Caractéristiques des médecins.....	171
4.2.2.	Variables de motivations extrinsèques	172
4.2.1.	Construction et analyse de la variabilité du score de MI	174
4.3.	Résultats.....	175
4.3.1.	Distribution de l'indicateur composite de MI	175
4.3.2.	Les déterminants de la variabilité du score	176
4.3.3.	Interaction avec les variables de ME.....	177
4.4.	Discussion.....	178
Annexes.....		181
Annexe C.1.	Questionnaire.....	181
Annexe C.2.	Graphiques additionnels	187
Conclusion générale		189

Bibliographie..... 196

Liste des Tableaux

Tableau 1. Statistiques descriptives et modèles univariés.....	22
Tableau 2. Modèles multivariés de la durée de consultation (log).....	25
Tableau 3. Modèles multivariés de la durée de consultation (log) selon le motif de consultation	27
Tableau 4. Distribution des variables indépendantes en fonction de l'année d'enquête.....	42
Tableau 5. Modèles logistiques multivariés du non recours au dépistage du cancer du col de l'utérus.....	45
Tableau 6. Modèles logistiques multivariés du non recours au dépistage cancer du sein	46
Tableau 7. Modèles logistiques multivariés du non recours au dépistage du cancer colorectal	48
Tableau 8. Statistiques descriptives sur les échantillons pre- et post- CAPI	61
Tableau 9. Date de signature du CAPI (N=16 428 individus)	62
Tableau 10. Relation entre les covariables et la réalisation d'une mammographie	63
Tableau 11. Résultats du modèle conditionnel de recours à la mammographie	65
Tableau 12. Dispositifs retenus (exemple dans le cadre du dépistage du cancer du sein)	82
Tableau 13. Description du design à 12 scénarios	88
Tableau 14. Statistiques descriptives sur les trois versions du questionnaire	91
Tableau 15. Variables médecins introduites dans les modèles mixtes et hypothèses	101
Tableau 16. Réponses aux épreuves DCE selon le type de cancer	104
Tableau 17. Résultats des modèles logit conditionnels à effets fixes pour le dépistage du cancer du sein (N=110 MGs)	106
Tableau 18. Résultats des modèles logit conditionnels à effets fixes pour le dépistage du cancer du col de l'utérus (N=113 MGs).....	107
Tableau 19. Résultats des modèles logit conditionnels à effets fixes pour le dépistage du cancer colorectal (N= 115 MGs).....	107
Tableau 20. Estimation des $TMSrem \rightarrow x$ en fonction du niveau de rémunération additionnelle et du type de dépistage	108
Tableau 21. Résultats des modèles mixtes pour le dépistage du cancer du sein.....	110
Tableau 22. Résultats des modèles mixtes pour le dépistage du cancer du col de l'utérus ...	111
Tableau 23. Résultats des modèles mixtes pour le dépistage du cancer colorectal.....	112
Tableau 24. Représentation simplifiée de la base de données dans le modèle des paires.....	121

Tableau 25. Représentation simplifiée de la base de données dans le modèle marginal.	124
Tableau 26. Fréquence de sélection des niveaux d'attributs comme préférés/moins préférés	125
Tableau 27. Score relatif, score standardisé et rang des niveaux d'attribut	126
Tableau 28. Comparaison des résultats des modèles des paires et marginaux.....	127
Tableau 29. Résultats des modèles logit conditionnels multinomiaux (modèles des paires).	128
Tableau 30. Résultats des modèles logit conditionnels multinomiaux avec interactions (modèles des paires).....	130
Tableau 31. Description des variables de motivations intrinsèques.....	156
Tableau 32. Résultats de l'Analyse des Correspondances Multiples.....	159
Tableau 33. Synthèse des indicateurs de MI et estimation des poids basés sur l'ACM	165
Tableau 34. Résultats des estimations des modèles CFA et Birnbaum	166
Tableau 35. Statistiques descriptives sur la distribution des ICMI	167
Tableau 36. Coefficients de Kappa pondérés entre les trois ICMI	168
Tableau 37. Statistiques descriptives (MI, ME et variables médecin)	173
Tableau 38. Résultats des modèles multiniveaux : impact des caractéristiques médecins sur la variabilité de l'ICMI.....	176
Tableau 39. Coefficients de corrélation de Spearman entre l'ICMI et les variables de ME et tests des rangs de Wilcoxon	177

Liste des Figures

Figure 1. Distribution de la durée de consultation (N= 20 665 consultations).	20
Figure 2. Evolution du recours au dépistage du cancer du col de l'utérus entre 2006 et 2010	39
Figure 3. Evolution du recours au dépistage du cancer du col de l'utérus entre 2006 et 2010	40
Figure 4. Evolution du recours au dépistage du cancer colorectal entre 2006 et 2010	40
Figure 5. Modélisation du recours à la mammographie en fonction de la visite chez le médecin traitant.	54
Figure 6. Histogramme du rang de classement de 30 MGs en fonction de l'ICMI calculé ...	167

Abréviations

ACM : Analyse des Correspondances Multiples

AFSSAPS : Agence Française de Sécurité Sanitaire des Produits de Santé

ALD : Affection de Longue Durée

AM : Assurance Maladie

AME : Aide Médicale d'Etat

ANSM : Agence Nationale de Sécurité du Médicament

ARS : Agence Régionale de Santé

BWAS : Best-Worst Attribute Scaling

CAPI : Contrat d'Amélioration des Pratiques Individuelles

CCAM : Classification Communes des Actes Médicaux

CET : Cognitive Evaluation Theory

CFA : Confirmatory Factor Analysis

CMU : Couverture Maladie Universelle

CMU-C : Couverture Maladie Universelle Complémentaire

CNAMTS : Caisse Nationale d'Assurance Maladie des Travailleurs Salariés

CNIL : Commission Nationale de l'Informatique et des Libertés

DCE : Discrete Choice Experiment

DO : Dépistage Organisé

ECOGEN : Eléments de la COnsultation en médecine GENérale

EGB : Echantillon Généraliste des Bénéficiaires

ESPS : Enquête Santé et Protection Sociale

HAS : Haute Autorité de Santé

IC : Indicateur Composite

INCa : Institut National contre le Cancer

INSEE : Institut National de la Statistique et des Etudes Economiques

INVS : Institut National de Veille Sanitaire

IRDES : Institut de Recherche et de Documentation en Economie de la Santé

IRT : Item Response Theory

MCO : Motivation Crowding Out

ME : Motivations Extrinsèques

MG : Médecin Généraliste

MGT : Médecin Généraliste Traitant

MI : Motivations Intrinsèques

MSA : Mutualité Sociale Agricole

OCDE : Organisation de Coopération et de Développement Economiques

OMS : Organisation Mondiale de la Santé

P4P : Pay-for-Performance

PROSPERE : Partenariat pluridisciplinaire de Recherche sur l'Organisation des Soins de PrEmiers REcours

RIAP : Relevé Individuel d'Activité Professionnelle

ROSP : Rémunération sur Objectifs de Santé Publique

RG : Régime Général

RSI : Régime Social des Indépendants

SDT : Self Determination Theory

SNIR-AM : Système National d'Information Inter-Régimes de l'Assurance Maladie

TMS : Taux Marginal de Substitution

UNCAM : Union Nationale des Caisses d'Assurance Maladie

URML : Union Régionale des Médecins Libéraux

Introduction générale

La médecine générale a fait l'objet de nombreuses réformes en France ces 10 dernières années. En particulier, la réforme de l'assurance maladie d'août 2004 a reconnu un rôle clé au médecin généraliste (MG) dans la prise en charge des patients en instaurant le dispositif du médecin traitant. Toutefois, à l'instar de nombreux pays de l'OCDE (OCDE 2013), la médecine générale, en France, est en crise. Malgré son récent statut de spécialité depuis 2004, elle souffre d'un manque d'attractivité pour les jeunes : ainsi, en 2010, 33% des postes proposés en médecine générale à l'issue des épreuves nationales classantes sont restés vacants (Fauvet et al 2012). Par ailleurs, les conditions d'exercice des médecins généralistes ont beaucoup évolué du fait notamment du vieillissement de la population, de la prévalence croissante des maladies chroniques et des progrès techniques. Ces modifications ont transformé les modèles de pratique et de prise en charge (ONDPS 2007). Dans ce contexte, les différentes réformes tentent d'adapter les conditions de la pratique médicale et les modes d'organisations des soins de premiers recours.

Le mode de rémunération des médecins et particulièrement celui des MGs, installés pour une large majorité d'entre eux en libéral, s'est diversifié ces dernières années : au seul paiement à l'acte se sont progressivement ajoutés différents forfaits (suivi ALD, coordinations, etc.) et une rémunération additionnelle sur objectifs de santé publique. La possibilité d'introduire des contrats individualisés - dits contrats de bonne pratique entre l'Assurance Maladie (AM) et les médecins libéraux - est apparue pour la première fois dans la loi de financement de la sécurité sociale de 2000. Ces contrats ouvraient droit à des majorations forfaitaires en contrepartie de la réalisation de pratiques ciblées, telles que la permanence et la coordination des soins, et étaient aussi destinés à soutenir l'activité médicale dans certaines zones géographiques (Bras et Duhamel 2008). Puis, l'article 43 (L 162-12-21) de la loi de financement pour 2008 offre la possibilité pour les organismes locaux de l'AM de passer des contrats individuels avec des médecins libéraux prévoyant des « contreparties financières », liées à l'atteinte d'objectifs concernant « *la prescription, la participation à des actions de dépistage et de prévention (...) ainsi que toute action d'amélioration des pratiques, de la formation et de l'information des professionnels* ». Cette loi conduit à l'introduction des Contrats d'Amélioration des Pratiques Individuelles (**CAPi**) en Juillet 2009. Ce dispositif

perdure pendant une période de 2 ans et demie sur la base du volontariat, permettant à tout médecin généraliste libéral éligible de signer un contrat de gré à gré avec une caisse primaire d'assurance maladie. Le CAPI prévoit le versement de bonus financiers en contrepartie de l'amélioration d'indicateurs définis par le contrat. Plus précisément, 16 objectifs cibles communs à tous les médecins sont définis sur la base de recommandations émises par la Haute Autorité de la Santé, l'Agence Nationale de la Sécurité du Médicament (ex AFSSAPS) et de comparaisons internationales (Commission des comptes 2011). A la date anniversaire du contrat, le médecin perçoit une rémunération associée à l'atteinte ou à l'amélioration de ces objectifs cibles, cette prime pouvant atteindre un montant maximal de 5000€ pour les médecins les plus « performants ». Par ailleurs, aucune conséquence négative ou sanction n'est attachée au non-respect des objectifs, et chaque indicateur est jugé indépendamment des autres.

Dans un contexte de très forte opposition des principaux syndicats de médecins généralistes, le succès de ces contrats est assez inattendu : entre juillet 2009 et décembre 2011, plus du tiers des MGs éligibles ont volontairement signé un CAPI (38%). Le dispositif du CAPI est généralisé à tous les médecins généralistes par la convention médicale de juillet 2011 alors signée par tous les syndicats de médecins. Cette convention instaure une Rémunération sur Objectif de Santé Publique (**ROSP**) dès janvier 2012 pour tous les MGs (sauf refus explicite, 97% y ont adhéré) et élargit le nombre d'indicateurs définis dans le contrat. Ainsi, elle offre un complément de revenus d'un montant maximal de 9 100 euros par an pour améliorer la qualité des soins et de la prise en charge, approchée par une trentaine d'indicateurs : des indicateurs cliniques (prise en charge de pathologies chroniques), des indicateurs d'actions de prévention (dépistage des cancers, vaccination), des indicateurs sur l'organisation du cabinet (tenue du dossier médical, utilisation de logiciels certifiés d'aide à la prescription) et enfin, des indicateurs d'efficience de la pratique (atteinte d'un taux prédéfini de prescriptions de génériques). Ainsi en 2013, la ROSP représentait en moyenne 12% des revenus des médecins généralistes français et son coût brut s'est élevé à 341 millions d'euros (Cour des comptes 2014).

Le CAPI et la ROSP appartiennent à « la famille » des paiements à la performance (**P4P**). Appliqués au contexte de la santé, ces paiements visent le plus souvent à rémunérer davantage les professionnels de santé en fonction de leurs résultats en termes de réalisation d'objectifs prédéfinis par le contrat. Recourir à ces dispositifs est cohérent avec les résultats de la théorie des contrats : un agent économique rationnel est sensible aux incitatifs

économiques, de sorte qu'associer une rémunération à la performance est un moyen d'augmenter sa productivité notamment en cas d'asymétrie d'information (Bolton and Dewatripont 1957; Laffont and Martimort 2002). En effet, la difficulté voire l'impossibilité de mesurer les résultats de l'activité médicale et d'observer le niveau d'effort consenti par le médecin, même *ex post* (Arrow 1963) rendent impossible la construction d'un système de rémunération basé sur la productivité des médecins ; la mise en place d'un système d'incitations économiques est un moyen de pallier les problèmes d'asymétrie d'information comme le problème d'aléa moral rencontré dans la double relation d'agence médecin-patient et médecin-financeur (Ellis and McGuire 1986; Rochaix 1989; Robinson 2001; Choné and Ma 2011). Les incitatifs instaurés permettent en général d'obtenir une situation de second rang ou « second best » où une rente informationnelle est cédée par le principal (le financeur des soins / le patient) à la partie informée (l'agent, le médecin). Cette rente informationnelle se traduit par l'instauration de mécanismes monétaires (sanction / bonus).

Toutefois, le design d'un incitatif économique est complexe, car il va dépendre de l'objectif fixé et de la nature de l'asymétrie d'informations à « contrôler », mais aussi du contexte et dans notre cas, des schémas de paiement propres aux différents systèmes de santé. En effet, les différents pays qui mobilisent des systèmes de P4P n'ont pas un système uniforme d'organisation de la médecine ambulatoire et par conséquent, n'ont pas un même système initial de paiement des médecins : par exemple, au Royaume-Uni les médecins généralistes sont payés à la capitation alors qu'en France, les médecins libéraux sont payés à l'acte. Or, ces systèmes initiaux produisent des incitations différentes concernant par exemple le partage des tâches entre soins curatifs et préventifs (Rochaix 2004; Franc and Lesur 2004; De Jaegher 2010). Ainsi, en comparant les incitations produites par différents schémas de rémunération en vigueur (paiement à l'acte, capitation, salariat), Franc et Lesur (2004) ont montré que le paiement à la capitation était celui qui fournissait le plus d'incitations à la prévention et le paiement à l'acte, le moins. En outre, plusieurs travaux ont montré que le paiement à l'acte pouvait engendrer une sur-production de soins curatifs (Gosden et al 2001; Hennig-Schmidt et al 2011), phénomène souvent associé au mécanisme de la demande induite (Delattre and Dormont 2003). De fait, la juxtaposition d'incitations financières en sus de systèmes de paiements aux propriétés incitatives différentes est susceptible de produire des résultats contrastés, rendant difficile toute transposition des résultats d'évaluations d'un pays à un autre.

Par ailleurs, la notion de paiement à la performance nécessite de définir ce qu'est un système de santé performant : s'agit-il d'un système qui améliore les performances cliniques et la qualité des soins, d'un système qui améliore la satisfaction des patients, ou encore d'un système qui améliore l'efficacité ? Après avoir parfaitement identifié l'objectif ciblé par le P4P, car tous les objectifs ne sont pas toujours compatibles, il s'agit ensuite de définir les modalités de distribution des éventuels bonus alloués par le dispositif. Selon Bras et Duhamel (2008), cette distribution peut prendre trois formes : 1) le « tournoi », dans lequel seuls les individus les plus performants sont récompensés, 2) le « progrès », i.e. les rémunérations sont attribuées indépendamment du niveau des résultats obtenus mais en considérant l'évolution des indicateurs ciblés, 3) « la cible », i.e. les bonus sont attribués à tous ceux qui dépassent un niveau de performance. Le choix d'un type d'allocation a des conséquences importantes sur les incitations produites ainsi que les coûts engendrés. Ainsi, la 3^e modalité est plus coûteuse que les deux autres, et peut entraîner un effet d'aubaine pour les médecins initialement proches de la cible. En France, le calcul des bonus du CAPI ou de la ROSP est basé sur un mélange des modalités 2) et 3), puisque les progrès sont récompensés en exigeant un niveau de progrès d'autant plus faible que le niveau d'atteinte des objectifs au départ est élevé (Bras et Duhamel 2008).

Dans le contexte de la rémunération des médecins, les premiers P4P ont été instaurés dès le début des années 2000 dans les pays anglo-saxons, notamment au travers du « California Pay for Performance Program » aux Etats-Unis (2001) et du « Quality and Outcomes Framework » (QOF) en Angleterre (2004). Ces systèmes se sont rapidement étendus à d'autres pays de l'OCDE. Malgré leur généralisation en cours dans un nombre croissant de pays, des questions subsistent quant à la capacité de ces incitatifs, dits extrinsèques, à générer une plus grande efficacité des organisations et plus particulièrement des organisations de soins de premiers recours. En effet, les résultats des études sur l'impact des paiements à la performance sur la qualité des soins sont controversés (Petersen et al 2006; Mullen et al 2010; Houle et al 2012; Eijkenaar et al 2013) parce que souvent partiels et/ou mettant en évidence, dans le domaine de la prévention par exemple, des effets faibles voire inexistantes (Grady et al 1997a; Hillman et al 1998; Town et al 2005a; Rosenthal and Frank 2006; Conrad and Perry 2009; Kiran et al 2014). Les résultats d'évaluations du QOF anglais par exemple, sont mitigés, montrant un retour sur investissement faible comparé aux sommes engagées, de l'ordre de 3.5 Milliards d'euros en 2005 (Bras et Duhamel 2008). Il faut toutefois noter que ce dispositif est peu comparable à la ROSP instaurée en France, car le

QOF comporte un nombre beaucoup plus élevé d'indicateurs (146 indicateurs *versus* 29 pour la ROSP), et implique des montants financiers bien plus élevés (représentant près du tiers du revenu des médecins généralistes anglais).

Par ailleurs, Selon Mannion et Davies (2008), malgré leur diversité, les P4P appliqués à un domaine tel que la santé ont des caractéristiques communes et génèrent potentiellement des effets indésirables importants. A partir d'expérimentations de schémas incitatifs, il a notamment été montré au Royaume-Uni et en Espagne que l'augmentation de la charge de travail et/ou de stress associée à ces politiques incitatives ont eu un impact négatif sur la qualité de vie et la satisfaction des médecins, diminuant en retour la qualité des soins prodigués (Hadley et al 1999; Gené-Badia et al 2007). De même, il est à craindre que des politiques incitatives « agressives » puissent favoriser un sentiment de « burnout » (épuisement professionnel) déjà assez répandu chez les médecins généralistes (Smith 2001; Cathébras et al 2008). Certains auteurs pointent d'autres effets indésirables potentiels des P4P, comme la sélection de patients et/ou de pathologies plus « rémunérateurs » pour le médecin, même si de tels phénomènes n'ont pas été clairement mis en évidence par la littérature empirique (Gosden et al 2001; Doran et al 2008; Mckalip 2009). Enfin, Mannion et Davies (2008) comme d'autres auteurs (Forsberg et al 2001; Forsberg et al 2002; Mangin and Toop 2007; Maynard 2012) listent un certain nombre d'effets indésirables à plus long terme comme une baisse du sentiment d'autonomie du médecin, une éviction des motivations intrinsèques ou encore une dégradation de la relation médecin-patient.

Les résultats mitigés du P4P pourraient être la conséquence de l'existence d'autres types de motivations, appelées motivations intrinsèques (**MI**), liées à l'activité professionnelle en elle-même (Deci et Ryan, 2000), et/ou au respect d'une norme ou d'une éthique professionnelle (Frey 1997; Janus 2010). Il a été montré que les MI ont des effets bénéfiques sur l'engagement au travail, l'apprentissage et le partage des connaissances (Osterloh and Frey 2000). Des travaux menés en psychologie puis en économie ont émis l'hypothèse que pour certaines activités, les motivations extrinsèques (**ME**), associées le plus souvent aux récompenses monétaires¹, évinçaient les MI, conduisant à l'effet d'éviction communément

¹ Il existe un débat entre certains psychologues et économistes concernant le contenu exact des motivations extrinsèques. Notamment, les premiers considèrent le plus souvent que la recherche de l'estime d'autrui ou d'une reconnaissance des pairs peut être synonyme de motivations extrinsèques (Deci 1971). Pour les économistes, les motivations extrinsèques sont le plus souvent associées à des récompenses monétaires, et la recherche de l'approbation d'autrui est généralement considéré comme une préférence sociale qui n'est parfois pas distinguée des motivations intrinsèques (Bénabou and Tirole 2005; Sliwka 2006).

appelé de « crowding-out » (Frey 1994; Frey and Jegen 2001). Des travaux issus de la théorie cognitive en psychologie montrent que les incitatifs économiques peuvent diminuer le sentiment d'autonomie et de compétence des agents, qui sont les déterminants de base des MI (Deci et al 1999b). Des travaux sociologiques montrent que les incitations monétaires peuvent nuire à la recherche de l'approbation sociale (Gneezy and Rustichini 2000a) et à l'adoption d'un comportement civique (Frey and Oberholzer-Gee 1997).

D'un point de vue économique, les ME, incitations économiques, nous l'avons vu, permettent de remédier aux problèmes d'asymétrie d'information qui compte tenu des comportements stratégiques des acteurs en présence d'aléa moral sont susceptibles de ne pas produire le niveau d'effort optimal (Laffont and Martimort 2002). Dans le modèle principal - agent, Benabou et Tirole (Bénabou and Tirole 2003) reprennent les résultats de Gneezy et Rustichini (2000) afin d'étudier la nature de l'interaction entre MI et ME : les auteurs montrent notamment que l'introduction d'incitations monétaires décourage les efforts de l'agent, car celui-ci interprète ces incitatifs comme le signal d'une perte de confiance de la part du principal ou le signal que la tâche est moins attractive ou plus difficile à réaliser.

Dans le domaine particulier de la santé, plusieurs auteurs montrent en quoi l'exercice même de la médecine avec ses spécificités (par exemple, le respect d'une norme ou d'une éthique médicale) peut être source de MI (Videau 2010; Janus 2010; Feng et al 2011). La prise en compte de ces MI pourrait sérieusement remettre en cause l'efficacité théorique attendue des mécanismes d'incitations monétaires de type P4P dans le domaine de la production de soins. Ainsi, Feng et al (2011) étudient l'impact des incitatifs financiers générés par le QOF sur les MI des médecins généralistes anglais. Les auteurs définissent les MI à partir de la proportion de patients traités sans paiement QOF additionnel. Ils mettent en évidence un effet d'éviction des MI, i.e. une augmentation du revenu d'une livre sterling de P4P réduisait l'indicateur calculé de MI de 0.12%. En France, Videau (2010) conclut que l'effet de « crowding out », correspondant à l'érosion des MI du fait de l'instauration de ME, chez les médecins généralistes libéraux est particulièrement plausible pour trois types de raisons : d'une part, la pratique libérale de la médecine sous tend un profond désir d'autonomie au travail ainsi qu'une aspiration forte à l'individualisation des pratiques; d'autre part, la médecine constitue un ensemble de tâches à haut intérêt et haute responsabilité (diagnostic, gestion de l'urgence, décision thérapeutique dans des situations complexes de polyopathologies, etc.). Enfin, la médecine est un domaine largement soumis à l'existence de

normes professionnelles et où l'on considère souvent que le médecin est mû par une éthique ou un altruisme médical (Videau et al 2010).

Face à la complexité des motivations, certains auteurs ont élargi la définition de la fonction d'utilité du médecin à la prise en compte des préférences à la fois monétaires et non monétaires de la pratique. Ainsi, Scott (Scott 2001) a estimé les déterminants de la fourniture de services d'un médecin en formalisant une fonction d'utilité par la méthodologie des choix discrets : il a montré que les médecins avaient des préférences marquées pour les dispositifs non monétaires (diminution du temps de travail, utilisation de guides de bonne pratique) et des dispositions à payer élevées pour ces dispositifs. En France, Ammi et Peyron (2010) se sont intéressés aux préférences des médecins pour les dispositifs susceptibles d'être instaurés afin de favoriser la prévention en médecine de ville. Leurs résultats ont confirmé que les médecins étaient sensibles aux dispositifs non monétaires destinés à favoriser leur activité de prévention (guides de bonne pratique, formation à la prévention, retour d'information sur la pratique) et avaient une préférence pour l'activité en groupe. Par ailleurs, dans un article cherchant à comprendre les sources de mécontentement des médecins généralistes allemands ayant conduit à la grève de 2006, Janus et al (2007) ont mis en évidence l'importance relative des déterminants non monétaires (reconnaissance, formation continue, coopération entre professionnels de santé) *vs* les déterminants monétaires, comme facteurs explicatifs de la satisfaction au travail.

Analyser la perception des P4P par les médecins et leurs préférences pour des dispositifs éventuellement complémentaires semble d'autant plus pertinent dans le contexte français qu'entre 2009 et le milieu de 2011, les médecins généralistes ont eu le choix d'adhérer au CAPI et seuls un tiers ont signé ce contrat qui ne pouvait pourtant qu'accroître leur rémunération (pas de sanction ni d'irréversibilité). Or d'un point de vue économique, la non adhésion à ce dispositif peut être considérée comme irrationnelle et pourrait s'expliquer par l'existence de motivations non financières et/ou la perception par les médecins de risques éthiques associés au P4P. Ainsi, Saint-Lary et al. ont montré que les tensions éthiques associées à la signature du CAPI et le sentiment que ces indicateurs ne permettent pas d'améliorer la qualité des soins ont constitué des freins importants à la non adhésion au contrat (Saint-Lary et al 2011; Saint-Lary et al 2013).

Problématique et questions de recherche

Alors que depuis Janvier 2012, le P4P s'applique à tous les médecins généralistes (MGs) français, de nombreuses questions restent pourtant en suspens :

(1) Quelle est l'efficacité de ces mécanismes incitatifs en termes d'amélioration des indicateurs de qualité des soins définis par le contrat ?

(2) Les MGs sont-ils sensibles à d'autres mécanismes susceptibles d'améliorer les pratiques et pouvant compléter de façon efficiente le P4P?

(3) Doit-on craindre un effet indésirable des P4P sur les motivations intrinsèques des MGs français, engendrant *in fine* une perte d'efficacité économique du dispositif ?

Du fait de la place croissante des P4P dans la rémunération totale des médecins et du rôle attribué aux P4P dans la régulation de l'offre de soins ambulatoires, il apparaît tout d'abord indispensable de s'interroger sur l'efficacité de ces schémas incitatifs et sur leur impact sur les pratiques. En effet, les transpositions des résultats d'études internationales ne sont pas simples du fait de la complexité des systèmes de santé où de nombreux agents interagissent (patients, médecins, assureurs, autres professionnels de santé, financeur public). Par ailleurs, si l'assurance maladie a produit des résultats agrégés des indicateurs avant et après la mise en place du CAPI et de la ROSP, aucune évaluation à proprement parlé de ces dispositifs n'a réellement été mise en œuvre en France. Pour évaluer un dispositif tel que le CAPI, il est nécessaire recours à une modélisation économétrique de l'évolution des indicateurs, permettant notamment de prendre en compte les facteurs de confusion tels que la structure de patientèle (case-mix) et de contrôler ou du moins de discuter des phénomènes de sélection probables associés à la signature du contrat. De façon complémentaire, il apparaît important de s'interroger sur les conséquences des P4P sur des indicateurs plus généraux de la qualité des soins tels que par exemple la durée de consultation, déjà utilisée dans certains mesures multidimensionnelles de la qualité (Howie et al 1997; Howie et al 2000; Mercer and Howie 2006). En effet, à notre connaissance, aucune étude n'a étudié l'impact des P4P sur la durée de consultation des médecins, en particulier dans le contexte du paiement à l'acte qui ne favorise *a priori* pas l'allongement de la durée de consultation (Delattre and Dormont 2003).

D'un point de vue économique, les P4P sont des dispositifs incitatifs parmi d'autres et surtout ils ont un coût non négligeable pour le système de santé. En ce sens, il semble essentiel, dans un second temps, de s'interroger sur la pertinence d'autres mesures

éventuellement complémentaires permettant d'améliorer la qualité des soins de premiers recours en France. En particulier, mieux comprendre les préférences des médecins concernant certaines mesures susceptibles d'être mises en œuvre pour améliorer les pratiques permettrait de mieux anticiper et discuter la pertinence des P4P au regard de dispositifs potentiellement moins coûteux.

Dans le cadre de cette thèse, nous nous concentrons sur un domaine de soins en particulier aussi bien pour évaluer l'impact des P4P que pour analyser les préférences des MGs: les actions de prévention/dépistage des cancers. En effet, il est peu probable que l'impact d'un P4 soit uniforme indépendamment de la nature de l'indicateur ciblé de même que les attentes et les préférences des MGs sont sûrement différentes pour différents domaines de leurs pratiques. Aussi, notre choix se justifie pour plusieurs raisons. Alors que le paiement à l'acte incite peu aux actions de prévention comparé par exemple, à un paiement par capitation (Franc and Lesur 2004), il est intéressant d'analyser si l'instauration de mécanismes de P4P ciblant le dépistage des cancers peut être un outil efficace permettant de « corriger » les défaillances incitatives du paiement à l'acte. En outre, les enjeux majeurs en termes de santé publique associés au dépistage des cancers (Kerlikowske et al 1995; Quinn et al 1999) et la reconnaissance accrue du rôle du MG comme coordinateur des différents dépistages (Bungener et al 2010) justifient une analyse précise des dispositifs (monétaires ou non) pouvant être mis en œuvre pour améliorer les pratiques.

Enfin, les incitatifs économiques appliqués à la médecine peuvent engendrer des effets indésirables potentiels ; de fait, leur évaluation est conditionnée à une meilleure compréhension des motivations et de l'interrelation entre les différents types de motivations des médecins. En particulier, l'efficacité économique des P4P repose sur l'hypothèse d'indépendance entre motivations intrinsèques et extrinsèques des agents et plus particulièrement des professionnels de soins de premiers recours. Autrement dit, l'instauration de récompenses financières/extrinsèques ne doit pas impacter négativement les motivations propres des MGs à fournir des soins de qualité. Ainsi, nous souhaitons examiner la nature des déterminants des motivations intrinsèques afin de les mesurer au mieux. L'idée est ensuite de valider la pertinence de l'hypothèse d'un impact négatif des incitations économiques sur les motivations intrinsèques dans le contexte français, en examinant les relations de substituabilité ou de complémentarité entre MI et ME des médecins généralistes français.

Plan de la thèse

Dans le **premier chapitre**, nous nous intéressons à l'impact en France des paiements à la performance sur différents aspects et indicateurs de la pratique des médecins généralistes français. Dans un premier temps, nous analysons l'impact du CAPI sur une mesure générale supposée être corrélée à la qualité des soins : la durée de consultation. Nous montrons que les signataires du CAPI ont une durée moyenne de consultation plus faible que les non signataires, qui s'explique principalement par un nombre d'actes réalisés dans l'année plus élevé. En revanche, après contrôle de l'ensemble des facteurs de confusion, le CAPI ne semble pas avoir eu d'impact sur la durée de consultation. Si ce premier résultat ne témoigne pas d'une amélioration de la qualité des soins (approchée par la durée de consultation), il ne semble pas non plus s'accompagner d'une dégradation de la qualité (sous certaines hypothèses). Néanmoins, il est intéressant de se demander si ce dispositif coûteux a un impact sur les indicateurs qu'il cible en particulier.

Dans un second temps, nous concentrons notre analyse sur les indicateurs de P4P à proprement parlé, c'est à dire ciblant les actions de prévention/dépistage des cancers. De manière à mieux comprendre le contexte global des incitations (côté offre et côté demande) dans lequel ces indicateurs sont évalués, nous nous intéressons tout d'abord aux déterminants du recours au dépistage des cancers et leur évolution dans le temps, et mettons en évidence que les inégalités socioéconomiques et de recours aux soins n'ont pas diminué dans le temps malgré la mise en place de différentes politiques favorisant le dépistage.

Puis, nous modélisons à partir des données de l'Echantillon Généraliste des Bénéficiaires (EGB) l'évolution du recours au dépistage du cancer du sein lié à la mise en place du CAPI en 2009. A partir d'une modélisation en deux temps, considérant les périodes pre- et post implémentation du CAPI, nous montrons d'une part que les signataires du CAPI n'avaient pas de pratiques de dépistage plus efficaces que les non signataires avant la mise en place du contrat et que, d'autre part, le CAPI semble avoir eu un impact limité sur le recours au dépistage. Ces résultats justifient la recherche de dispositifs complémentaires aux P4P qui permettraient une amélioration des pratiques dans le domaine particulier du dépistage des cancers.

Dans le **deuxième chapitre**, nous étudions les modalités d'une meilleure implication du médecin généraliste dans les dépistages des cancers du sein, du col de l'utérus et du côlon, où coexistent incitations du côté de l'offre (les P4P), et dispositifs collectifs pour favoriser la demande (les dépistages organisés). L'objectif est de mettre en évidence le point de vue des MGs eux-mêmes sur la désirabilité d'autres formes de leviers d'actions susceptibles de compléter efficacement les dispositifs incitatifs existants.

Nous utilisons la méthodologie des choix discrets et construisons un questionnaire pour interroger plus de 400 MGs. L'objectif est de leur faire révéler leurs préférences concernant les conditions ou mesures susceptibles d'améliorer leur implication dans le dépistage des cancers du sein, du col de l'utérus et du côlon et/ou d'améliorer l'efficacité de leur démarche. Nous analysons les résultats en mobilisant deux types de méthodologies qui sont toutes les deux de plus en plus utilisées en économie de la santé: la méthodologie des choix discrets d'une part (Ryan and Gerard 2003; de Bekker-Grob et al 2012) et celle dite du « best-worst scaling » d'autre part (Flynn et al 2007). Ces deux méthodes consistent à présenter différentes épreuves de choix aux médecins de façon à identifier les déterminants structurants de leur fonction d'utilité. Elles diffèrent toutefois dans la conception des épreuves, et permettent de répondre à des questions différentes et complémentaires de la recherche.

Nous confirmons que les MGs sont sensibles à des dispositifs non monétaires de la pratique tels que la formation ou des résumés pluriannuels de patients dépistés et seraient prêts à renoncer à des montants élevés de P4P pour en disposer, notamment dans le cadre des dépistages des cancers féminins. En effet, nous montrons que les préférences varient en fonction de la nature du cancer concerné par le dépistage. Ces résultats attestent d'une grande diversité des préférences et des motivations des médecins généralistes français, justifiant une analyse précise des différents registres de motivation et de leurs relations d'interdépendance.

Dans le **troisième chapitre**, nous analysons empiriquement la relation entre motivations intrinsèques (MI) et motivations extrinsèques (ME) des médecins généralistes français à partir d'une enquête effectuée auprès de MGs de la région Midi-Pyrénées. L'objectif de ce travail est de mettre en évidence les risques potentiels liés à l'augmentation des ME par l'instauration d'incitatifs financiers (CAPI, ROSP), sur les MI des MGs français.

Dans un premier temps, nous comparons différentes méthodes statistiques pour la construction d'un score composite de la MI des médecins. Les méthodes statistiques mises en œuvre permettent dans un premier temps de valider des indicateurs de MI, préalablement sélectionnés *a priori* notamment à partir des résultats de la littérature. Puis, dans un second temps, il s'agit de construire un score composite permettant de classer les médecins le long d'une échelle latente de motivations intrinsèques. Les indicateurs des MI sont sélectionnés de façon à refléter différentes dimensions du concept tel que développé par les psychologues (Deci 1971; Deci and Ryan 2000) et les chercheurs en sciences sociales (Osterloh and Frey 2000; Lindenberg 2001). Les méthodes développées pour la construction du score permettent toutes d'analyser le nombre de dimensions communes aux indicateurs présélectionnés et de prendre en compte les erreurs de mesures inhérentes aux données déclaratives. Nous montrons que le classement des médecins le long de l'échelle de MI varie peu selon la méthode utilisée, validant la robustesse du score obtenu.

Dans un second temps, nous sélectionnons le modèle le plus parcimonieux de construction du score et étudions la relation de complémentarité ou substituabilité entre ce score et des indicateurs définis de motivations extrinsèques. Nous montrons que des médecins motivés intrinsèquement ont un degré plus faible de ME telles que mesurées par différents indicateurs qualitatifs (non satisfaction de la rémunération, sentiment d'être contraints par les exigences des patients), témoignant ainsi d'une possible relation de substitution entre MI et ME. Dans le contexte d'intensification des incitations économiques au travers notamment des P4P, ces résultats invitent à davantage de prudence dans la conception des indicateurs et mettent en garde contre des effets potentiellement contre-productifs de ces schémas.

Chapitre 1

Impact des incitatifs économiques sur différents indicateurs de la qualité des soins en médecine générale

1. Impact du CAPI sur la durée de consultation des médecins généralistes français²

1.1. Introduction

Le paiement à la performance (P4P) appliqué aux soins primaires a connu un fort développement ces dix dernières années (Rosenthal MB and Dudley R 2007). Au Royaume-Uni, pionnier en la matière, le P4P est généralisé à l'ensemble des médecins généralistes (MG) depuis le printemps 2004 avec le *Quality of Outcomes Framework* (Doran et al 2006). D'autres pays comme les Etats-Unis, l'Australie, la Nouvelle-Zélande ou encore Israël ont également adopté ce mode additionnel de paiement (Morgan and Dunbar 2011). Depuis 2009, un système national de P4P a été introduit en France à destination des médecins généralistes (MGs) au travers du *Contrat d'Amélioration des Pratiques Individuelles* (CAPI). Un des objectifs du CAPI est de fournir une incitation aux médecins à améliorer la qualité des soins mesurée au travers de 16 indicateurs couvrant trois principaux domaines : la prévention/dépistage, le suivi des maladies chroniques et l'optimisation de la prescription. Le CAPI a été initié par l'UNCAM en 2009 contre l'avis des syndicats de médecins. Ces derniers avaient la possibilité de signer le contrat à titre volontaire, leur permettant de recevoir un bonus représentant un montant maximal de 5 000€ (environ 7% du chiffre d'affaire moyen) conditionné à la réalisation ou l'amélioration des indicateurs définis par le contrat. Aucune sanction n'était prévue, et il était possible de quitter le programme à tout moment sur demande. Le choix des indicateurs et leur niveau ainsi que la formule permettant de calculer le montant des bonus ont été définis unilatéralement par l'Assurance Maladie (cf. figure A.1, Annexe A, pour une description du calcul des indicateurs). Entre le début et la fin du dispositif (i.e. Juillet 2009 et Décembre 2011), 38% des médecins éligibles ont souscrit au contrat. Les médecins non éligibles au contrat étaient les médecins remplaçants ou récemment installés et n'ayant pas une patientèle suffisamment grande. A la fin du mois de Juin 2009, 70% des signataires avaient reçu un bonus de 3 100€ en moyenne par an variant entre 1 500€ (premier décile) et 4 900€ (dernier décile) (Commission des comptes 2011).

L'efficacité des programmes de P4P fait l'objet de vifs débats, notamment concernant leur capacité à générer des soins de meilleure qualité (Mannion and Davies 2008; Ryan 2010;

² Cette section a fait l'objet d'un article coécrit avec Olivier Saint-Lary accepté dans la revue *Health Policy* sous le titre : « Impact of a pay-for performance programme on French GPs' consultation length ».

Woolhandler et al 2012; Jha 2013; Eijkenaar et al 2013). Si de nombreux travaux ont tenté d'évaluer l'efficacité du P4P en termes d'amélioration d'indicateurs mesurés (Scott et al 1996; Campbell et al 2009; Ryan and Doran 2012) ou de réduction des coûts (Cheng et al 2012b), à notre connaissance, aucune étude n'a évalué l'impact de ce type de dispositif sur la durée de consultation des médecins. Il a déjà été montré que la durée de la consultation est un facteur déterminant de la satisfaction des patients (Howie et al 1989; Wilson 1991; Wilson and Childs 2002), qu'elle est souvent associée au traitement de problèmes de santé chroniques et d'actions de prévention (Howie et al 1991) ainsi qu'à des soins de meilleure qualité dans les cas plus difficiles faisant intervenir des problèmes psychologiques et sociaux (Howie et al 1997; Freeman 2002; Gude et al 2013). Ainsi, dans leurs recommandations concernant les patients atteints de maladies chroniques, le Collège Australien des Médecins Généralistes encourage les médecins à prévoir des durées de consultation plus longues afin de déterminer les besoins des patients et de formuler un plan de gestion approprié (RACGP 2011). De plus, l'indice anglais de consultation de qualité (« Consultation Quality Index », CQI), qui est un marqueur globale de la qualité pour les MGs, repose sur trois dimensions parmi lesquelles figure la durée de consultation (Howie et al 2000; Mercer and Howie 2006).

Considérer la durée de consultation comme un indicateur de la qualité des soins semble particulièrement pertinent dans le contexte de paiement à l'acte qui *a priori* ne favorise pas l'allongement de la durée de consultation (Delattre and Dormont 2003; Rochaix 2004). Dans ce contexte, nous faisons l'hypothèse que la durée de consultation peut être considérée comme un proxy de la qualité des soins.

L'analyse de l'impact des incitatifs économiques sur le comportement des producteurs de soins a déjà fait l'objet d'une abondante littérature à la fois théorique et empirique (Ellis and McGuire 1986; Ellis and McGuire 1993; Whynes and Baines 1998; Eggleston 2005; Kaarboe and Siciliani 2011; Hennig-Schmidt et al 2011; Brosig-Koch et al 2013). Conformément à la théorie économique standard (modèle principal-agent) et à la théorie des contrats en particulier, l'introduction de bonus concernant la réalisation de certaines tâches fixées par le contrat permet, sous certaines conditions, d'augmenter le niveau d'effort consacré à ces tâches, effort qui peut représenter une durée du service produit qui s'allongerait sous l'effet de l'incitation (Dewatripont et al 2000; Eggleston 2005). Toutefois, il faut que certaines hypothèses soient vérifiées : dans le cas des modèles d'arbitrage travail-loisir, il faut notamment que le travail et le loisir soient substituables et que l'effet substitution du travail au loisir l'emporte sur l'effet revenu (Rizzo and Blumenthal 1994; Chanel et al 2010). Dans le

cadre des modèles multi-tâches, il est nécessaire que les indicateurs mesurent la qualité de façon exhaustive (Eggleston 2005). Ainsi, la littérature empirique a montré des résultats contradictoires. Chanel & al (2010) ont étudié de façon expérimentale la modification de l'offre de consultation (heures travaillées) suite à un changement hypothétique des honoraires des MGs, et ont montré que la plupart des médecins ne changeaient pas leur offre de travail et que certains la diminuaient, traduisant une élasticité nulle voire négative du travail au loisir. D'autres auteurs ont analysé l'impact du passage d'un mode contractuel moins incitatif et ont montré que ce changement a réduit la durée de consultation des médecins (Dumont et al 2008). L'opinion des médecins généralistes interrogés sur ce point en 2010 allait également dans ce sens, considérant le CAPI comme un moyen d'augmenter le temps consacré à certaines activités telles que les pratiques de prévention ou le suivi des maladies chroniques (Saint-Lary et al, 2011).

Le caractère volontaire de l'adhésion au CAPI peut engendrer des phénomènes d'auto-sélection. D'un côté, les signataires du CAPI pourraient être plus « productifs » en termes de nombre d'actes (Devlin and Sarma 2008) et donc pourraient avoir des durées de consultation plus courtes. D'un autre côté, les signataires du CAPI pourraient être ceux qui considèrent qu'ils font des soins de meilleure qualité, se traduisant par une durée de consultation plus longue. Cette étude constitue un travail préliminaire dont le but est d'étudier la relation entre signature du CAPI et durée de consultation des MGs français, de façon à générer des hypothèses quant à l'impact de mécanismes de P4P sur la qualité des soins. Du fait de la diversité des motifs de consultations engendrant une grande variété du contenu de la consultation et donc de sa durée, nous choisissons par ailleurs de nous concentrer sur certains motifs de consultation spécifiques ciblés par des indicateurs du CAPI, et associés au traitement de maladies chroniques (diabète, hypertension) ou des actions de prévention (vaccination). Nous décrivons les données et méthodes utilisées dans la section 2, les résultats sont présentés dans la section 3 et discutés dans la section 4.

1.2. La base de données Ecogen

Les données que nous avons utilisées proviennent de la base de données *Ecogen* ('Eléments de CONsultation en medecine GENérale'), constituée à partir du travail de 54 internes en médecine générale ayant retranscrit les motifs et résultats de consultation de 128 médecins généralistes recrutés sur la France entière, entre le 1er novembre 2011 et le 30 avril

2012³. Ces médecins étaient tous maîtres de stage et ont participé au projet de manière volontaire. Les données ont été collectées de façon continue durant les 6 mois d'inclusion : chaque interne passait au moins une journée par semaine au cabinet du médecin tuteur et les consultations étaient retranscrites de manière exhaustive. Ainsi, la base de données contient 20 779 observations de consultation, soit environ 162 observations par médecin. Chaque observation donne des informations précises sur le contexte de la consultation (date, consultation ou visite), les caractéristiques du patient (âge, sexe, profession, statut d'exonération, nouveau patient), et les caractéristiques du médecin (âge, sexe, localisation géographique, nombre total de consultations en 2011 à partir du RIAP⁴, secteur conventionnel, mode d'exercice, signature du CAPI). Par ailleurs, pour chaque consultation sont renseignés les différents motifs et résultats de consultation (jusqu'à sept motifs/résultats) suivant l'ICPC-2 (*International Classification of Primary Care*, 2^e Edition) ainsi que la durée précise de consultation chronométrée par l'interne.

1.3. Méthodologie statistique

1.3.1. Le modèle multi-niveaux de la durée de consultation

La méthodologie statistique adoptée est une modélisation multi-niveau comportant deux niveaux : le niveau séance (niveau 1) et le niveau médecin (niveau 2). Au niveau 1 sont introduites les variables caractérisant le patient ou la séance i ⁵. Au niveau 2 sont introduites les variables caractérisant le médecin j . Ce type d'approche permet de distinguer les composantes de la variabilité de la durée de consultation liées à la variance inter- ou intra-médecin.

Le modèle général à estimer s'écrit de la façon suivante :

$$\begin{cases} y_{ij} = \beta_{0j} + \beta X_{ij} + \varepsilon_{ij} & (\text{niveau 1}) \\ \beta_{0j} = \gamma_{00} + \delta CAPI_j + \gamma_1 K_j + \gamma_2 X_{.j} + u_{0j} & (\text{niveau 2}) \end{cases}$$

³ Nous remercions le Professeur Laurent Letrillart, coordinateur du projet, pour nous avoir permis d'exploiter ces données.

⁴ Le 'Relevé Individuel d'Activité Professionnelle' (RIAP) est transmis annuellement aux médecins par l'Assurance maladie.

⁵ Nous ne pouvons pas construire de modèle à trois niveaux en distinguant le niveau patient et séance car aucun identifiant patient n'a été codé dans la base.

La variable d'intérêt y_{ij} est le logarithme de la durée de la consultation i pour le médecin j ; γ_{00} est la constante générale du modèle et β_{0j} est la constante à estimer pour le médecin j , représentant sa durée de consultation moyenne. Le vecteur X_{ij} contient les covariables caractérisant le patient ou la consultation i , la variable $CAP I_j$ est une variable indicatrice indiquant si le médecin j a signé ou non le CAPI entre 2009 et 2011. Le vecteur K_j contient les covariables caractérisant le médecin j et le vecteur $X_{.j}$ contient les variables patients ou consultation introduites en moyenne afin de contrôler des structures de patientèle (Mundlak 1978). Enfin, deux termes d'erreur sont introduits dans le modèle : u_{0j} est le terme d'erreur au niveau 2 représentant l'influence des caractéristiques inobservées variant d'un médecin à l'autre, et ε_{ij} est le terme d'erreur au niveau 1 représentant l'influence sur la durée de consultation des autres variables omises variant d'une consultation à l'autre et d'un patient à l'autre.

1.3.2. Les hypothèses du modèle

Dans ce modèle, plusieurs hypothèses d'exogénéité des variables sont implicitement faites. Les variables au niveau 1 (X_{ij}) ne sont pas corrélées aux termes d'erreur de niveau 1 (*hypothèse 1*) et de niveau 2 (*hypothèse 2*). Une source de violation de l'hypothèse 2 peut provenir d'effets d'auto sélection de la patientèle : certaines catégories de patients (par exemple les cadres) sont susceptibles de privilégier les médecins pratiquant des séances plus longues (Breuil-Genier and Goffette 2006). L'introduction parmi les variables explicatives des moyennes des caractéristiques des patients ($X_{.j}$) est censée pallier ce problème (approche de Mundlack). Par ailleurs, la signature du CAPI n'est pas corrélée aux termes d'erreur de niveau 1 (*hypothèse 3*) et de niveau 2 (*hypothèse 4*). Or il est peu probable que ces hypothèses soient vérifiées dans la mesure où les signataires du CAPI se distinguent probablement des non signataires sur un ensemble de variables observables (âge, sexe, nombre de consultations, caractéristiques des patients) et inobservables (motivation, style de pratique). L'introduction parmi les variables de contrôle des variables caractérisant le médecin (K_j) et du case-mix ($X_{.j}$) permet de contrôler des phénomènes de sélection sur les variables observables.

1.3.3. Les différents modèles multi-niveaux estimés

Trois modèles multiniveaux sont estimés, incorporant successivement les variables séance/patients (modèle 1), les variables médecins (modèle 2), et les moyennes des

caractéristiques des patients (modèle 3). Les trois modèles s'écrivent sous leur forme développée de la façon suivante :

$$y_{ij} = \gamma_{00} + \beta X_{ij} + \delta CAPI_j + u_{0j} + \varepsilon_{ij} \quad (\text{modèle 1})$$

$$y_{ij} = \gamma_{00} + \beta X_{ij} + \delta CAPI_j + \gamma_1 K_j + u_{0j} + \varepsilon_{ij} \quad (\text{modèle 2})$$

$$y_{ij} = \gamma_{00} + \beta X_{ij} + \delta CAPI_j + \gamma_1 K_j + \gamma_2 X_{.j} + u_{0j} + \varepsilon_{ij} \quad (\text{modèle 3})$$

Parmi les caractéristiques des médecins incluses dans les modèles 2 et 3 figure le nombre de consultations en 2011, permettant de mesurer l'intensité de l'activité médicale du médecin supposée très corrélée à la durée de consultation. En revanche, le sens de la causalité est incertain puisque, dans le contexte du paiement à l'acte, le médecin peut aussi choisir son activité médicale en fonction de sa durée moyenne de consultation. Économétriquement parlant, le nombre de consultation par an en 2011 est potentiellement endogène dans les modèles 2 et 3, ce qui peut produire des estimations inconsistantes des paramètres du modèle. Ainsi, des modèles distincts sont estimés incluant cette variable ou non : cette variable n'est pas incluse dans les modèles 2a et 3a, mais nous l'incluons les modèles 2b et 3b.

Ces trois modèles sont estimés dans un premier temps sur les observations tout motif de consultation confondu. Puis, afin de s'assurer que les résultats ne varient pas d'un type de consultation (pour une raison donnée) à un autre, le modèle 3 - supposé le plus robuste - est estimé pour des différents motifs de consultations ciblés par les indicateurs du CAPI: le diabète, l'hypertension artérielle et la vaccination contre la grippe. À nouveau, des analyses séparées sont effectuées incluant ou non le nombre de consultation en 2011 comme variable de contrôle.

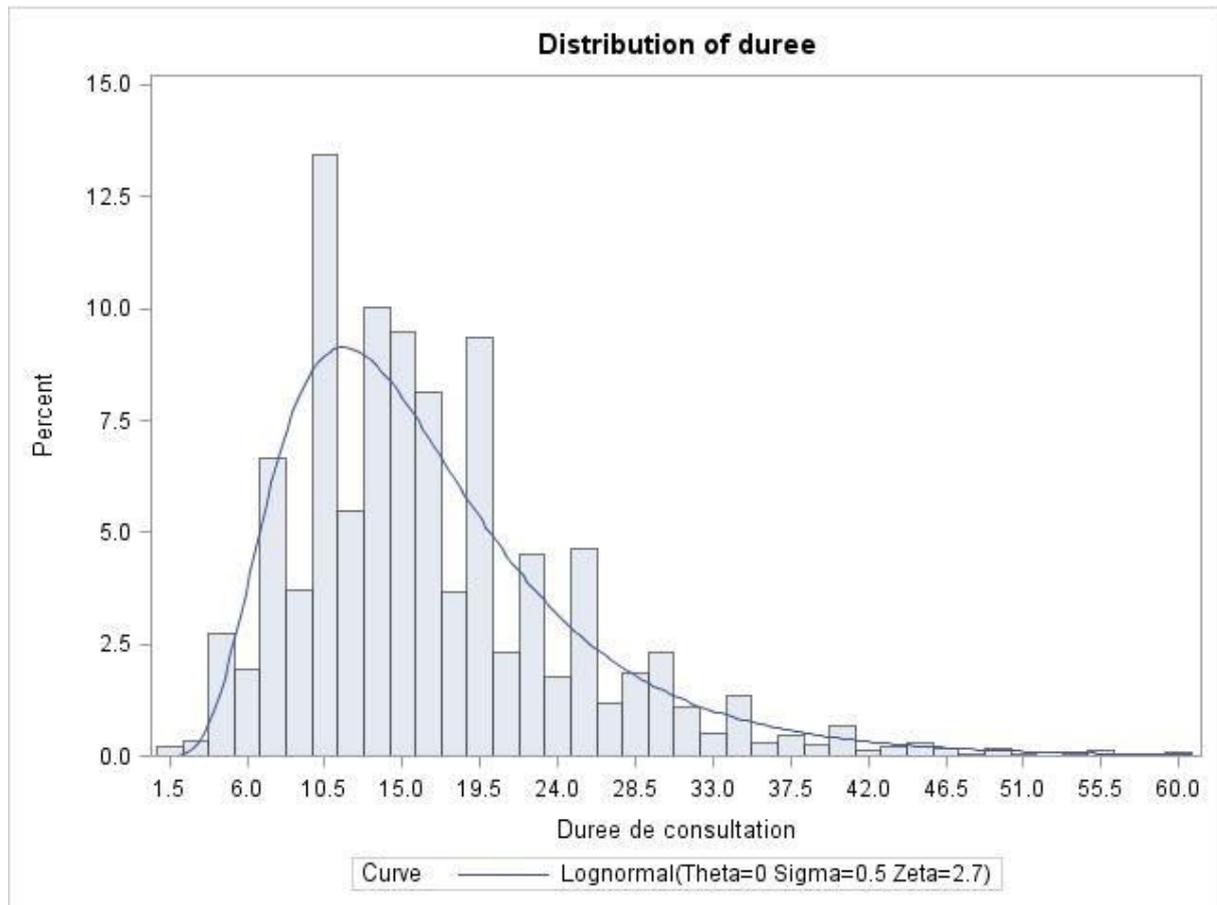
1.4. Statistiques descriptives et modèles univariés

1.4.1. Distribution de la durée de consultation

La durée de consultation moyenne dans l'échantillon est de 16.8 minutes, après suppression des valeurs extrêmes. Nous avons considéré comme valeurs extrêmes les durées supérieures à 1 heure de consultation, soit 114 consultations (i.e. 0.5% des consultations). Trois consultations sur quatre durent entre 5 et 20 minutes et plus de 90% des consultations

durent entre 5 et 30 minutes (Figure 1). Cette distribution a une allure log-normale, justifiant ainsi de modéliser le logarithme de la durée de consultation.

Figure 1. Distribution de la durée de consultation (N= 20 665 consultations).



1.4.2. Déterminants de la durée de consultation

Les statistiques descriptives sur les déterminants des durées de consultation sont présentées dans le tableau 1. Les médecins de l'échantillon sont plus nombreux à ne pas avoir signé le CAPI (57%) et sont majoritairement des hommes (66%), ayant plus de 45 ans (80%), en secteur 1 (92%), ayant un mode d'exercice regroupé (79%). Par ailleurs, ils sont assez bien représentés dans toutes les régions françaises, mise à part le Sud de la France où ils sont minoritaires (6%) et exercent à la fois en région rurale/semi-rurale (48%) et en région urbaine.

Par ailleurs, nous avons effectué des tests d'indépendance du Chi-deux mesurant les différences de caractéristiques individuelles entre signataires et non signataires du CAPI. Ainsi, il n'y a pas de différence significative entre eux deux groupes, à l'exception du

nombre de consultations en 2011, plus élevée pour les signataires du CAPI ($p=0,0262$). Comme le nombre de consultations annuel est observé seulement après la signature du CAPI, il n'est pas possible de déterminer s'il s'agit d'un effet d'auto-sélection ou si, à l'inverse, le CAPI a entraîné une intensification de l'activité médicale⁶.

Les variables relatives à la séance ou au patient telles que l'âge, le genre ou la catégorie socio professionnelle ainsi que les variables médecins ont un impact significatif sur la durée de consultation dans les analyses univariées (tableau 1). Les médecins ayant signé le CAPI ont une durée de consultation significativement plus courte (15,6 minutes en moyenne) que les non signataires (17,6 minutes en moyenne) ($p<0,001$).

⁶ Les deux hypothèses sont discutées dans la section suivante. Par ailleurs, sous l'hypothèse d'auto-sélection, le nombre total de consultations en 2011 n'a pas pu être utilisé comme instrument de la signature du CAPI car il est aussi un prédicteur significatif de la durée de consultation.

Tableau 1. Statistiques descriptives et modèles univariés

	Statistiques descriptives			Analyses univariées	
	CAPI (non) N (col %) ^a	CAPI (oui) N (col %)	Total N (%)	Durée moyenne (en min)	Ecart brut à la situation de référence
Variables séance/patient (N=20 779)					
Age du patient					*** ^b
0 à 16 ans	2111 (17.6)	1392 (15.9)	3503 (16.9)	14.8	-18%
17 à 44 ans	3427 (28.5)	2337 (26.6)	5764 (27.7)	15.9	-12%
45 à 69 ans	4007 (33.4)	3074 (35)	7081 (34.1)	17.6	-3%
70 ans et plus	2461 (20.5)	1969 (22.4)	4430 (21.3)	18.1	ref
Sexe du patient					***
Homme	4937 (41.1)	3736 (42.6)	8673 (41.7)	16.3	-5%
Femme	7070 (58.9)	5036 (57.4)	12106 (58.3)	17.1	ref
Profession du patient					***
Agriculteur-Artisan	357 (3)	292 (3.3)	649 (3.1)	17	10%
Cadre sup	725 (6)	327 (3.7)	1052 (5.1)	19	23%
Prof intermédiaire	790 (6.6)	457 (5.2)	1247 (6.0)	17	10%
Employé	2253 (18.8)	1761 (20.1)	4014 (19.3)	16.4	6%
Ouvrier	452 (3.8)	368 (4.2)	820 (4.0)	14.9	-4%
Retraité	3821 (31.8)	2979 (34.0)	6800 (32.7)	17.9	15%
Inactif	3604 (30)	2585 (29.5)	6189 (29.8)	15.5	ref
Statut d'exonération du patient					***
Motif ALD	2502 (20.8)	2030 (23.2)	4532 (21.8)	18.6	14%
CMU ou AME	428 (3.6)	406 (4.6)	834 (4.0)	14.8	-9%
Autre	550 (4.6)	372 (4.2)	922 (4.4)	17	4%
Aucun	8523 (71.0)	5957 (68.0)	14480 (69.7)	16.3	ref
Contexte					***
Cabinet	11191 (93.2)	8313 (94.8)	19504 (93.9)	16.6	-14%
Visite	816 (6.8)	459 (5.2)	1275 (6.1)	19.2	ref
Nouveau patient					***
Oui	719 (6.0)	427 (4.9)	1146 (5.5)	17.6	5%
Non	11288 (94.0)	8345 (95.1)	19633 (94.5)	16.7	ref
Année					***
Année 2011	2825 (23.5)	2142 (24.4)	4933 (23.9)	16.3	-4%
Année 2012	9182 (76.5)	6630 (75.6)	15732 (76.1)	16.9	ref
Variables médecin (N=128)					
Signature CAPI					***
Oui	-	-	55 (43.0)	15.6	-11%
Non	-	-	73 (57.0)	17.6	ref
Age		$p^c = 0.8507$			ns
30-44 ans	13 (17.8)	12 (21.8)	25 (19.5)	16.1	-8%
45-54 ans	29 (39.7)	21 (38.2)	50 (39.1)	16.2	-7%
55 ans et plus	31 (42.5)	22 (40.0)	53 (41.4)	17.5	ref
Sexe		$p = 0.5767$			**
Homme	47 (64.4)	38 (69.1)	85 (66.4)	16	-12%
Femme	26 (35.6)	17 (30.9)	43 (33.6)	18.2	ref

^a Lecture (% en colonne): parmi les médecins non signataires du CAPI, 17.6% des patients ont entre 0 et 16 ans, contre 15.9% chez les médecins signataires du CAPI, ce chiffre étant de 16.9% pour l'ensemble des médecins.

^b Test de Fisher: relation entre les variables indépendantes et la durée de consultation (en log).

Niveau de significativité: ***=1% ; **=5% ; *=10%.

^c Test d'indépendance du Chi-deux : relation entre les caractéristiques du médecin et la signature du CAPI.

Tableau 1. Statistiques descriptives et modèles univariés (suite)

	Statistiques descriptives			Analyses univariées	
	CAPI (non) <i>N (col %)</i>	CAPI (oui) <i>N (col %)</i>	Total <i>N (%)</i>	Durée moyenne (en min)	Ecart brut à la situation de référence
Region		<i>p = 0.5624</i>			<i>ns</i>
Centre	12 (16.9)	8 (14.8)	20 (16.0)	20.9	25%
Est	20 (28.2)	14 (25.9)	34 (27.2)	21.3	28%
Ile de France	13 (18.3)	10 (18.5)	23 (18.4)	23.5	41%
Nord	14 (19.7)	6 (11.1)	20 (16.0)	16	-4%
Sud	3 (4.2)	5 (9.3)	8 (6.4)	16.5	-1%
Ouest	9 (12.7)	11 (20.4)	20 (16.0)	16.7	ref
Nombre de consultations en 2001		<i>p = 0.0262</i>			***
Moins de 4000	26 (35.6)	9 (16.4)	35 (27.3)	18.8	32%
De 4000 à 5000	24 (32.9)	15 (27.3)	39 (30.5)	18.2	28%
De 5000 à 6000	9 (12.3)	11 (20)	20 (15.6)	14.9	5%
Plus de 6000	14 (19.2)	20 (36.4)	34 (26.6)	14.2	ref
Secteur		<i>p = 0.5137</i>			*
Secteur 1	66 (90.4)	52 (94.5)	118 (92.2)	17.8	-20%
Secteur 2 (honoraires libres)	7 (9.6)	3 (5.5)	10 (7.8)	22.2	ref
Milieu d'exercice		<i>p = 0.3358</i>			ns
Rural	14 (19.2)	15 (27.3)	29 (22.7)	16.1	-8%
Semi-rural	22 (30.1)	11 (20)	33 (25.8)	15.8	-10%
Urbain	37 (50.7)	29 (52.7)	66 (51.6)	17.5	ref
Mode d'exercice		<i>p = 0.9463</i>			ns
Groupe médecins	45 (61.6)	34 (61.8)	79 (61.7)	16.6	-2%
Groupe pluridisciplinaire	12 (16.4)	10 (18.2)	22 (17.2)	17.1	1%
Solo	16 (21.9)	11 (20)	27 (21.1)	16.9	ref

1.5. Résultats des modèles multi-niveaux

1.5.1. Modèles tout motif de consultation confondu

Lorsqu'on considère les observations tout motif de consultation confondu, plus d'un quart (28%⁷) de la variabilité de la durée de consultation est liée à des caractéristiques au niveau médecin (variabilité inter-médecin), et 72% est liée aux caractéristiques relatives aux séances ou aux patients (variabilité intra-médecin).

Les résultats des trois modèles multivariés sont présentés dans le tableau 2. Dans le modèle 1, à caractéristiques des patients données, les médecins signataires du CAPI ont une durée de consultation significativement inférieure à celle des non signataires (-14,2%, $p < 0.001$). Par ailleurs, la durée de consultation est significativement plus faible pour les

⁷ Ce résultat correspond au coefficient de corrélation intra-classe. Il est estimé à partir du modèle multi niveau « vide », c'est-à-dire sans variable explicative.

patients plus jeunes (moins de 44 ans), de sexe masculin, employés ou ouvriers, bénéficiant de la Couverture Maladie Universelle (CMU) ou de l'Aide Médicale d'Etat (AME)⁸ et dans le contexte d'une consultation en cabinet (par rapport à une visite); elle est significativement plus longue pour les nouveaux patients, les agriculteurs, artisans, cadres ou retraités, et les individus ayant une affection de longue durée (ALD).

Dans le modèle 2, les résultats sont contrôlés des caractéristiques du médecin. Les modèles 2a et 2b montrent des résultats différents quant à l'impact de la signature du CAPI sur la durée de consultation : dans le modèle 2a, les signataires ont une durée de consultation significativement plus courte que les non signataires (-13.1%, $p=0.005$), alors que dans le modèle 2b (qui contrôle du nombre de consultations en 2011), cette différence n'est plus significative (-7.4%, $p=0.102$). En ce qui concerne les variables médecin, être un homme, avoir un nombre de consultation élevé en 2011 (modèle 2b) et être en secteur 1, sont associés à une durée de consultation plus courte. Les caractéristiques du médecin introduites expliquent 20% (resp. 31%) de la variabilité inter-médecin dans le modèle 2a (resp. dans le modèle 2b).

Dans le modèle 3, les moyennes des caractéristiques des patients sont introduites en sus des variables médecin et patient. Plus la proportion de femmes et/ou d'agriculteurs dans la patientèle est élevée, plus la durée de consultation est élevée. Toutes choses égales par ailleurs, l'impact du CAPI est toujours négatif mais n'est plus significatif (au seuil de 5%) dans les modèles 3a et 3b. Les caractéristiques de la patientèle ainsi introduites dans les modèles 3a et 3b expliquent 13% de la variabilité inter-médecin.

⁸ La CMU permet à toute personne résidant en France de façon stable et régulière et ne pouvant relever à aucun autre titre d'une couverture maladie obligatoire, de bénéficier de la sécurité sociale. Au 28 février 2014, 2.3 millions de personnes bénéficiaient de la CMU (www.ameli.fr). L'AME est destinée aux personnes en situation irrégulière au regard de la réglementation française sur le séjour en France. Elle est attribuée sous conditions de résidence et de ressources.

Tableau 2. Modèles multivariés de la durée de consultation (log)

	Modèle 1		Modèle 2		Modèle 3	
			Modèle 2a	Modèle 2b	Modèle 3a	Modèle 3b
Intercept	2.85***		2.943***	2.831***	2.606***	2.448***
Variables de niveau 1						
<i>Age</i>				<i>ref= + 70 ans</i>		
0 à 16 ans	-0.088***		-0.087***	-0.087***	-0.069***	-0.069***
25 à 44 ans	-0.073***		-0.071***	-0.071***	-0.063***	-0.063***
45 à 69 ans	0.008		0.009	0.009	0.015	0.015
<i>Sexe</i>				<i>ref= femme</i>		
Homme	-0.028***		-0.029***	-0.029***	-0.033***	-0.033***
<i>Profession</i>				<i>ref= employé</i>		
Agriculteur-Artisan	0.041**		0.039**	0.039**	0.032	0.032
Cadre sup	0.05***		0.05***	0.05***	0.051***	0.051***
Prof intermédiaire	0.019		0.018	0.018	0.022	0.022
Employé	-0.025		-0.022	-0.022	-0.022	-0.022
Ouvrier	0.021*		0.022*	0.022*	0.025*	0.025*
Retraité	0.001		0.002	0.001	-0.005	-0.005
<i>Statut d'exonération</i>				<i>ref= aucun</i>		
Motif ALD	0.087***		0.089***	0.089***	0.094***	0.094***
CMU ou AME	-0.046***		-0.044***	-0.044***	-0.041**	-0.041**
Autre	0.059***		0.064***	0.064***	0.077***	0.077***
<i>Contexte</i>				<i>ref= visite</i>		
Cabinet	-0.081***		-0.08***	-0.08***	-0.098***	-0.098***
<i>Nouveau patient</i>				<i>ref= non</i>		
Yes	0.096***		0.097***	0.096***	0.096***	0.096***
<i>Année</i>				<i>ref= 2012</i>		
Année 2011	-0.008		-0.008	-0.008	-0.006	-0.007
Variables de niveau 2						
<i>Signature CAPI</i>				<i>ref= non</i>		
Oui	-0.141***		-0.131***	-0.074	-0.087*	-0.032
Autres variables médecin						
<i>Age</i>				<i>ref= + 55 years</i>		
30-44 ans			-0.051	-0.070	0.048	0.023
45-54 ans			-0.035	-0.031	-0.052	-0.040
<i>Sexe</i>				<i>ref= femme</i>		
Homme			-0.123**	-0.092*	-0.054	-0.044
<i>Region</i>				<i>ref= ouest</i>		
Centre			0.063	-0.030	0.009	-0.080
Est			0.121	0.046	0.095	0.026
Ile de France			0.025	-0.078	0.019	-0.072
Nord			0.120	0.018	0.075	-0.028
Sud			-0.022	-0.022	0.068	0.057
<i>Secteur</i>				<i>ref= secteur 1</i>		
Secteur 2			0.176*	0.18*	0.004	0.099
<i>Milieu d'exercice</i>				<i>ref= urbain</i>		
Rural			-0.060	-0.087	-0.027	-0.101
Semi-rural			-0.009	-0.033	0.046	0.000
<i>Mode d'exercice</i>				<i>ref= solo</i>		
Groupe médecins			-0.044	-0.039	0.015	0.015
Groupe pluridisciplinaire			-0.061	-0.028	0.029	0.067

Les p-values sont indiquées par des étoiles. Niveau de significativité: ***=1% ; **=5% ; *=10%.

Tableau 2. Modèles multivariés de la durée de consultation (log) (suite)

	Modèle 1		Modèle 2		Modèle 3	
			Modèle 2a	Modèle 2b	Modèle 3a	Modèle 3b
<i>Nbr consultations en 2011</i>				<i>ref= + 6000</i>		
Moins de 4000				0.227***		0.209***
De 4000 à 5000				0.232***		0.232***
De 5000 à 6000				0.031		0.043
Variabiles de patientèle						
Proportion de femmes					0.005	0.005
Proportion de 0-16 ans					-0.004	-0.001
Proportion de +70 ans					0.000	0.001
Proportion d'agriculteurs					0.012	0.017
Proportion de cadres					0.011*	0.009
Proportion d'ouvriers					-0.012	-0.009
Proportion de nouveaux patients					-0.005	-0.007
Variance des effets aléatoires						
Niveau 1 (ε _{ij})	0.1833***	0.1911***	0.1911***	0.1911***	0.1914***	0.1914***
Niveau 2 (u _{oj})	0.0683***	0.0605***	0.0518***	0.0518***	0.0496***	0.0428***

1.5.2. Modèles par motif de consultation

Les résultats des modèles multivariés par motif de consultation sont présentés dans le tableau 3. Trois motifs de consultation faisant l'objet d'un ou plusieurs indicateurs CAPI sont sélectionnés : le diabète (N= 2821 consultations), l'hypertension (N= 3417 consultations) et la vaccination anti-grippale (N= 1356 consultations). Le fait de travailler sur des consultations plus homogènes diminue la variabilité intra-médecin pour les motifs de consultations liées au diabète (62% de variabilité intra-médecin) et à l'hypertension (61% de variabilité intra-médecin) mais pas pour le motif de consultation lié à la vaccination antigrippale (79% de variabilité intra-médecin).

Pour ces trois motifs de consultation, les durées de consultation sont plus élevées pour les nouveaux patients. Par ailleurs, les patients plus jeunes et ceux bénéficiant de la CMU ou de l'AME ont une durée de consultation plus courte pour les consultations ayant pour motif le diabète. L'effet du CAPI sur la durée de consultation est similaire d'un motif de consultation à l'autre, le coefficient étant négatif mais non significatif. En revanche, la valeur du coefficient varie légèrement en fonction du motif de consultation et selon le fait que la variable « nombre de consultation en 2011 » soit incluse ou non dans le modèle. Dans les modèles 3a, le coefficient associé à l'effet du CAPI varie de -9.3% (diabète) à -3.5% (vaccination) ; dans les modèles 3b, il varie de -4.3% (diabète) à 0.4% (vaccination).

Tableau 3. Modèles multivariés de la durée de consultation (log) selon le motif de consultation

	Diabète		Hypertension		Vaccination anti-grippale	
	Modèle 3a	Modèle 3b	Modèle 3a	Modèle 3b	Modèle 3a	Modèle 3b
Constante	2.276***	2.168***	2.47***	2.332***	2.667***	2.624***
Variables de niveau 1						
<i>Age</i>	<i>ref= + 70 years</i>					
0 à 16 ans	-0.161***	-0.161***	0.021	0.023	0.056	0.066
17 à 44 ans	-0.1***	-0.099***	0.063	0.064	-0.127	-0.118
45 à 69 ans	-0.014	-0.014	0.025	0.025	0.028	0.037
<i>Sexe</i>	<i>ref= femme</i>					
Homme	0.023	0.022	-0.023	-0.024	-0.034	-0.034
<i>Profession</i>	<i>ref = employé</i>					
Agriculteur-Artisan	0.049	0.049	0.048	0.047	0.041	0.037
Cadre sup	0.075*	0.075*	-0.004	-0.004	0.073	0.067
Prof intermédiaire	0.036	0.035	0.036	0.036	0.023	0.017
Employé	0.022	0.02	-0.008	-0.01	0.252**	0.247**
Ouvrier	0.002	0.002	0.018	0.018	0.064	0.069
Retraité	0.014	0.014	0.047	0.048	-0.017	-0.021
<i>Statut d'exonération</i>	<i>ref= non</i>					
Motif ALD	0.055***	0.055***	0.092***	0.092***	0.019	0.024
CMU ou AME	-0.129***	-0.129***	0.096	0.096	-0.087	-0.085
Autre	0.073*	0.073*	0.057	0.055	0.14	0.124
<i>Contexte</i>	<i>ref= visite</i>					
Cabinet	-0.043	0.042	-0.038	-0.038	-0.077	-0.076
<i>Nouveau patient</i>	<i>ref= non</i>					
Oui	0.197***	0.197***	0.254***	0.252***	0.152***	0.158***
<i>Année</i>	<i>ref= 2012</i>					
Année 2011	-0.002	-0.002	0.007	0.007	-0.026	-0.025
Variables de niveau 2						
<i>Signature CAPI</i>	<i>ref= non</i>					
Oui	-0.093	-0.043	-0.07	-0.026	-0.035	0.004
Autres variables médecin						
<i>Age</i>	<i>ref = + 55 years</i>					
30-44 ans	-0.079	-0.107	-0.042	-0.057	0.12	0.084
45-54 ans	-0.088	-0.063	-0.076	-0.056	-0.065	-0.041
<i>Sexe</i>	<i>ref= femme</i>					
Homme	-0.06	-0.058	-0.057	-0.049	-0.057	-0.062
<i>Region</i>	<i>ref = ouest</i>					
Centre	-0.005	-0.087	-0.086	-0.152	0.084	0.019
Est	0.19*	0.124	0.03	-0.019	-0.071	-0.126
Ile de France	0.102	0.02	-0.019	-0.087	0.027	-0.047
Nord	0.158	0.058	0.035	-0.045	-0.059	-0.156
Sud	0.068	0.059	-0.047	-0.057	0.121	0.111
<i>Secteur</i>	<i>ref = secteur 1</i>					
Secteur 2	-0.078	0.013	0.014	0.103	0.024	0.123
<i>Milieu d'exercice</i>	<i>ref= urbain</i>					
Rural	-0.001	-0.073	-0.122	-0.179*	0.019	-0.041
Semi-rural	0.035	-0.015	0.047	0.011	0.03	-0.016
<i>Mode d'exercice</i>	<i>ref= solo</i>					
Groupe de médecins	-0.009	-0.011	-0.073	-0.072	0.069	0.056
Groupe pluridisciplinaire	0.044	0.078	-0.013	0.012	0.081	0.107

Les p-values sont indiquées par des étoiles. Niveau de significativité: ***=1% ; **=5% ; *=10%.

Tableau 3. Modèles multivariés de la durée de consultation (log) selon le motif de consultation (suite)

	Diabète		Hypertension		Vaccination anti-grippale	
	<i>Modèle 3a</i>	<i>Modèle 3b</i>	<i>Modèle 3a</i>	<i>Modèle 3b</i>	<i>Modèle 3a</i>	<i>Modèle 3b</i>
<i>Nbr consultations en 2011</i>			<i>ref = + 6000</i>			
Moins de 4000		0.173*		0.151*		0.133
De 4000 à 5000		0.199**		0.192**		0.183**
De 5000 à 6000		-0.012		0.037		-0.046
Variables de patientèle						
Proportion de femmes	0.008	0.007	0.008	0.008	0.006	0.005
Proportion de 0-26 ans	0.001	0.004	-0.001	0.001	-0.007	-0.005
Proportion de +70 ans	0	0.001	0	0.001	-0.003	-0.002
Proportion d'agriculteurs	0.022	0.029*	0.004	0.008	0.003	0.009
Proportion de cadres	0.015*	0.012	0.009	0.007	0.014**	0.011*
Proportion d'ouvriers	-0.01	-0.007	-0.008	-0.005	-0.021**	-0.019**
Proportion de nouveaux patients	0	-0.002	-0.003	-0.004	-0.002	-0.003
Variance des effets aléatoires						
Niveau 1 (ε _{ij})	0.1406***	0.1406***	0.1552***	0.1552***	0.2313***	0.2313***
Niveau 2 (u _{0j})	0.0694***	0.0646***	0.0698***	0.0677***	0.0407***	0.0338***

1.6. Discussion

L'objectif de ce travail était de comprendre l'impact du dispositif de paiement à la performance introduit au travers du CAPI sur la durée de consultation des médecins généralistes français. Une méthodologie en deux étapes a été utilisée. A partir de plus de 20 000 observations de consultations effectuées par 128 médecins généralistes (MGs) et retranscrites entre 2011 et 2012, trois modèles emboîtés ont été estimés par une approche multi-niveaux. Dans ces modèles, nous avons ajouté progressivement les caractéristiques des patients, les caractéristiques des médecins et les caractéristiques de la patientèle. Le modèle le plus adéquat a ensuite été estimé sur trois principaux motifs de consultations faisant l'objet d'un indicateur CAPI : le diabète, l'hypertension, et la vaccination antigrippale.

La principale difficulté de cette étude provenait de l'hypothèse de phénomènes d'auto-sélection associés à la signature du contrat, c'est-à-dire le fait que les signataires du CAPI aient des caractéristiques différentes des non signataires, pouvant biaiser les estimations de l'impact du CAPI sur la durée de consultation. Les résultats des tests d'indépendance n'ont pas révélé de différence significative entre signataires et non signataires en terme d'âge, sexe, localisation géographique ou encore pratique en groupe. En revanche, les signataires fournissaient un plus grand nombre de consultations en 2011 que les non signataires, mais il n'était pas possible de conclure sur le sens de la relation, le nombre d'actes n'ayant pas été observé avant la signature du contrat.

D'un côté, nous pourrions faire l'hypothèse (H1) que les médecins signataires du CAPI étaient plus productifs en termes d'activité médicale (hypothèse d'auto-sélection). Sous cette hypothèse, les résultats de l'étude révèlent un arbitrage quantité/qualité différent entre signataires et non signataires du contrat, les signataires privilégiant plutôt la quantité avec un nombre de consultations plus élevé, et les non signataires, la qualité, avec une durée moyenne de consultation plus élevée. Sous cette hypothèse, le CAPI n'aurait donc pas eu d'effet incitatif sur l'augmentation de la qualité des soins telle que mesurée par la durée de consultation.

D'un autre côté, il est possible que le CAPI ait engendré une augmentation de l'activité médicale telle que mesurée par le nombre de consultations (H2). Comme la durée de consultation est reliée de façon négative au nombre de consultations, cela signifierait que le CAPI ait engendré indirectement une diminution de la durée de consultation. En termes de politiques publiques, ce résultat doit être jugé au regard des objectifs du régulateur : l'objectif est-il d'augmenter la quantité de soins pour faire face à des problèmes d'offre et

potentiellement d'accès aux soins ? Dans ce cas, sous l'hypothèse H2, les P4P pourraient être un outil intéressant. L'objectif est-il d'améliorer la qualité des soins ? Dans ce cas, sous l'hypothèse H2, les P4P auraient un effet délétère sur la qualité. Toutefois, l'hypothèse H2 semble peu vraisemblable pour différentes raisons. La raison principale est que le CAPI ne contient pas en soi d'incitatifs à augmenter le nombre de consultations car le dispositif ne comptabilise pas le nombre de sessions dédiées à un problème de santé. Il est vrai que les médecins pourraient « recruter » davantage de patients pour augmenter le montant des bonus mais, à notre connaissance, il n'existe pas de travaux empiriques ayant montré cet effet. En outre, les montants associés au CAPI (seulement 7% du chiffre d'affaire moyen pour une amélioration de l'ensemble des indicateurs) sont faibles comparé aux montants proposés par exemple dans le cadre du QOF en Angleterre, ces montants s'élevant à environ 26% du chiffre d'affaire moyen des MGs anglais (Bras et Duhamel 2008).

Le résultat principal de cette étude est que le CAPI ne semble pas avoir eu d'effet significatif sur la durée de consultation, après ajustement sur l'ensemble des facteurs de confusion possibles. Les résultats obtenus sur les trois motifs de consultation spécifiques ont permis de valider cette conclusion. Néanmoins, il ne s'agit pas de conclure qu'un P4P en général n'a pas de conséquence sur la qualité des soins approchée par la durée de consultation. En effet, la rémunération additionnelle maximale du CAPI ne correspond finalement qu'à 7% du chiffre d'affaire des médecins ; ces incitatifs, sans être négligeables, ne constituent probablement pas des incitatifs suffisamment élevés pour engendrer une modification significative de la pratique médicale.

Limites

Une première limite de l'étude concerne le nombre limité de variables médecin pouvant expliquer la signature du CAPI et ainsi contrôler des phénomènes d'auto-sélection. Par ailleurs, même si nous avons ajusté par un certain nombre de variables observables, il n'est pas possible d'exclure que certaines variables inobservables telles que le style de pratique, les préférences pour le loisir ou encore l'éthique, puissent influencer à la fois la signature du CAPI et la durée de consultation. Une variable potentiellement cachée dans la relation entre caractéristiques des médecins, signature du CAPI et durée de consultation est la motivation intrinsèque à fournir des soins de qualité. En effet, il a été montré que cette dernière a un impact positif sur le temps dédié à une activité (Deci 1972; Deci and Ryan 2000). Par ailleurs, un travail complémentaire à cette étude a cherché à comprendre les

déterminants de la non signature du CAPI et a révélé le poids prépondérant des considérations éthiques associées au CAPI (risques de conflits d'intérêt, risque d'exclusion des patients les plus précaires, manquement à l'éthique professionnelle) comme déterminants de la non signature du contrat (Saint-Lary et al 2013). En ce sens, une analyse précise des motivations et des préférences des médecins associés à une pratique de qualité semble constituer un prolongement nécessaire à ce travail.

Une seconde limite de l'étude concerne la taille de l'échantillon. L'effectif de 128 médecins, bien que raisonnable, peut générer des problèmes de puissance dans l'estimation des paramètres au niveau médecin. Une autre limite concerne le mode de recrutement des médecins généralistes, qui sont tous maîtres de stages. Ainsi, par rapport à la population des MGs français, les médecins en groupe étaient légèrement sur-représentés (62 % vs 54 % (Gautier 2011)) tout comme les médecins signataires du CAPI (43% vs 38%). Une quatrième limite est liée au fait que pendant les deux tiers de la période d'étude (de Janvier à Avril 2012) le CAPI a été généralisé à l'ensemble des MGs avec la ROSP. De fait, les médecins non signataires du CAPI ont été exposés au nouveau dispositif de paiement à la performance. En revanche, une analyse complémentaire a été effectuée en incluant une indicatrice temporelle (avant/après 2012) dans le modèle, mais n'a pas révélé d'impact significatif de la période considérée sur la durée de la consultation. Ainsi, l'implémentation de la ROSP ne semble pas avoir modifié la durée de la consultation. Enfin, comme aucun identifiant patient n'a été recueilli, il n'était pas possible de prendre en compte les corrélations entre les consultations d'un même patient.

Conclusion

En France, le CAPI ne semble pas avoir modifié la durée de consultation des médecins généralistes. Ce résultat apporte un message d'avertissement quant à la généralisation des paiements à la performance en France et leur capacité à générer des soins de meilleure qualité. Comme la durée de consultation ne constitue pas une mesure unique de la qualité des soins, les recherches futures devraient examiner la question en tenant compte d'autres mesures de la qualité (contenu de la consultation, suivi des patients, satisfaction des patients) et comparer les résultats avec ceux obtenus dans cette étude. Ainsi, l'objectif de la deuxième section du chapitre est d'analyser l'impact du CAPI sur des indicateurs ciblés par le contrat, en particulier le dépistage des cancers.

2. Impact des incitatifs économiques sur le dépistage des cancers

2.1. Introduction

Le cancer est une des premières causes de mortalité dans le monde (Organization 2008). Le cancer du sein et du côlon sont parmi les cancers les plus diagnostiqués (Ferlay et al 2010). Des études publiées dès le milieu des années 1990 ont montré que que les pratiques de prévention et de dépistage des cancers jouent un rôle important dans la diminution de la prévalence, de la morbidité et de la mortalité de ces cancers (Kerlikowske et al 1995; Quinn et al 1999; Walter et al 2005). Un rapport récent de l'Institut National contre le Cancer⁹ (INCa) réaffirme ce constat en mettant en avant le rôle essentiel des médecins généralistes (MGs) dans l'offre ciblée de dépistage des cancers du sein, du col de l'utérus, et du côlon. En effet, le rôle du MG dans le dépistage est important puisqu'il est en général le premier acteur à être contacté pour toute forme de soins non urgents. Il est donc idéalement placé pour éduquer et informer le patient sur l'importance du dépistage (Carrieri, 2009). Par ailleurs, plusieurs auteurs soutiennent la nécessité d'un rôle renforcé du médecin généraliste en tant que coordinateur des différents dépistages (Carcaise-Edinboro and Bradley 2008; Fon Sing et al 2013). En effet, le médecin généraliste est l'un des acteurs les plus pertinents dans la prescription de dépistage du cancer colorectal (Federici et al 2006) et joue un rôle important (consultatif) dans le dépistage des cancers du sein et du col de l'utérus (Jensen et al 2012).

Par ailleurs, en 2009, un paiement à la performance - le CAPI- est introduit pour, entre autres objectifs, favoriser l'offre de dépistage du cancer du sein. Pour cet objectif spécifique, le taux à atteindre pour percevoir le « bonus » correspond à 80% de patientes âgées entre 50 à 74 ans ayant participé au dépistage du cancer du sein (individuel ou organisé) dans les deux dernières années, patientes pour lesquelles le MG est le « médecin traitant »¹⁰. La ROSP, mise

⁹ « La situation du cancer en France en 2011 », INCa.

¹⁰ Le dispositif du médecin traitant a été instauré en 2005 à la suite de la réforme de l'Assurance Maladie d'août 2004. Chaque assuré a la possibilité de désigner un médecin traitant, qu'il soit généraliste ou non, son rôle étant de tenir à jour le dossier médical du patient, de coordonner le parcours de soins et de centraliser les avis des autres soignants. En outre, ce dispositif s'est accompagné d'incitations financières à partir de 2006, de façon à mieux réguler l'accès aux spécialités : les consultations sont ainsi mieux remboursées lorsqu'elles ont été prescrites par le MG traitant. Ainsi, en 2011, 90% des assurés ont choisis un médecin traitant et parmi eux, 95% ont désigné médecin généraliste traitant un (rapport annuel de la cour des comptes, Février 2013).

en place en 2012 pour tous les MGs compte un plus grand nombre d'objectifs et intègre notamment le dépistage du cancer du col de l'utérus : pour ce dépistage, le MG doit atteindre également un taux de 80% de ses patientes âgées entre 25 à 65 ans ayant réalisé un frottis au cours des 3 dernières années. Pour ces deux cancers, 35 points sont consacrés à l'amélioration des indicateurs. Sachant qu'un point correspond à 7€, les MGs peuvent obtenir un montant maximal de 245€. Pour le dépistage du cancer colorectal, en revanche, aucun dispositif incitatif de type P4P n'a été mise en œuvre à l'heure actuelle.

Cette généralisation d'un paiement à la performance (P4P) en France a été décidée par le financeur publique alors même que les études internationales n'ont pas montré d'effets significatifs de ces incitatifs sur l'efficacité des pratiques (Petersen et al 2006; Mannion and Davies 2008; Eijkenaar et al 2013), notamment dans le cadre de l'offre de prévention (Grady et al 1997b; Town et al 2005b; Kiran et al 2014; Li et al 2014). Par exemple, Li et al (2014) montrent un impact très modéré voir non significatif des P4P instaurés à destination des MGs pour améliorer le recours au dépistage des cancers du col de l'utérus, du sein et du côlon. Pourtant, les auteurs ont pu bénéficier d'une expérience naturelle dans la région de l'Ontario et utilisent la méthode de différence de différences qui permet de contrôler des effets (observables et inobservables) de sélection des médecins exposés au paiement à la performance. Ainsi malgré un montant relativement élevé en faveur des dépistages : \$2200 (soit environ 1700€) pour l'atteinte d'un taux de 80% de patients dépistés, les résultats semblent peu probants.

Un élément qui renforce la difficulté à évaluer l'impact de ces schémas de P4P est qu'ils se cumulent avec d'autres dispositifs incitatifs mobilisés pour encourager le recours au dépistage des cancers. Entre autres, un programme de dépistage organisé a été mis en place depuis 2004 pour le cancer du sein et depuis 2009 pour le cancer colorectal, alors que des discussions sont en cours pour la mise en place d'un tel programme pour le dépistage du cancer du col de l'utérus. Ainsi, malgré son rôle essentiel d'éducation et d'information (Kendel, 1994), notamment chez les femmes (Lurie et al, 1993), le MG n'est pas le seul acteur du dépistage. Compte tenu de l'organisation de l'offre en France, il est en « concurrence » avec d'autres spécialistes tels que les gynécologues (accès direct) mais aussi avec le dépistage organisé puisque sur invitation du financeur, les femmes peuvent effectuer le dépistage sans consulter préalablement leur médecin traitant. Par ailleurs, il semble difficile de pouvoir évaluer l'impact d'un mécanisme tel que le CAPI ou la ROSP sur le dépistage des cancers sans étudier préalablement les disparités individuelles de recours au dépistage. Ces

disparités peuvent être liées à des facteurs socioéconomiques incluant notamment l'âge, le sexe, le niveau d'éducation ou encore le fait de bénéficier d'une couverture complémentaire, ainsi qu'à des facteurs psychologiques tels que l'aversion ou la perception du risque, la préférence pour le présent, les antécédents familiaux de cancer, etc. Ces déterminants ne sont pas exhaustifs et des recherches récentes sont menées pour mettre en évidence des freins éventuels tels que le stress et l'angoisse face à l'information.

Ainsi, avant d'étudier l'impact des P4P sur le recours au dépistage des cancers, nous nous intéressons dans un premier temps à la demande de dépistage et plus spécifiquement aux freins associés à la réalisation d'un dépistage du cancer du sein, du col de l'utérus, et du côlon (section 2). Dans la section 3, nous présentons les données et méthodes utilisées pour analyser l'impact du CAPI sur le recours au dépistage du cancer du sein.

2.2. Analyse des freins au dépistage des cancers¹¹

2.2.1. Contexte

Malgré la reconnaissance, par les autorités de santé françaises et internationales, de la nécessité d'un dépistage de masse des cancers du sein, du col de l'utérus, et du côlon, les analyses récentes mettent en évidence de nombreuses disparités dans le recours au dépistage de ces cancers. Tout d'abord, il semble exister des inégalités de recours liées notamment au niveau de revenu, d'éducation, d'accès à l'assurance complémentaire ou encore liées au statut marital (Lantz et al 1997; Chattopadhyay et al 2005; Arrossi et al 2008; Couture et al 2008; Lofters et al 2010). En ce qui concerne les cancers féminins, différents travaux ont montré que les femmes pauvres, non mariées, avec un niveau d'éducation faible et un accès plus difficile au système de santé, avaient moins de chance d'être dépistées (Arrossi et al 2008; Martín-López et al). D'autres travaux ont montré le rôle de l'âge, du fait de vivre en couple et/ou d'avoir des enfants, et de consulter plus régulièrement son médecin généraliste sur le dépistage du cancer colorectal (Beck and Gautier 2010). Il a aussi été montré que l'illettrisme et le dégoût pour le test constituaient des barrières au dépistage du cancer colorectal (Kobayashi et al 2014). Enfin, la littérature en psychologie et notamment en psychologie de la santé a mis en évidence le rôle des comportements et des perceptions du risque sur le recours

¹¹ Cette section a fait l'objet d'un article coécrit avec Carine Franc accepté dans la revue *BMC Health Services Research* sous le titre: « Obstacles to the uptake of breast, cervical, and colorectal cancer screenings. What remains to be achieved by French national programmes ? ».

au dépistage (Eaker et al 2001; Katapodi et al 2004; Mills et al 2008), et en particulier le lien entre l'histoire familiale de cancer et le recours à la mammographie (Kash et al 1992).

Face à ces constats, des dispositifs ont été mis en place, et la lutte contre les inégalités d'accès et de recours aux dépistages des cancers a été définie comme une priorité du 2^e plan cancer 2009-2013. Des programmes de dépistages organisés ont été instaurés en France en 2004 pour le cancer du sein et en 2009 pour le cancer colorectal. Le dépistage organisé du cancer du sein cible les femmes âgées entre 50 et 74 ans en les invitant à réaliser (gratuitement) une mammographie bilatérale des seins (avec 2^e lecture) tous les deux ans dans un centre radiologique participant au programme. Le dépistage organisé du cancer colorectal cible les hommes et femmes âgés entre 50 et 74 ans ; il consiste à envoyer par courrier un test hemocult (test par recherche de sang occulte dans les selles) ainsi qu'une notice explicative sur les modalités de réalisation du test et le dépôt dans un laboratoire biologique pour analyse gratuite. Le dépistage du cancer du col de l'utérus ne fait pas encore l'objet d'un dépistage organisé. D'après les recommandations de la HAS, il concerne les femmes âgées entre 25 et 65 ans qui doivent réaliser un frottis cervico-utérin tous les trois ans.

De nombreuses études se sont intéressées aux déterminants du recours au dépistage des cancers en coupe transversale (Janz et al 2007; Arrossi et al 2008; Couture et al 2008; Rigal et al 2011). Toutefois, peu d'études se sont intéressées à l'évolution de ces déterminants dans le temps, évolution associée par exemple à la mise en place de dépistages organisés ou de campagnes de dépistage de masse. Pourtant, dans le contexte actuel de renforcement des politiques de prévention via l'extension des dépistages organisés et la médiatisation des campagnes de dépistages, d'instaurations d'incitatifs économiques - côté demande via la gratuité des dépistages comme côté offre via les paiements à la performance- on pourrait s'attendre à une augmentation des taux de recours aux dépistages des cancers du sein et du côlon en particulier et à une diminution des inégalités socioéconomiques. Afin de tester cette hypothèse, nous modélisons les déterminants du non recours au dépistage des cancers du col de l'utérus, du sein et du côlon en France et leur évolution dans le temps, c'est-à-dire entre 2006 et 2010. Notons que dans ce travail, le terme « inégalités » de recours fait référence aux disparités observées de recours au dépistage qui sont supposées provenir de barrières d'accès aux soins ou liées à l'organisation du système de santé. Le terme « différence » fait référence aux disparités de recours observées dont il n'est pas possible de déterminer si elles sont liées aux choix des individus ou à l'équité du système de santé.

2.2.2. L'enquête ESPS

L'enquête santé protection sociale (ESPS) est une enquête par sondage en population générale réalisée depuis 1988 par l'Institut de Recherche et de documentation en économie de la santé (IRDES). Annuelle de 1988 à 1997, elle est devenue bisannuelle depuis 1998. Tous les bénéficiaires de l'assurance maladie appartenant à l'échantillon ESPS ainsi que les membres de leurs ménages sont sollicités tous les quatre ans (par vague). Jusqu'en 2008, la base de sondage de l'enquête ESPS était l'échantillon permanent des assurés sociaux (EPAS) ; depuis 2010 celui-ci a été remplacé par l'échantillon généraliste des bénéficiaires (EGB). Ces deux sources sont réputées pour être représentatives de la population générale. Ainsi, les vagues 2006, 2008 et 2010 sont constituées de différentes sous-populations et les taux de réponses sont respectivement de 63%, 65% et 66%. Les bénéficiaires échantillonnés ainsi que l'ensemble des membres de leur ménage sont enquêtés. Les données déclaratives renseignent donc sur un ensemble d'informations concernant le ménage : revenus, type de ménage, région d'habitation, etc. et concernant les différents membres du ménage: caractéristiques socio démographiques, état de santé, utilisation des services de santé, prévention et comportements à risque.

2.2.3. Variables de recours au dépistage des cancers

Nous avons considéré trois variables dépendantes : le recours au dépistage du cancer du col de l'utérus, du sein et du côlon suivant les recommandations des autorités de santé françaises.

- La variable de dépistage du cancer du col de l'utérus est codée 1 si la femme âgée entre 25 et 65 ans déclare avoir effectué un frottis de dépistage il y a moins de trois ans, 0 sinon.
- La variable de dépistage du cancer du sein est codée 1 si la femme âgée entre 50 et 74 ans déclare avoir effectué une mammographie dans les deux ans, 0 sinon.
- La variable de dépistage du cancer colorectal est codée 1 pour les hommes et les femmes âgés entre 50 et 74 ans déclarant avoir effectué un test hemocult dans les deux ans, 0 sinon.

2.2.4. Déterminants du non recours au dépistage des cancers

Les déterminants du recours au dépistage des cancers sont regroupés en 3 catégories : (1) les caractéristiques socio-économiques, (2) l'état de santé et le recours aux soins, (3) les comportements à risque et de prévention/dépistage.

Pour la catégorie 1 nous incluons l'âge, la zone d'aménagement du territoire (8 zones : Ile de France, Bassin Parisien, Nord, Est, Ouest, Sud-Ouest, Centre, Méditerranée), la classe sociale (en 8 catégories socio-professionnelles suivant la catégorisation de l'INSEE), le statut marital (en couple/célibataire), le fait de posséder une couverture maladie complémentaire privée ou la CMU-C (trois catégories : oui/CMU-C/non)¹².

Pour la catégorie 2 nous incluons deux variables d'état de santé : une variable dichotomique codée 1 si l'individu déclare une maladie chronique, 0 sinon, et une variable d'état de santé auto-déclarée codée en 4 catégories à partir d'une échelle de Likert à dix niveaux (1-4 = « mauvais » / 5-6 = « moyen » / 7-8 = « bon » / « 9-10 » = « très bon-excellent »). L'utilisation des services de santé est mesurée au travers de deux variables codées en 2 catégories: le nombre de consultations chez le médecin généraliste au cours des 12 derniers mois (1= supérieur à la médiane de l'échantillon / 0= inférieur à la médiane), et le nombre de consultations chez le spécialiste au cours des 12 derniers mois (1= supérieur à la médiane de l'échantillon / 0= inférieur à la médiane).

Dans la troisième catégorie, les comportements à risque sont mesurés au travers du niveau de consommation d'alcool et de tabac. Quatre classes sont définies pour décrire les comportements à risque associés à la consommation de tabac : les non-fumeurs, les ex-fumeurs, les fumeurs « légers » (moins de 10 cigarettes par jour), les fumeurs « à risque » (strictement plus de 10 cigarettes par jour). Pour le profil d'alcoolisme, une variable en quatre catégories progressives est construite en suivant la typologie établie par Com-Ruelle et al sur la base de l'AUDIT-C¹³ (Bush et al 1998; Com-Ruelle et al 2005). Ainsi, une volumétrie hebdomadaire d'alcool consommé est calculée avant d'être comparée aux seuils recommandés par l'OMS. La construction de la typologie permet alors de définir quatre catégories de consommateurs d'alcool : les personnes déclarant ne jamais consommer

¹² En France, l'assurance maladie complémentaire privée est souscrite afin de couvrir les dépenses non prises en charges par l'Assurance Maladie (restes à charge, dépassements d'honoraires). En dessous d'un seuil de revenu, les individus peuvent bénéficier de l'Assurance Maladie Universelle Complémentaire (la CMU-C).

¹³ L'AUDIT-C est version courte de l'Alcohol Use Disorders Identification Test (AUDIT) développé et recommandé par l'OMS.

d'alcool, les consommateurs sans risque, les consommateurs à risque ponctuel et les consommateurs à risque chronique/avec dépendance. L'algorithme mis en œuvre par l'IRDES permettant de classer les individus est présenté dans la figure A.2 (Annexe A).

Concernant les comportements de prévention/dépistage, nous incluons les autres variables indicatrices de dépistage pour lesquelles les individus sont éligibles. Ainsi, pour le modèle concernant le dépistage du cancer du sein, nous incluons comme prédicteur le fait pour la femme d'avoir ou non effectué un test hémocult dans les deux ans, et dans le modèle concernant le dépistage du cancer colorectal (pour les femmes uniquement), nous incluons comme prédicteur le fait pour les femmes d'avoir ou non effectué une mammographie dans les deux ans¹⁴.

2.2.5. Méthodologie d'analyse statistique

Dans un premier temps, nous effectuons des statistiques descriptives afin d'étudier l'évolution du recours au dépistage des trois cancers entre 2006 et 2010 et d'examiner si la distribution des variables indépendantes est similaire d'une année d'enquête à l'autre.

Dans un deuxième temps, les données des trois vagues d'enquêtes sont regroupées et un modèle logistique multivarié sur données « poolées » est estimé pour déterminer l'association entre le non recours au dépistage des cancers et les variables indépendantes. Comme tous les dépistages concernent les femmes et seulement un (le côlon) concerne les hommes, des analyses séparées homme femmes sont menées pour ce cancer. Par ailleurs, deux modèles distincts sont estimés pour le dépistage du cancer du sein (respectivement, le cancer colorectal) : le premier modèle n'inclue pas la réalisation d'un test hémocult (respectivement, d'une mammographie) comme prédicteur, contrairement au second modèle. En effet, comme ces variables sont potentiellement endogènes, leur inclusion est suspectée d'introduire des biais dans les modèles. Par ailleurs, des catégories de non réponses sont créées pour chaque variable indépendante et incluses dans les modèles, mais nous ne reportons par les coefficients associés dans les tableaux par soucis de parcimonie.

Dans un troisième temps enfin, nous analysons les évolutions des déterminants du non recours au dépistage dans le temps en intégrant des interactions entre les variables d'intérêt (i.e. l'âge, la classe sociale, l'assurance complémentaire, les comportements à risque) et des

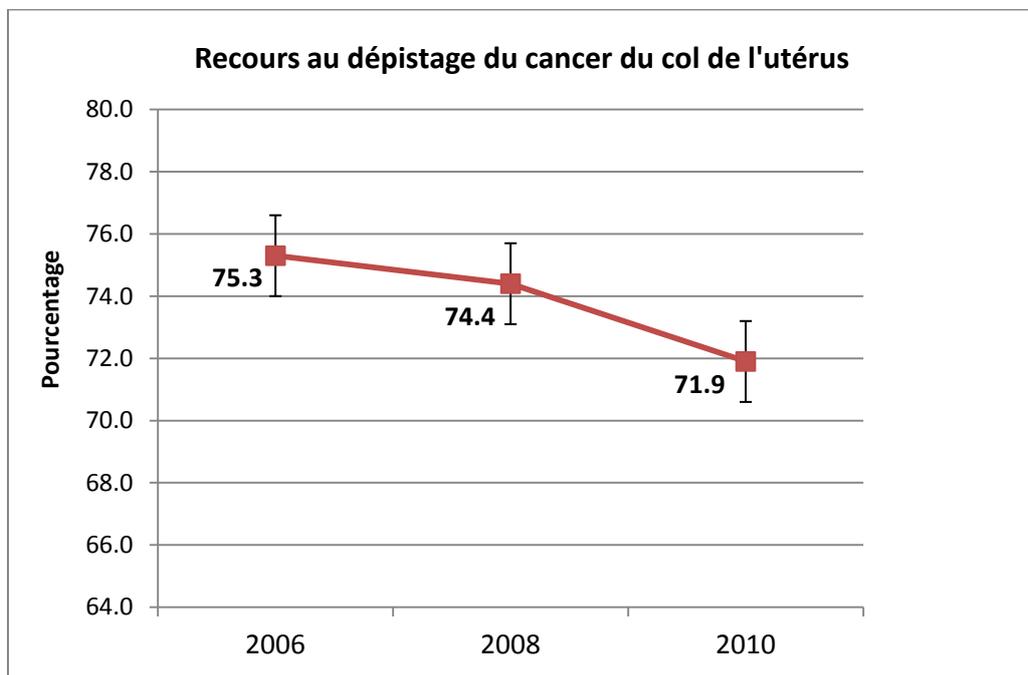
¹⁴ Nous n'avons pas inclus le dépistage du cancer du sein ou du côlon comme prédicteur du dépistage du cancer du col de l'utérus car ces dépistages ne concernent pas les mêmes tranches d'âge.

indicatrices années (i.e. 2008 vs. 2006 et 2010 vs. 2006) pour les modèles concernant le dépistage des cancers féminins, et 2010 vs. 2008 pour le dépistage du cancer colorectal¹⁵). Dans ce modèle final, nous n'incluons que les interactions significatives au seuil de 5%.

2.2.6. Evolution du recours au dépistage des cancers et de leurs déterminants

Les résultats concernant l'évolution du recours au dépistage des cancers du col de l'utérus, du sein et du côlon entre 2006 et 2010 sont présentés dans les figures 2, 3 et 4.

Figure 2. Evolution du recours au dépistage du cancer du col de l'utérus entre 2006 et 2010



Alors que 75,3% (95% IC = [74,1%-76,5%]) des femmes déclarent avoir effectué un frottis de dépistage dans les trois ans en 2006, ce taux décroît significativement pour atteindre 71,9% [70,5%-73,3%] en 2010. Le taux de recours au dépistage du cancer du sein décroît aussi entre 2006 et 2010, passant de 77,6% [75,8%-79,3%] en 2006 à 74,0% [72,1%-75,9%] en 2010, mais cette baisse n'est significative qu'au seuil de 10%.

Sur l'ensemble, le taux de recours au dépistage du cancer colorectal augmente significativement entre 2006 et 2010, de 18,2% [17,0%-19,4%] en 2006 à 38,9% [37,4%-40,5%] en 2010. Comme le montre la Figure 4, le taux de croissance du recours est

¹⁵ En effet, pour ce cancer en particulier, nous nous intéressons aux évolutions des différences/inégalités de recours suite à la mise en place du dépistage organisé en 2009.

légèrement plus fort chez les femmes, passant de 16,9% (vs. 19,6% pour les hommes) en 2006 à 40,0% (vs. 38,0% pour les hommes).

Figure 3. Evolution du recours au dépistage du cancer du col de l'utérus entre 2006 et 2010

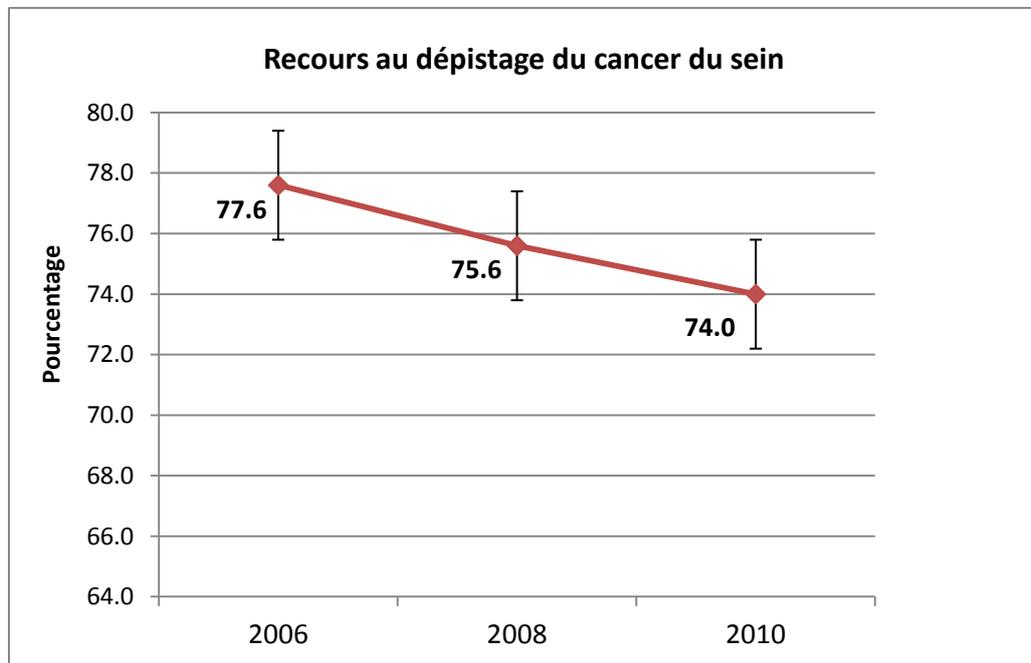
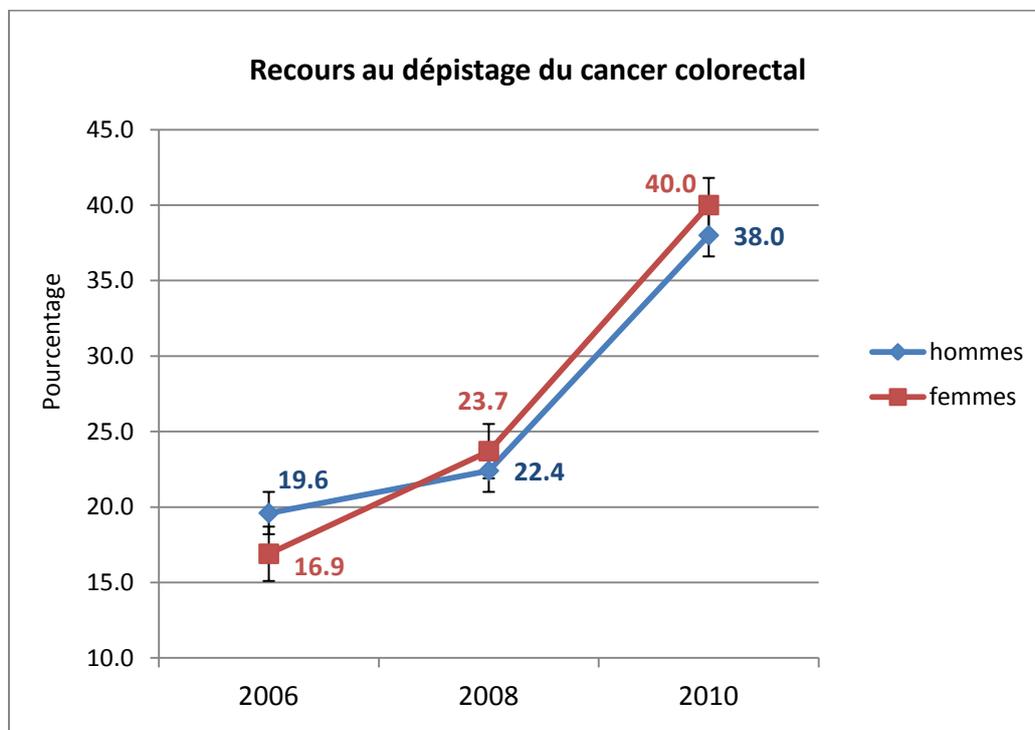


Figure 4. Evolution du recours au dépistage du cancer colorectal entre 2006 et 2010



Le tableau 4 présente la distribution des variables indépendantes en fonction de l'année d'enquête. Pour la plupart des variables, la proportion des individus dans chacune des catégories des variables ne varie pas ou peu entre les années d'enquête. Les principales exceptions sont le fait de posséder la CMUC (9,7% en 2010 *vs.* 6,5% en 2006), de déclarer une maladie chronique (58,5% en 2010 *vs.* 68,8% en 2006), d'avoir vu un spécialiste au moins une fois dans l'année (47,1% en 2010 *vs.* 42,0% en 2008) et de déclarer ne jamais consommer d'alcool (19,3% en 2008 *vs.* 24,2% en 2010).

Tableau 4. Distribution des variables indépendantes en fonction de l'année d'enquête

	2006		2008		2010	
	N	%	N	%	N	%
Caractéristiques socio-économiques						
Age						
[25-49]	5876	56.1	5205	53.2	4825	54.0
[50-65]	3404	32.5	3372	34.5	3092	34.6
[66-74]	1203	11.5	1209	12.4	1021	11.4
Localisation géographique						
Ile de France (hors Paris)	1574	15.0	1350	13.8	1256	14.1
Bassin Parisien	1870	17.8	1895	19.4	1614	18.1
Nord	741	7.1	709	7.3	657	7.4
Est	958	9.1	851	8.7	795	8.9
Ouest	1500	14.3	1443	14.8	1511	16.9
Sud Ouest	1277	12.2	1134	11.6	1030	11.6
Centre	1468	14.0	1279	13.1	1143	12.8
Méditerranée	1095	10.5	1125	11.5	914	10.3
Catégorie socio-professionnelle						
Agriculteurs exploitants	353	3.4	302	3.1	259	2.9
Artisans, commerçants et chefs d'entreprise	581	5.5	595	6.1	521	5.8
Cadres et prof. intellectuelles	1359	13.0	1291	13.2	1201	13.4
Professions intermédiaires	2159	20.6	2097	21.4	1666	18.6
Employés	2904	27.7	2761	28.2	2561	28.7
Ouvriers qualifiés	1766	16.9	1572	16.1	1491	16.7
Ouvriers non qualifiés	1046	10.0	939	9.6	880	9.9
Inactifs	301	2.9	215	2.2	316	3.5
Statut marital						
En couple	8425	80.4	7805	79.8	7142	79.9
Célibataire	1115	10.6	1049	10.7	978	10.9
Non réponse	943	9.0	932	9.5	818	9.2
Assurance maladie complémentaire						
Oui (privée)	9150	87.3	8628	88.2	7721	86.4
CMU-C	682	6.5	674	6.9	868	9.7
Non	618	5.9	454	4.6	300	3.4
Etat de santé et recours aux soins						
Maladie chronique déclarée						
Oui	2610	24.9	2561	26.2	3196	35.8
Non	7217	68.8	6574	67.2	5227	58.5
Ne sais pas	656	6.3	651	6.7	515	5.8
Etat de santé auto-déclaré						
[0-4] : mauvais	365	3.5	403	4.1	386	4.3
[5-6] : moyen	1604	15.3	1709	17.5	1362	15.2
[7-8] : bon	4870	46.5	4699	48.0	4090	45.8
[9-10] : excellent	3225	30.8	2555	26.1	2688	30.1
Non réponse	419	4.0	420	4.3	412	4.6
Nb. de visites chez le généraliste (12 derniers mois)						
Moins de 2	4884	46.6	4539	46.4	4133	46.2
Plus de 3	4983	47.5	4693	48.0	4208	47.1
Non réponse	616	5.9	554	5.7	597	6.7
Nb. de visites chez le spécialiste (12 derniers mois)						
Aucune	4154	39.6	3690	37.7	3302	36.9
Au moins une	4578	43.7	4112	42.0	4208	47.1
Non réponse	1751	16.7	1984	20.3	1428	16.0

Tableau 4. Distribution des variables indépendantes en fonction de l'année d'enquête (suite)

	2006		2008		2010	
	N	%	N	%	N	%
Comportements à risque						
Consommation de tabac						
Non fumeur	3608	34.4	3434	35.1	2992	33.5
Ex-fumeur	2825	27.0	2562	26.2	2101	23.5
<= 10 cigarettes/jour	1357	12.9	1412	14.4	1495	16.7
> 10 cigarettes/jour	1428	13.6	1241	12.7	1161	13.0
Non réponse	1265	12.1	1137	11.6	1189	13.3
Consommation d'alcool						
Jamais	2234	21.3	1893	19.3	2164	24.2
Consommateur sans risque	4590	43.8	4338	44.3	3676	41.1
Consommateurs à risque ponctuel	2575	24.6	2288	23.4	1997	22.3
Consommateurs à risque chronique	531	5.1	748	7.6	629	7.0
Non réponse	553	5.3	519	5.3	472	5.3

2.2.7. Analyses multivariées du non recours au dépistage des cancers

Dépistage du cancer du col de l'utérus

Les résultats des modèles avec effets principaux et interactions pour le dépistage du cancer du col de l'utérus sont présentés dans le tableau 5 (seulement une interaction est significative au niveau 5%).

Toutes les variables indépendantes expliquent l'absence de recours au dépistage du cancer dans l'analyse multivariée (N= 13 177 femmes). Ainsi, comparées aux femmes plus jeunes (25-34 ans), les femmes âgées entre 45 et 54 ans ou 55-65 ans ont une probabilité plus élevée de ne pas être dépistées ($OR_{45-54} = 1.56 [1.37-1.78]$, $OR_{55-65} = 2.39 [2.08-2.74]$). L'impact négatif de l'âge sur le recours au dépistage du cancer du col semble s'être accentué dans le temps: les femmes âgées entre 45 et 54 ans ou entre 55 et 65 ans ont une probabilité plus forte de ne pas être dépistées en 2010 par rapport à 2006 ($OR_{45-54}=1.54$, $OR_{55-65} = 1.40$). Comparées aux femmes ayant une profession intermédiaire, vivant en couple, et ayant une couverture maladie complémentaire privée, les femmes agricultrices, employées, ouvrières qualifiées ou non qualifiées ou sans emplois, ainsi que les femmes célibataires, n'ayant pas de complémentaire santé ou bénéficiant de la couverture maladie universelle complémentaire (CMU-C), ont une probabilité plus forte de ne pas recourir au dépistage.

De plus, les femmes déclarant une maladie chronique, un état de santé « bon », « moyen » ou « mauvais » et un moindre recours au généraliste ou au spécialiste, ont plus de chance de ne pas avoir été dépistées. Enfin, en ce qui concerne les comportements à risque, la probabilité de recourir au dépistage suit une courbe en U : comparé aux non fumeurs, être un

fumeur « léger » ou un ex-fumeur décroît la probabilité de non recours au dépistage ($OR_{\text{ex-fumeur}} = 0.84 [0.74-0.94]$, $OR_{\leq 10c/j} = 0.81 [0.70-0.92]$) alors qu'être un fumeur « à risque » accroît cette probabilité ($OR_{>10c/j} = 1.36 (1.18-1.56)$). De façon similaire, comparé aux fumeurs occasionnels, ne pas consommer d'alcool ou être un consommateur à risque chronique augmente la probabilité de ne pas recourir au dépistage ($OR_{\text{jamais}} = 1.44 [1.30-1.59]$, $OR_{\text{risque-chronique}} = 1.42 [1.09-1.85]$).

Toutes choses égales par ailleurs, les femmes ont une probabilité plus élevée de ne pas être dépistées en 2010 par rapport à 2006 ($OR_{2010} = 1.15 [1.03-1.27]$).

Dépistage du cancer du sein

Les résultats du modèle à effets principaux pour le recours au dépistage du cancer du sein sont présentés dans le tableau 6 (aucune interaction n'était significative au seuil de 5%). Les déterminants associés au non recours au dépistage du cancer du sein (N= 6 220 femmes) incluent le fait d'être un ouvrier non qualifié ($OR = 1.62 [1.26-2.09]$), d'être sans emploi ($OR = 1.61 [1.14-2.28]$), célibataire ($OR = 1.45 [1.17-1.80]$), ne pas avoir d'assurance maladie complémentaire ($OR = 2.08 [1.57-2.75]$) ou avoir la CMU-C ($OR = 1.79 [1.36-2.37]$). Par ailleurs, les femmes déclarant une maladie chronique, un mauvais état de santé, un nombre plus faible de consultations chez le médecin généraliste et aucune consultation chez le spécialiste dans l'année, ont une probabilité plus élevée de ne pas avoir été dépistées. Ceci est aussi le cas des fumeurs à risque et des individus ne consommant jamais d'alcool.

Toutes choses égales par ailleurs, les femmes ont une probabilité plus faible d'être dépistées en 2010 par rapport à 2006.

Dans le modèle 2 où le recours au dépistage du cancer colorectal est inclu comme prédicteur, l'effet de ce dépistage est significatif : toutes choses égales par ailleurs, ne pas avoir effectué de test hémocult dans les deux dernières années augmente la probabilité de ne pas recourir à la mammographie ($OR = 3.82 [3.18-4.60]$).

Tableau 5. Modèles logistiques multivariés du non recours au dépistage du cancer du col de l'utérus

	Absence de dépistage du cancer du col de l'utérus	Absence de dépistage du cancer du col de l'utérus
	<i>Effets principaux</i>	<i>Interactions</i>
	OR (95% IC)	OR (95% IC)
Caractéristiques socio-économiques		
Age	<i>(ref= 25-34)</i>	
[35-44]	1.05 (0.92-1.20)	1.01 (0.81-1.25)
[45-54]	1.56*** (1.37-1.78)	1.19 (0.97-4.47)
[55-65]	2.39*** (2.08-2.74)	2.05*** (1.65-2.56)
Age * année d'enquête	<i>(ref: age = '25-34'; année = '2006')</i>	
[45-54]*2008	...	1.50*** (1.11-2.03)
[45-54]*2010	...	1.54*** (1.13-2.09)
[55-65]*2010	...	1.40** (1.03-1.91)
Catégorie socio-professionnelle	<i>(ref= prof. intermédiaire)</i>	
Agriculteurs exploitants	1.78*** (1.29-2.45)	1.78*** (1.29-2.45)
Artisans, commerçants et chefs d'entreprise	1.19 (0.91-1.54)	1.18 (0.91-1.54)
Cadres et prof. intellectuelles	0.89 (0.74-1.08)	0.89 (0.74-1.08)
Employés	1.19*** (1.05-1.34)	1.18*** (1.05-1.33)
Ouvriers qualifiés	1.24** (1.01-1.53)	1.24** (1.00-1.52)
Ouvriers non qualifiés	1.64*** (1.38-1.95)	1.64*** (1.38-1.95)
Inactifs	2.31*** (1.85-2.89)	2.29*** (1.84-2.86)
Statut marital	<i>(ref= en couple)</i>	
Célibataire	1.82*** (1.61-2.06)	1.82*** (1.60-2.05)
Assurance maladie complémentaire	<i>(ref= privée)</i>	
CMU-C	1.54*** (1.33-1.78)	1.54*** (1.33-1.78)
Non	2.05*** (1.68-2.51)	2.06*** (1.68-2.52)
État de santé et recours aux soins		
Maladie chronique déclarée	<i>(ref= non)</i>	
Oui	1.28*** (1.15-1.43)	1.28*** (1.14-1.42)
État de santé auto-déclaré	<i>(ref= excellent)</i>	
[0-4] : mauvais	2.04*** (1.62-2.58)	2.04*** (1.61-2.58)
[5-6] : moyen	1.52*** (1.31-1.76)	1.53*** (1.32-1.77)
[7-8] : bon	1.21*** (1.08-1.34)	1.21*** (1.09-1.35)
Nb. de visites chez le généraliste (12 derniers mois)	<i>(ref= plus de 3)</i>	
Moins de 2	1.20*** (1.09-1.33)	1.21*** (1.10-1.33)
Nb. de visites chez le spécialiste (12 derniers mois)	<i>(ref= au moins une)</i>	
Aucune	3.43*** (3.11-3.79)	3.43*** (3.11-3.79)
Comportements à risque		
Consommation de tabac	<i>(ref= non fumeur)</i>	
Ex-fumeur	0.84*** (0.74-0.94)	0.83*** (0.74-0.94)
<= 10 cigarettes/jour	0.81*** (0.70-0.92)	0.80*** (0.70-0.92)
> 10 cigarettes/jour	1.36*** (1.18-1.56)	1.36*** (1.18-1.56)
Consommation d'alcool	<i>(ref= consommateur sans risques)</i>	
Jamais	1.44*** (1.30-1.59)	1.45*** (1.30-1.60)
Consommateurs à risque ponctuel	1.09 (0.95-1.25)	1.09 (0.95-1.25)
Consommateurs à risque chronique	1.42*** (1.09-1.85)	1.43*** (1.10-1.86)
Année d'enquête	<i>(ref= 2006)</i>	
2008	1.02 (0.92-1.13)	0.86 (0.68-1.09)
2010	1.15** (1.03-1.27)	0.90 (0.71-1.15)

Les résultats sont ajustés par la variable de localisation géographique. Les p-values sont indiquées par des étoiles. Niveau de significativité: ***=1% ; **=5% ; *=10%.

Tableau 6. Modèles logistiques multivariés du non recours au dépistage cancer du sein

	Absence de dépistage du cancer du sein
	<i>Effets principaux</i>
	OR (95% IC)
Caractéristiques socio-économiques	
Age	(ref= 50-54)
[55-59]	0.75*** (0.62-0.89)
[60-64]	0.78*** (0.64-0.94)
[65-74]	1.06 (0.89-1.27)
Catégorie socio-professionnelle	(ref= prof. intermédiaire)
Agriculteurs exploitants	1.08 (0.76-1.55)
Artisans, commerçants et chefs d'entreprise	1.09 (0.78-1.51)
Cadres et prof. intellectuelles	1.00 (0.76-1.33)
Employés	1.18 (0.98-1.42)
Ouvriers qualifiés	1.28 (0.96-1.70)
Ouvriers non qualifiés	1.62*** (1.26-2.09)
Inactifs	1.61*** (1.14-2.28)
Statut marital	(ref= en couple)
Célibataire	1.45*** (1.17-1.80)
Assurance maladie complémentaire	(ref= privée)
CMU-C	1.79*** (1.36-2.37)
Non	2.08*** (1.57-2.75)
État de santé et recours aux soins	
Maladie chronique déclarée	(ref= non)
Oui	1.29*** (1.11-1.50)
État de santé auto-déclaré	(ref= excellent)
[0-4] : mauvais	1.87*** (1.38-2.54)
[5-6] : moyen	1.20 (0.96-1.49)
[7-8] : bon	1.06 (0.88-1.28)
Nb. de visites chez le généraliste (12 derniers mois)	(ref= plus de 3)
Moins de 2	1.40*** (1.21-1.62)
Nb. de visites chez le spécialiste (12 derniers mois)	(ref= au moins une)
Aucune	3.00*** (2.59-3.47)
Comportements à risque	
Consommation de tabac	(ref= non fumeur)
Ex-fumeur	0.92 (0.77-1.10)
<= 10 cigarettes/jour	1.08 (0.85-1.36)
> 10 cigarettes/jour	1.84*** (1.44-2.37)
Consommation d'alcool	(ref= consommateur sans risques)
Jamais	1.27*** (1.09-1.48)
Consommateurs à risque ponctuel	1.14 (0.90-1.44)
Consommateurs à risque chronique	1.26 (0.88-1.80)
Année d'enquête	(ref= 2006)
2008	1.16 (1.00-1.35)
2010	1.25*** (1.07-1.46)
Dépistage du cancer colorectal (modèle 2 uniquement)	
	(ref= oui)
Non	3.82*** (3.18-4.60)

Les résultats sont ajustés par la variable de localisation géographique. Les p-values sont indiquées par des étoiles. Niveau de significativité: ***=1% ; **=5% ; *=10%.

Dépistage du cancer colorectal

Les résultats des modèles à effets principaux et avec interactions pour le dépistage du cancer colorectal sont présentés dans le tableau 7 (N= 5 927 hommes et 6229 femmes). Pour les deux sexes, les déterminants associés au non recours sont le fait d'être plus jeune (50-74 ans), être un ouvrier non qualifié, avoir un moindre recours aux services de santé, et être fumeur à risques (>10 cigarettes/jour). Les femmes déclarant un mauvais état de santé et ne consommant jamais d'alcool ont une probabilité plus élevée de ne pas être dépistées (OR= 1.48 [1.07-2.03], OR= 1.22 [1.05-1.42], respectivement) mais ce n'est pas le cas pour les hommes. Au contraire, être un fumeur « léger » augmente la probabilité de ne pas être dépisté chez les hommes (OR= 1.28 [1.03-2.22]) mais non chez les femmes. Une interaction est significative dans le modèle des hommes : comparé au fait de posséder une complémentaire privée, le fait de posséder la CMU-C est associé à une probabilité plus forte de ne pas recourir au dépistage du cancer colorectal en 2010 (OR= 4.31 [1.90-9.76]).

Les résultats du modèle 2 incluant le recours à la mammographie pour les femmes montrent un résultat significatif de cette variable : ne pas avoir recouru à la mammographie dans les deux dernières années est associé à une probabilité plus forte de ne pas recourir au dépistage du cancer colorectal (OR= 3.85 [3.20-4.63]).

Tableau 7. Modèles logistiques multivariés du non recours au dépistage du cancer colorectal

	Absence de dépistage du cancer colorectal <i>Effets principaux</i>	Absence de dépistage du cancer colorectal <i>Interactions</i>	Absence de dépistage du cancer colorectal <i>Effets principaux</i>
	Hommes	Hommes	Femmes
Caractéristiques socio-économiques			
Age		<i>(ref= 50-54)</i>	
[55-59]	0.64*** (0.54-0.77)	0.64*** (0.53-0.77)	0.55*** (0.46-0.66)
[60-64]	0.57*** (0.48-0.69)	0.57*** (0.48-0.69)	0.48*** (0.40-0.58)
[65-74]	0.54*** (0.45-0.65)	0.54*** (0.45-0.64)	0.47*** (0.39-0.56)
Catégorie socio-professionnelle		<i>(ref= prof. intermédiaire)</i>	
Agriculteurs exploitants	1.46** (1.08-1.99)	1.46*** (1.07-1.99)	1.06 (0.76-1.50)
Artisans, commerçants, chefs d'entreprise	1.19 (0.95-1.50)	1.20 (0.95-1.51)	1.22 (0.88-1.68)
Cadres et prof. intellectuelles	0.99 (0.82-1.19)	0.99 (0.82-1.19)	0.96 (0.75-1.22)
Employés	1.15 (0.88-1.51)	1.15 (0.88-1.51)	0.87 (0.74-1.03)
Ouvriers qualifiés	1.08 (0.90-1.29)	1.08 (0.90-1.28)	1.10 (0.83-1.47)
Ouvriers non qualifiés	1.60*** (1.23-2.08)	1.62*** (1.25-2.11)	1.03 (0.80-1.33)
Inactifs	7.18 (0.92-55.82)	6.07 (0.77-47.70)	1.08 (0.75-1.57)
Statut marital		<i>(ref= en couple)</i>	
Célibataire	1.12 (0.77-1.62)	1.13 (0.77-1.64)	1.14 (0.90-1.44)
Assurance maladie complémentaire		<i>(ref= privée)</i>	
CMU-C	1.18 (0.81-1.72)	0.53** (0.30-0.94)	1.05 (0.76-1.45)
Non	1.40 (1.00-1.97)	1.10 (0.64-1.90)	1.20 (0.85-1.70)
Assurance maladie compl.* année		<i>(ref= assurance maladie compl. = 'privée'; année = '2008')</i>	
CMU-C * 2010	...	4.31*** (1.90-9.76)	...
État de santé et recours aux soins			
Maladie chronique déclarée		<i>(ref= non)</i>	
Oui	0.95 (0.82-1.10)	0.96 (0.83-1.11)	0.97 (0.84-1.12)
État de santé auto-déclaré		<i>(ref= excellent)</i>	
[0-4] : mauvais	1.31 (0.94-1.82)	1.32 (0.95-1.83)	1.48** (1.07-2.03)
[5-6] : moyen	1.21 (0.97-1.50)	1.20 (0.97-1.49)	1.17 (0.95-1.45)
[7-8] : bon	1.02 (0.86-1.21)	1.03 (0.86-1.22)	1.14 (0.96-1.36)
Nb. de visites chez le généraliste (12 derniers mois)		<i>(ref= plus de 3)</i>	
Moins de 2	1.22*** (1.06-1.41)	1.22*** (1.06-1.41)	1.18** (1.02-1.36)
Nb. de visites chez le spécialiste (12 derniers mois)		<i>(ref= au moins une)</i>	
Aucune	1.29*** (1.12-1.49)	1.30*** (1.12-1.50)	1.68*** (1.45-1.96)
Comportements à risque			
Consommation de tabac		<i>(ref= non fumeur)</i>	
Ex-fumeur	1.08 (0.93-1.27)	1.08 (0.93-1.27)	1.00 (0.85-1.17)
<= 10 cigarettes/jour	1.28** (1.03-1.59)	1.29*** (1.04-1.62)	1.20 (0.94-1.52)
> 10 cigarettes/jour	1.70*** (1.31-2.22)	1.71*** (1.31-2.23)	1.68*** (1.23-2.28)
Consommation d'alcool		<i>(ref= consommateur sans risques)</i>	
Jamais	1.19 (0.96-1.48)	1.18 (0.95-1.47)	1.22** (1.05-1.42)
Consommateurs à risque ponctuel	0.95 (0.82-1.11)	0.95 (0.82-1.11)	1.19 (0.95-1.49)
Consommateurs à risque chronique	1.13 (0.93-1.37)	1.13 (0.93-1.38)	1.44 (0.98-2.09)
Année d'enquête		<i>(ref= 2008)</i>	
2008	1.15 (0.99-1.35)	1.12 (0.95-1.23)	1.52*** (1.30-1.77)
2010	0.45*** (0.39-0.52)	0.42*** (0.36-0.49)	0.45*** (0.39-0.52)
Dépistage du cancer du sein (modèle 2 uniquement)			
Non	3.85*** (3.20-4.63)

Les résultats sont ajustés par la variable de localisation géographique. Les p-values sont indiquées par des étoiles. Niveau de significativité: ***=1% ; **=5% ; *=10%.

2.2.8. Discussion

L'objectif de cette étude était d'étudier les déterminants du non recours au dépistage du cancer du sein, du col de l'utérus et du côlon ainsi que leur évolution dans le temps, entre 2006 et 2010. Nous montrons que les taux de recours au dépistage des cancers du sein et du col de l'utérus ont diminué entre 2006 et 2010. Ces résultats vont dans le sens d'autres études menées dans différents contextes (Martinez-Huedo et al 2012) mais sont particulièrement décevants dans le contexte du renforcement des politiques de promotion du dépistage en France (le *Plan Cancer 2009-2013*, notamment). En revanche, le taux de recours au dépistage du cancer colorectal a augmenté fortement entre 2006 et 2010, ce qui peut s'expliquer par l'introduction d'un programme de dépistage organisé en 2009.

D'après les résultats de cette étude, les déterminants de la demande de dépistage des cancers agissent différemment selon la nature du cancer et semblent pour certains avoir varié dans le temps : l'âge, le statut socioéconomique et l'état de santé auto-déclaré avaient par exemple l'impact le plus fort sur le recours au dépistage du cancer du col de l'utérus comparé aux autres cancers. L'absence d'un programme de dépistage organisé pour le cancer du col de l'utérus et, par conséquent, l'attention plus faible portée par les médias de même que le coût nécessairement plus élevé du dépistage pourrait expliquer ces résultats (Walsh et al 2011). Dans la mesure où le frottis est souvent réalisé durant une consultation chez le gynécologue, les inégalités liées à l'accès aux soins sont plus fortes. En outre, les femmes vivant en couple ont plus de chance de consulter un gynécologue dans le cadre d'une demande de contraception par exemple, et donc d'être dépistées. Le taux de recours au frottis plus élevé pour les femmes âgées entre 25 et 34 ans est sans doute lié aux consultations associées à la contraception et/ou à une grossesse augmentant la fréquence des visites chez le gynécologue. Par extension, le recours au dépistage du cancer du col de l'utérus est certainement lié à la proximité d'un gynécologue. Des analyses complémentaires menées en incluant comme variable explicative le fait d'avoir consulté un gynécologue dans les 12 derniers mois ont montré un impact très significatif de cette variable sur le recours au frottis (OR = 7.8 [6.2-9.4]).

En ce qui concerne le dépistage du cancer du sein, les caractéristiques socio-économiques avaient un impact sur la probabilité d'effectuer une mammographie, mais moins prononcé que pour le dépistage du cancer du col de l'utérus. Ainsi, malgré l'existence d'un dépistage organisé depuis 2004 et de campagnes informatives touchant a priori une grande

partie de la population, les inégalités socioéconomiques liées au fait de ne pas posséder une couverture maladie complémentaire (ou de l'avoir gratuitement au travers de la CMU-C) persistent dans le temps. En France, l'absence de couverture complémentaire ou le fait de posséder la CMU-C sont généralement associés à des revenus plus faibles du ménage. De plus, même si la mammographie est gratuite dans le cadre du dépistage organisé, un accès plus faible au système de santé constitue un obstacle à la réalisation du dépistage. Ces résultats montrent donc la nécessité de politiques complémentaires au dépistage organisé permettant de réduire les inégalités d'accès aux soins.

En ce qui concerne le dépistage du cancer colorectal, malgré l'introduction d'un programme de dépistage organisé en 2009, les différences de recours liées à l'âge ont persisté : les individus âgés de moins de 54 ans avaient une probabilité plus faible d'être dépistés. Le fait de posséder une assurance complémentaire n'avait pas d'effet sur le recours au dépistage du cancer colorectal sur la période d'étude. En revanche, les inégalités socioéconomiques semblent s'être accrues : chez les hommes, le fait de posséder la CMU-C diminuait la probabilité de recourir au dépistage en 2010, mais pas en 2006 ou en 2008. En outre, les individus ayant un moindre recours au généraliste et au spécialiste recouraient moins au dépistage. Cela montre que des politiques visant à réduire les inégalités d'accès aux soins doivent nécessairement accompagner le dépistage organisé pour augmenter son efficacité.

Un résultat original de cette étude concerne l'impact des comportements à risque, tels que mesurés par la consommation de tabac et d'alcool, sur le recours au dépistage. D'un point de vue économique, ces pratiques à risque sont le reflet d'arbitrages guidés par des préférences temporelles (Becker and Mulligan 1997) ou des perceptions/aversion au risque pouvant engendrer des comportements de prévention différents (Katapodi et al 2004). Le moindre recours observé chez les individus ayant des comportements risqués en termes de consommation de tabac et d'alcool est cohérent avec des études antérieures (Hsia et al 2000; Selvin and Brett 2003) : ces individus pourraient avoir une préférence plus forte pour le présent et/ou considérer un retour sur investissement plus faible du dépistage (Grossman, M 1972). Toutefois, de façon plus surprenante, les individus n'ayant aucun comportement à risque (ne consommant jamais d'alcool et ne fumant jamais) étaient aussi moins susceptibles de recourir au dépistage. Ce résultat pourrait s'expliquer soit par la peur de la maladie engendrant une crainte face au résultat du dépistage : en ce sens, ces individus pourraient être considérés comme plus averses au risque ; soit par une perception du risque plus faible liée à

des comportements sains engendrant un sentiment de moindre vulnérabilité (Mills et al 2008). Dans tous les cas, ce résultat suggère la nécessité d'un renforcement des politiques de prévention primaire, de manière à combattre par exemple l'idée que le cancer n'affecte que les individus ayant des pratiques à risque.

Limites

Les principales limites de cette étude sont liées à la nature déclarative des données. En effet, il n'est pas possible d'exclure des biais de désirabilité sociale ayant conduit à une surrestimation du recours au dépistage du cancer. En France, l'Institut National de Veille Sanitaire (INVS) fournit une estimation annuelle du recours au dépistage du cancer du sein (depuis 2004) et du cancer colorectal (depuis 2009). En ce qui concerne le cancer du sein, les taux s'élèvent à 49.3% des femmes dépistées en 2006, 52.5% en 2008 et 52.0% en 2010¹⁶ et pour le cancer colorectal, 34% des individus ont été dépistés entre 2009 et 2010¹⁷. Ainsi, les taux de dépistage déclarés dans les différentes vagues des enquêtes SPS semblent surrestimer les taux de l'INVS (même s'ils ne sont pas directement comparables car ils sont annuels et non bi-annuels), mais restent cohérents par rapport aux résultats d'autres enquêtes basées sur des données déclaratives (Beck and Gautier 2010).

Une deuxième limite de l'étude est liée à l'absence de certains déterminants du recours au dépistage, principalement subjectifs, liés aux antécédents familiaux de cancer (Simou et al 2011), aux croyances et aux connaissances des individus (Lostao et al 2001; Klug et al 2005; Tacken et al 2007; Hall et al 2013), ainsi qu'à la disponibilité des médecins (Bingham et al 2003; Agurto et al 2004). D'autres études ont aussi mis en évidence l'impact du revenu sur la probabilité de recours au dépistage (Duport and Ancelle-Park 2006; Rigal et al 2011). Bien que des données sur le revenu aient été collectées dans les enquêtes SPS, nous n'avons pas ajusté par cette variable, du fait d'une proportion importante de non réponses (plus de 25%) et pour éviter des biais de multicollinéarité avec les variables de classe sociale et d'assurance complémentaire très corrélées au revenu.

¹⁶ <http://www.invs.sante.fr/Dossiers-thematiques/Maladies-chroniques-et-traumatismes/Cancers/Evaluation-des-programmes-de-depistage-des-cancers/Evaluation-du-programme-de-depistage-du-cancer-du-sein>.

¹⁷ <http://www.invs.sante.fr/Dossiers-thematiques/Maladies-chroniques-et-traumatismes/Cancers/Evaluation-des-programmes-de-depistage-des-cancers/Evaluation-du-programme-de-depistage-du-cancer-colorectal>.

Conclusion

Nos résultats montrent la persistance d'obstacles au dépistage, qui sont principalement liés aux caractéristiques socioéconomiques, au recours aux soins et aux comportements à risques. Cependant, nous ne sommes pas capables de déterminer précisément si le recours plus faible observé est lié à des choix personnels ou à un accès plus faible au système de santé. Ainsi, les études futures devraient étudier plus précisément les impacts respectifs des attitudes, croyances et perceptions du risque d'un côté, et des difficultés d'accès aux soins de l'autre côté, sur le recours au dépistage.

Du côté de l'offre de soins, les résultats sont probants : le recours au MG est un déterminant significatif du recours au dépistage. Ce résultat atteste qu'il y a bien des marges de manœuvre pour des schémas visant à améliorer l'offre de dépistage des cancers par le médecin généraliste. En particulier, les paiements à la performance pourraient être un outil susceptible d'engendrer une pratique plus efficace du MG. Dans ce contexte, il semble ainsi indispensable d'analyser l'impact des incitations financières sur le dépistage des cancers ciblés par les indicateurs du CAPI.

2.3. Impact du CAPI sur le recours au dépistage du cancer du sein

2.3.1. Objectif et hypothèse à tester

Dans la section précédente, nous nous sommes intéressés aux déterminants, côté demande, du recours au dépistage des cancers. Ce travail constituait la première étape de la recherche visant à comprendre l'impact des incitatifs financiers, implémentés à destinations des médecins généralistes français, sur le recours au dépistage des cancers. Dans la suite de ce travail, l'objectif est d'étudier précisément l'impact du CAPI sur le recours au dépistage du cancer du sein. En effet, en 2009, seul le dépistage du cancer du sein faisait partie des indicateurs évalués par le dispositif¹⁸.

L'hypothèse que nous souhaitons tester est donc la suivante : le CAPI a incité les médecins signataires à augmenter leur temps ou leurs efforts dédiés à convaincre/conseiller les femmes âgées entre 50 et 74 ans, n'ayant pas effectué de mammographie dans les deux ans, à le faire. Ainsi, de manière empirique, nous souhaitons tester l'hypothèse selon laquelle avoir consulté un médecin généraliste traitant (MGT) signataire du CAPI durant la période d'exposition au contrat augmente la probabilité d'effectuer une mammographie. Le CAPI a été instauré en juillet 2009 et a été généralisé à tous les MGs en janvier 2012, nous limitons donc notre période d'intérêt à ces deux années et demi. Cette période constitue un délai suffisant pour effectuer une mammographie au regard des recommandations nationales. Afin d'identifier des biais de sélection potentiels liés au fait que les médecins signataires du CAPI seraient préalablement plus « performants » en termes de dépistage du cancer, il est nécessaire de mener l'analyse sur une période antérieure à la signature du CAPI, de même durée. Nous nous intéressons donc à deux périodes : la période 1 antérieure à la mise en place du CAPI (du 01/01/2007 au 30/06/2009) et la période 2 d'exposition au CAPI (du 01/07/2009 au 31/12/2011).

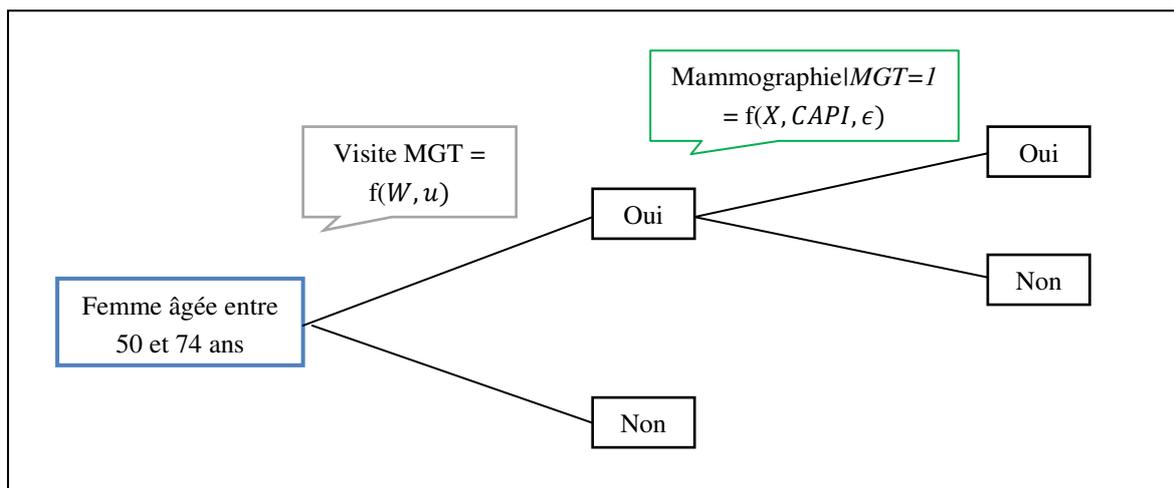
Durant chacune de ces périodes, une femme a la possibilité de consulter un MGT et d'effectuer une mammographie. Compte tenu des résultats de la littérature (Kenkel 1994; Carrieri and Bilger 2013) et des analyses menées dans la section précédente, il semble cohérent de poser l'hypothèse que la probabilité d'effectuer une mammographie est plus élevée pour les femmes ayant consulté un médecin généraliste traitant sur la période. De

¹⁸ Comme les incitatifs touchant le dépistage du cancer du col de l'utérus ont été implémentés à destination de l'ensemble des médecins (au travers de la ROSP), l'évaluation de ce dispositif nécessite l'utilisation de méthodes statistiques différentes, et fera l'objet d'un travail ultérieur qui dépasse le cadre de cette thèse.

même, nous faisons l’hypothèse que les déterminants inobservés expliquant le recours au MGT et le recours à la mammographie (niveau d’aversion/perception du risque, préférence pour le présent, facteurs psychosociaux) sont corrélés positivement (Carrieri and Bilger 2013).

Notre variable explicative d’intérêt étant la signature du CAPI par le médecin généraliste traitant (MGT), notre analyse ne porte donc que sur les femmes qui ont consulté un MGT sur la période considérée. Ainsi, nous nous plaçons dans un cadre d’analyse similaire à celui développé par Kenkel (1994), qui est schématisé par la figure 5. Ainsi, Kenkel (1994) modélise le recours au dépistage par un processus en deux étapes : en première étape, il modélise la probabilité de recourir au MG en fonction de caractéristiques exogènes des individus. En seconde étape, pour le sous-échantillon des individus ayant eu un recours au MG, il modélise la probabilité d’effectuer la mammographie comme une fonction de caractéristiques exogènes de l’individu ainsi que des caractéristiques des médecins. Dans notre modèle, la probabilité de consulter un MGT sur la période est spécifiée comme une fonction de variables sociodémographiques et d’état de santé (W) et d’un terme aléatoire u . Dans un second temps, pour les femmes ayant effectivement consulté un MGT, la probabilité d’effectuer une mammographie est spécifiée comme une fonction de variables sociodémographiques et de recours aux soins (X), ainsi qu’une indicatrice de visite chez un MGT signataire du CAPI (CAP) et d’un terme d’erreur aléatoire ϵ .

Figure 5. Modélisation du recours à la mammographie en fonction de la visite chez le médecin traitant.



Dans la section suivante, nous présentons les données de l’Echantillon Généraliste des Bénéficiaires, utilisées afin d’estimer ce modèle et de tester notre hypothèse de recherche.

2.3.2. L'échantillon généraliste des bénéficiaires (EGB)

En France, le Système National d'Information Inter- Régime de l'Assurance Maladie (SNIIR-AM) est le principal système d'information de l'assurance maladie ; il permet d'obtenir des informations exhaustives sur les consommations ayant donné lieu à remboursement de l'ensemble des affiliés aux trois principaux régimes d'assurance maladie (CNAMTS, MSA, RSI) et sur les professionnels de santé les ayant prescrit et/ou exécutés. Construit à partir d'une extraction du SNIIR-AM au 1/97^e, l'Echantillon Généraliste des Bénéficiaires (EGB) est un échantillon représentatif des personnes protégées des régimes de sécurité sociale en termes d'âge, de sexe et de localisation géographique, qui associe des informations sociodémographiques des bénéficiaires et des données de remboursement de soins de ces personnes. Ainsi, l'EGB produit une information détaillée sur les patients qui font l'objet de soins, les différents épisodes de soins (codes des prescriptions et tests, et codes CCAM¹⁹ détaillés de l'ensemble des actes techniques médicaux) ainsi que sur les professionnels qui les dispensent. Dans l'EGB, les données de remboursement sont conservées vingt ans au-delà de l'année en cours, ce qui permet de reconstituer l'ensemble du parcours de soins des patients sur une longue période et donc d'effectuer un suivi longitudinal des pratiques (De Roquefeuil et al 2009). Ces données sont disponibles depuis 2007 pour le régime général (RG) et permettent de suivre le parcours de soins de plus de 500 000 bénéficiaires et ayants-droits depuis 2003 jusqu'à aujourd'hui. Enfin, élément essentiel pour notre travail, l'EGB contient un listing des médecins ayant signé le CAPI ainsi que la date de signature du contrat.

2.3.3. Accès à l'EGB et constructions des bases de données

L'accès aux données de l'EGB nécessite une formation préalable, permettant de comprendre et de se familiariser avec l'architecture complexe du SNIIR-AM (cf. Figure A.3, Annexe A) et de l'EGB (cf. Figure A.4). Une étape préliminaire avant toute analyse consiste à sélectionner les observations et variables d'intérêt à partir des différentes bases de données de l'EGB au moyen d'une jointure adéquate. Les tables contenant l'information qui nous intéressent sont :

- La table *Prestations (EB_PRS_F)* est la pierre angulaire de l'EGB : elle contient l'ensemble des informations relatives à la date de prescription ou d'exécutions des

¹⁹ Classification Commune des Actes Médicaux.

soins, le numéro et la spécialité du médecin prescripteur ou exécutant, ainsi que la date de mise à jour des données. Cette table contient au total 412 519 231 observations.

- La table *Référentiels des bénéficiaires (EB_INB_F)* contient essentiellement les informations sur les bénéficiaires: âge, sexe, département, régime d'Assurance Maladie, date de naissance et de décès. Elle se « rattache » à la table *Prestations* par le biais du numéro de sécurité sociale (le NIR) anonymisé du bénéficiaire. Elle contient au total 708 208 observations correspondant à autant de bénéficiaires (y compris les ayants droit) présents dans l'EGB.
- La table *Codage CCAM (EB_CAM_F)* contient les informations relatives aux actes médicaux codés selon la Classification Commune des Actes Médicaux. C'est à partir de cette base que l'on retrouve les codes pour la mammographie. Elle est reliée à la table *Prestations* par une clé technique de jointure commune à l'ensemble des tables. Cette table contient 12 046 545 observations.
- La table *Prestations Affinée Biologie (EB_BIO_F)* regroupe les prestations affinées biologie et est reliée à la table *Prestations* grâce à une clé technique de jointure. Les prestations de biologie nous intéressent qu'indirectement: elles permettent de reconstituer un parcours de soins (visites chez certains spécialistes) sans se limiter aux prestations ayant donné lieu à des actes médicaux. Cette table contient 59 729 341 observations.
- Enfin, la table *CP_CAP_R_201307* qui contient les informations relatives à la date de signature du CAPI des 16 428 MGs signataires, ainsi qu'une clé d'anonymisation permettant de la relier à la table *Prestations*.

A partir de la table *Prestations* jointe avec le *Référentiel médicalisé des bénéficiaires*, les données de *Codage affiné CCAM* et les *Prestations de Biologie*, nous sélectionnons sur la période 2007-2011 les observations des femmes non décédées âgées de 50 à 74 ans, n'ayant pas d'ALD cancer et hors hospitalisations, soit 88 157 femmes et 100 610 110 observations au total.

A partir de cette base de données intermédiaire, deux bases de données finales sont constituées. Dans la première base de données (dite « pre-CAPI »), nous sélectionnons les

observations dont la date de prescription des soins (par n'importe quel professionnel de santé) est antérieure au 30 Juin 2009 et dont la date d'exécution est antérieure au 31 Juillet 2009, soit au total 61 286 femmes. Dans la seconde base de données (dite « post-CAPI »), nous sélectionnons les observations dont la date de prescription des soins est comprise entre le 1^{er} Juillet 2009 et le 31 Décembre 2011 et la date d'exécution est antérieure au 31 Janvier 2012, soit au total 74 532 femmes.

Une analyse descriptive succincte de ces bases révèle une répartition inégale des différents régimes d'Assurance Maladie : la première base comprend plus de 99% de bénéficiaires du RG alors que la seconde en compte 90%. Cette différence est due au fait que les observations des autres régimes (MSA, RSI et régimes spéciaux) ne sont présents dans l'EGB que depuis 2009. Ainsi, afin de travailler sur une population similaire, nous réduisons notre population d'étude aux seuls bénéficiaires du régime général, soit 61 141 femmes dans base « pre-CAPI » et 67 316 femmes dans la base « post-CAPI ».

2.3.4. Variables dépendantes et indépendantes

2.3.4.1. Réalisation de la mammographie

La variable dépendante est une indicatrice de réalisation d'une mammographie de dépistage, qui vaut 1 si la femme a effectué une mammographie sur la période considérée, 0 sinon. Dans l'EGB, la mammographie est codée de trois façons différentes : le code QEQK004 est utilisé pour coder les mammographies de dépistage organisé, et les codes QEQK001 et QEQK005 correspondent à des mammographies de dépistage individuel ou de suivi²⁰. Comme la rémunération associée au CAPI ne distingue pas entre mammographie de dépistage organisé et mammographie de dépistage individuel, nous considérons les trois codages pour la construction de la variable dépendante de réalisation d'une mammographie.

2.3.4.2. Déterminants du recours à la mammographie

Nous considérons trois types de variables indépendantes : les variables sociodémographiques, les variables de recours aux soins, et la variable de visite chez un MGT signataire du CAPI. Les deux premiers types de variables sont construites de la même façon quelque soit la période d'analyse. Concernant les variables sociodémographiques nous

²⁰ Plus précisément, le code QEQK001 est utilisé pour la mammographie bilatérale, et le code QEQK005 est utilisé pour la mammographie unilatérale.

considérons l'âge, la localisation géographique, et le fait de posséder la CMU-C. Nous considérons deux variables dichotomiques de recours aux soins : la première prend la valeur 1 si la femme a consulté un gynécologue sur la période considérée, 0 sinon ; la seconde prend la valeur 1 si la femme a consulté un dentiste ou un ophtalmologue sur la période considérée, 0 sinon²¹. La variable d'état de santé (ALD) n'est utilisée que pour expliquer le recours au MGT et non le recours à la mammographie car, en tant que soin curatif, le recours au MG est supposé être influencé par l'état de santé alors que le soin préventif, en tant que décision d'investissement, est a priori indépendant (Grossman, M 1972; Kenkel 1994; Lairson et al 2005; Carrieri and Bilger 2013).

La variable de visite chez un MGT signataire du CAPI est construite différemment selon la période d'analyse. Dans la période 1, cette variable est codée 1 si la femme a vu au moins une fois, entre 2007 et 2009, un MGT futur signataire du CAPI. Dans la période 2, cette variable est codée 1 si la femme a vu au moins une fois, entre 2009 et 2011, un MGT dont la date de signature du CAPI est antérieure à la date de la consultation (i.e. la date de prescription des soins).

2.3.5. Analyse économétrique

Pour chaque période d'analyse, nous estimons un modèle de sélection bivarié comprenant deux équations. Dans la première équation, nous modélisons la probabilité d'avoir consulté un MGT sur la période considérée. Dans la seconde équation, nous modélisons la probabilité d'avoir effectué une mammographie, conditionnellement à avoir vu un MGT. Ce type de modèle est similaire au modèle de sélection d'Heckman, excepté que la deuxième équation est aussi un modèle probit.

Notons z^* la variable latente associée au fait d'avoir vu un MGT. Le processus d'observation pour la femme i est le suivant:

$$z_i^* = W_i' \gamma + u_i, \quad z_i = \begin{cases} 1 & \text{si } z_i^* > 0 \\ 0 & \text{si } z_i^* \leq 0 \end{cases} \quad (1)$$

²¹ Comme nous ne disposons pas ou peu de caractéristiques individuelles, notamment socio-économiques, cette variable est utilisée comme un proxy du niveau de revenu et du fait de posséder une assurance maladie complémentaire, supposés corrélés au recours aux soins dentaires ou ophtalmologiques (Borrell et al 2001; Bedos et al 2004; Azogui-Levy and Rochereau 2005).

Avec $z_i = 1$ si la femme a vu un MGT, W_i est un vecteur de variables sociodémographiques et d'état de santé de la femme i sur la période considérée, γ est un vecteur de paramètres à estimer et u_i est un terme d'erreur supposé suivre une loi normale centrée réduite.

Notons y^* la variable latente associée au fait d'avoir effectué une mammographie, conditionnellement au recours au MGT. Le processus d'observation pour la femme i est le suivant :

$$\begin{cases} y_i^* = X_i' \beta + \varepsilon_i & \text{si } z_i = 1 \\ y_i^* = 0 & \text{si } z_i = 0 \end{cases}, \quad y_i = \begin{cases} 1 & \text{si } y_i^* > 0 \\ 0 & \text{si } y_i^* \leq 0 \end{cases} \quad (2)$$

Avec X_i un vecteur de variables sociodémographiques et de recours aux soins de la femme i sur la période considérée, incluant une indicatrice qui vaut 1 si la femme a consulté un MGT ayant signé le CAPI, 0 sinon. Par ailleurs, les hypothèses sur les termes d'erreur ε_i et u_i et leur relation sont les suivantes :

$$\begin{pmatrix} u_i | W_i \\ \varepsilon_i | X_i \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & \rho \\ \rho & \sigma^2 \end{pmatrix} \right)$$

Où ρ représente le coefficient de corrélation entre les deux termes d'erreur, et σ l'écart type de ε_i .

Il y a trois types d'observations dans ce modèle auxquelles sont associées les probabilités suivantes :

$$z_i = 0 \quad \Pr(z_i = 0) = \Phi(-W_i' \gamma)$$

$$z_i = 1, y_i = 0 \quad \Pr(z_i = 1, y_i = 0) = \Phi(W_i' \gamma) - \Phi_2(W_i' \gamma, X_i' \beta, \rho)$$

$$z_i = 1, y_i = 1 \quad \Pr(z_i = 1, y_i = 1) = \Phi_2(W_i' \gamma, X_i' \beta, \rho)$$

Avec $\Phi(\cdot)$ la fonction de distribution cumulée de loi normale centrée réduite.

La log vraisemblance de l'échantillon s'écrit donc de la façon suivante :

$$\begin{aligned} \ln L = & \sum_{i=1}^N \{z_i y_i \ln \Phi_2(W_i' \gamma, X_i' \beta, \rho) \\ & + y_i(1 - z_i) \ln [\Phi(W_i' \gamma) - \Phi_2(W_i' \gamma, X_i' \beta, \rho)] \\ & + (1 - z_i) \ln \Phi(-W_i' \gamma)\} \end{aligned}$$

La procédure de maximisation de cette log- vraisemblance telle qu'implémentée par exemple dans le logiciel SAS (PROC QLIM) utilise un algorithme d'optimisation non linéaire de type Quasi Newton (Dennis and Moré 1977).

2.3.6. Analyse descriptive du recours au dépistage du cancer du sein

Le tableau 8 présente les statistiques descriptives sur les deux périodes. Le pourcentage de réalisation d'une mammographie est similaire d'une période à l'autre, égal à 63,8% sur la période antérieure au CAPI (période 1) et 63,3% sur la période post-CAPI (période 2). Parmi les femmes ayant effectué au moins une mammographie sur la période considérée, la majorité n'en a effectué qu'une seule (78,9% en période 1 et 82,5% en période 2, respectivement) et la quasi-totalité moins de 2. Par ailleurs, en ce qui concerne les premières mammographies, plus de 3 mammographies sur 4 sont des mammographies de dépistage organisé (code QEQK004).

La distribution des variables indépendantes est similaire d'une période à l'autre. En période 1 (respectivement en période 2), 55,1% (resp. 53,6%) des femmes ont entre 50 et 59 ans, 6,0% (resp. 6,1%) ont la CMU-C, et 13,2% (resp. 13,5%) ont une affection de longue durée (ALD). Par ailleurs, les femmes sont réparties sur l'ensemble du territoire et de façon similaire d'une période à l'autre. En ce qui concerne les variables de recours aux soins, la majorité des femmes ont vu un MGT sur la période (87,9% en période 1 et 88,9% en période 2); elles sont 30,9% (resp. 32,2%) à avoir vu un MGT signataire du CAPI, 13,7% (resp. 12,8%) à avoir vu un gynécologue et 29,7% (resp. 31,0%) à avoir effectué une consultation chez un dentiste ou un ophtalmologue.

Tableau 8. Statistiques descriptives sur les échantillons pre- et post- CAPI

	Période Pre-CAPI (N=61 141)		Période Post-CAPI (N=67 316)	
	N	%	N	%
Mammographie effectuée*				
Non	22106	36.16	24702	36.70
Oui	39035	63.84	42614	63.30
Nombre mammographies effectuées*				
0	22106	36.16	24702	36.70
1	30808	50.39	35166	52.24
2	7712	12.61	7072	10.51
3	515	0.84	372	0.55
Code première mammographie				
QEQL001	9003	23.06	8340	19.57
QEQL004	29648	75.95	33877	79.50
QEQL005	384	0.98	397	0.93
Age				
50-54	19945	32.62	21348	31.71
55-59	13748	22.49	14735	21.89
60-64	10589	17.32	13385	19.88
65-74	16859	27.57	17848	26.51
Région				
Alsace	1823	2.98	1960	2.91
Aquitaine	2880	4.71	3179	4.72
Autre (ex: Outre-mer)	1830	2.99	2237	3.32
Auvergne	1384	2.26	1486	2.21
Basse Normandie	1528	2.50	1625	2.41
Bourgogne	1671	2.73	1812	2.69
Bretagne	2872	4.70	3126	4.64
Centre	2366	3.87	2616	3.89
Champagne-Ardenne	1234	2.02	1341	1.99
Franche-Comté	1116	1.83	1227	1.82
Hte Normandie	1936	3.17	2095	3.11
Ile de France (hors Paris)	8248	13.49	9114	13.54
Inconnu	1231	2.01	1066	1.58
Languedoc-Roussillon	2352	3.85	2723	4.05
Limousin	694	1.14	742	1.10
Lorraine	2264	3.70	2450	3.64
Midi-Pyrénées	2550	4.17	2903	4.31
Nord-Pas-De-Calais	3959	6.48	4193	6.23
PACA	5135	8.40	5699	8.47
Paris	1804	2.95	2074	3.08
Pays de la Loire	3274	5.35	3642	5.41
Picardie	1864	3.05	2058	3.06
Poitou-Charente	1602	2.62	1783	2.65
Rhone-Alpes	5524	9.03	6165	9.16
CMU-C				
Non	57487	94.02	63221	93.92
Oui	3654	5.98	4095	6.08
ALD				
Non	53085	86.82	58249	86.53
Oui	8056	13.18	9067	13.47
Visite MGT*				
Non	7396	12.10	7457	11.08
Oui	53745	87.90	59859	88.92
Visite MGT CAPI*				
Non	42267	69.13	45629	67.78
Oui	18874	30.87	21687	32.22
Visite chez le gynécologue*				
Non	52764	86.30	58680	87.17
Oui	8377	13.70	8636	12.83
Visite chez le dentiste ou ophtalmologue*				
Non	42976	70.29	46422	68.96
Oui	18165	29.71	20894	31.04

*Sur la période considérée (pre- ou post-CAPI)

Le tableau 9 présente les dates de signatures du CAPI des 16 428 médecins ayant adhéré à ce contrat. Plus de deux tiers (70%) des médecins ont signé le contrat dans les quatre mois suivant sa mise en place, i.e. avant octobre 2009, et 90% l'ont signé avant la date anniversaire du contrat. De fait, la majeure partie de la période 2 peut être considérée comme une période d'exposition au CAPI.

Tableau 9. Date de signature du CAPI (N=16 428 individus)

Date de signature du CAPI	N	%	Cum Freq	Cum %
juil-09	5419	33.0	5419	33.0
oct-09	6020	36.6	11439	69.6
janv-10	2030	12.4	13469	82.0
avr-10	712	4.3	14181	86.3
juil-10	586	3.6	14767	89.9
oct-10	419	2.6	15186	92.4
janv-11	695	4.2	15881	96.7
avr-11	367	2.2	16248	98.9
juil-11	124	0.8	16372	99.7
oct-11	54	0.3	16426	100.0
janv-12	2	0.0	16428	100.0

Le tableau 10 présente les résultats des tests d'indépendance du Chi-deux entre la réalisation de la mammographie et chacune des variables indépendantes. On constate une relation non linéaire entre l'âge et le recours à la mammographie : les femmes plus jeunes (50-54 ans) et plus âgées (65-74 ans) recourent moins à la mammographie que les femmes âgées entre 55 et 64 ans ($p < 0.0001$). Il y a aussi des disparités régionales de recours au dépistage : les femmes des régions Nord-Pas-de-Calais, Picardie et d'Outre-mer recourent moins à la mammographie, et celles des régions Auvergne, Basse Normandie, Centre, ou Pays de la Loire recourent davantage ($p < 0.0001$). Le fait de posséder la CMU-C, de ne pas avoir consulté un MGT, un gynécologue, un dentiste ou un ophtalmologue sur la période considérée, constituent des freins au recours à la mammographie. Ainsi, parmi les femmes n'ayant pas consulté une MGT en période 1, seulement 34.8% ont effectué une mammographie. Au contraire, parmi les femmes qui ont consulté un gynécologue en période 1, 86.3% des femmes ont effectué une mammographie.

Dans un second temps, parmi les individus ayant consulté un MGT sur la période, nous avons étudié le lien entre le fait d'avoir vu un MGT signataire du CAPI et le recours à la mammographie. Rappelons qu'en période 1, aucun MG n'a encore pu adhérer au CAPI, mais on considère là les futurs signataires pour éventuellement appréhender un biais de sélection. Ce lien est non significatif pour les deux périodes au seuil de 5%, même si en période 2, les

femmes ayant consulté un MGT signataire du CAPI sont 67.0% à avoir effectué une mammographie, contre 66.3% pour l'ensemble des femmes ayant vu un MGT.

Tableau 10. Relation entre les covariables et la réalisation d'une mammographie

	Période Pre-CAPI (N=61 141)			Période Post-CAPI (N=67 316)		
	Row %	Chi-Square	P-value	Row %	Chi-Square	P-value
Age		911.1376	<.0001		1041.9292	<.0001
50-54	58.6			56.5		
55-59	70.3			69.0		
60-64	72.0			71.3		
65-74	59.7			60.8		
Région		656.3754	<.0001		645.2272	<.0001
Alsace	64.5			63.9		
Aquitaine	67.2			66.8		
Autre (ex: Outre-mer)	56.5			54.4		
Auvergne	71.8			69.3		
Basse Normandie	68.5			68.7		
Bourgogne	61.9			60.0		
Bretagne	63.7			64.8		
Centre	69.0			67.7		
Champagne-Ardenne	68.1			65.9		
Franche-Comté	60.6			61.6		
Hte Normandie	64.1			62.1		
Ile de France (hors Paris)	63.9			63.8		
Inconnu	43.5			36.9		
Languedoc-Roussillon	61.0			60.3		
Limousin	66.6			65.9		
Lorraine	63.3			62.2		
Midi-Pyrénées	64.7			64.5		
Nord-Pas-De-Calais	55.9			58.9		
PACA	64.4			64.3		
Paris	58.2			59.0		
Pays de la Loire	70.9			69.4		
Picardie	59.0			60.8		
Poitou-Charente	67.0			65.1		
Rhone-Alpes	67.8			65.2		
CMU-C		423.9880	<.0001		453.2112	<.0001
Non	64.9			64.3		
Oui	48.0			47.8		
Visite MGT*		3065.6330	<.0001		2109.5049	<.0001
Non	34.8			39.1		
Oui	67.8			66.3		
Visite chez le gynécologue*		2126.4849	<.0001		2138.9197	<.0001
Non	60.3			60.0		
Oui	86.3			85.7		
Visite dentiste ou ophtalmologue*		663.7712	<.0001		485.8084	<.0001
Non	60.6			60.6		
Oui	71.5			69.4		
Sous échantillon MGT vu	Row %	Chi-Square	P-value	Row %	Chi-Square	P-value
MGT CAPI vu	67.7	0.4469	0.5038	67.0	2.3729	0.1235

Lecture : Parmi les femmes âgées entre 50 et 54 ans, 58.6% ont effectué une mammographie sur la période 1 alors que ce chiffre est de 63.8% pour l'ensemble des femmes.

*Sur la période considérée.

2.3.7. Résultats des modèles économétriques

Le tableau 11 présente les résultats des modèles de sélection bi-variés pour les deux périodes. Seuls les résultats de l'équation de recours à la mammographie sont présentés, les résultats de l'équation de recours au MGT étant reportés dans le tableau A.1 (annexe A). Tout d'abord, le coefficient de corrélation ρ entre les termes d'erreurs des deux équations est positif et significatif pour les deux périodes ($\rho = 0.79$), validant l'hypothèse selon laquelle les éléments inobservés du recours au médecin traitant et à la mammographie sont corrélés positivement, et justifiant par ailleurs l'estimation d'un modèle de sélection. Les résultats concernant l'impact des variables sociodémographiques et de recours aux soins confirment les résultats des analyses bi-variées et sont similaires d'une période à l'autre : comparé aux femmes âgées entre 65 et 74 ans, les femmes âgées entre 55 et 64 ans ont une probabilité plus élevée de recourir à la mammographie ($p < 0.0001$), alors que les femmes plus jeunes (âgées entre 50 et 54 ans) ont une probabilité plus faible de recourir ($p < 0.0001$). Toutes choses égales par ailleurs, les femmes qui ont la CMU-C ont une probabilité plus faible d'être dépistées, et les femmes ayant eu au moins une visite chez le gynécologue ainsi qu'une visite chez le dentiste ou l'ophtalmologue ont une probabilité plus forte d'être dépistées.

En revanche, les résultats concernant l'impact de la visite chez un MGT signataire du CAPI diffèrent d'une période à l'autre : avant la mise en place du CAPI, le fait d'avoir consulté un MGT futur signataire du CAPI n'a pas d'impact sur le recours au dépistage ($p = 0.5357$). En revanche, durant la période d'exposition au CAPI, le fait d'avoir vu un MGT signataire du CAPI a un impact positif mais ténu, non significatif au seuil de 5% ($p = 0.0667$).

En ce qui concerne les déterminants du recours au MGT présentés dans le tableau A.1 (Annexe A), être plus âgé, avoir une affection de longue durée et avoir effectué au moins une consultation chez le dentiste ou l'ophtalmologue sont des déterminants associés à une probabilité accrue de recourir au MGT.

Tableau 11. Résultats du modèle conditionnel de recours à la mammographie

	Période pre-CAPI			Période post-CAPI		
	Estimate	t-value	p-value	Estimate	t-value	p-value
Constante	0.1692	7.78	<.0001	0.1512	7.68	<.0001
Age	<i>ref = 65-74 ans</i>			<i>ref = 65-74 ans</i>		
50-54	-0.1284	-8.42	<.0001	-0.2147	-15.44	<.0001
55-59	0.2332	13.6	<.0001	0.1626	10.47	<.0001
60-64	0.2922	16.06	<.0001	0.2435	15.33	<.0001
Région	<i>ref = Rhones-Alpes</i>			<i>ref = Rhones-Alpes</i>		
Alsace	-0.0403	-1.11	0.268	0.0016	0.05	0.9623
Aquitaine	-0.0221	-0.71	0.4757	0.0329	1.13	0.2573
Autre (ex: Outre-mer)	-0.1547	-4.18	<.0001	-0.1593	-4.71	<.0001
Auvergne	0.1581	3.83	0.0001	0.1238	3.2	0.0014
Basse Normandie	0.0157	0.4	0.6871	0.0755	2.03	0.0424
Bourgogne	-0.1245	-3.34	0.0008	-0.1207	-3.45	0.0006
Bretagne	-0.0871	-2.8	0.0052	-0.0084	-0.29	0.7729
Centre	0.0672	2.01	0.0442	0.0814	2.62	0.0089
Champagne-Ardenne	0.0857	2.01	0.0442	0.0985	2.47	0.0135
Franche-Comté	-0.1442	-3.31	0.0009	-0.0427	-1.04	0.2997
Hte Normandie	-0.0496	-1.39	0.1636	-0.0321	-0.96	0.3355
Ile de France (hors Paris)	-0.1638	-6.96	<.0001	-0.0934	-4.21	<.0001
Inconnu	-0.7019	-16.27	<.0001	-0.8232	-17.96	<.0001
Languedoc-Roussillon	-0.1449	-4.43	<.0001	-0.1041	-3.45	0.0006
Limousin	0.0078	0.15	0.8841	0.0221	0.43	0.6667
Lorraine	-0.0634	-1.87	0.0613	-0.0164	-0.52	0.6038
Midi-Pyrénées	-0.0503	-1.56	0.1189	0.0052	0.17	0.8613
Nord-Pas-De-Calais	-0.2136	-7.33	<.0001	-0.0836	-3.14	0.0017
PACA	-0.0743	-2.85	0.0043	-0.0142	-0.58	0.5592
Paris	-0.4167	-10.53	<.0001	-0.3365	-9.22	<.0001
Pays de la Loire	0.1220	4.05	<.0001	0.1344	4.81	<.0001
Picardie	-0.1802	-5.02	<.0001	-0.0327	-0.98	0.3286
Poitou-Charente	-0.0014	-0.04	0.9706	0.0266	0.75	0.4545
CMU-C	<i>ref = non</i>			<i>ref = non</i>		
Oui	-0.3347	-13.95	<.0001	-0.3313	-14.96	<.0001
Visite chez le gynécologue*	<i>ref = non</i>			<i>ref = non</i>		
Oui	0.6802	23.81	<.0001	0.6875	27.75	<.0001
Visite dentiste ou ophtalmologue*	<i>ref = non</i>			<i>ref = non</i>		
Oui	0.2488	20.62	<.0001	0.1940	17.12	<.0001
Visite MGT CAPI*	<i>ref = non</i>			<i>ref = non</i>		
Oui	-0.0066	-0.62	0.5357	0.0184	1.83	0.0667
Rho	0.7900	16.6	<.0001	0.7932	18.9	<.0001
Log vraisemblance	-53612			-58584		

2.3.8. Discussion

L'objectif de ce travail était d'analyser l'impact d'incitatifs financiers mis en place au travers du Contrat d'Amélioration des Pratiques Individuelles (CAPI) sur le recours à la mammographie. Les médecins signataires pouvaient bénéficier de de bonus allant jusqu'à 245€ par an associés à l'atteinte d'un taux cible (environ 80%) de patientes âgées entre 50 et 74 ans, dont ils étaient le médecin traitant, ayant effectué une mammographie dans les deux ans. L'hypothèse que nous souhaitions tester était donc la suivante : le CAPI a-t-il incité les médecins traitants signataires à améliorer leur offre de dépistage du cancer du sein ? Dans ce but, nous avons utilisé les données de l'Echantillon Généraliste des Bénéficiaires (EGB) extraites du Système National d'Information Inter-Régimes de l'Assurance Maladie (SNIIR-AM). Ces données nous ont semblé idéales pour répondre à notre problématique, du fait de l'exhaustivité des données collectées permettant de reconstituer le parcours des soins des femmes sur une longue période.

Notre stratégie d'identification de l'impact du CAPI sur le recours à la mammographie a reposé sur trois principaux choix méthodologiques. Le premier a été de considérer une approche « patient », et ainsi d'estimer l'effet d'avoir consulté un médecin généraliste traitant sur le recours au dépistage du cancer du sein. Si la plupart des travaux ayant évalué l'impact d'incitatifs économiques sur le recours au dépistage ont utilisé une approche « médecin » (ils ont par exemple comparé l'évolution des indicateurs agrégés de dépistage de la patientèle avant/après la mise en place du dispositif incitatif (Li et al 2014), cette approche n'a pas été retenue pour deux raisons. La première est la difficulté voire l'impossibilité d'obtenir les résultats d'indicateurs agrégés par médecin auprès de la CNAMTS : pour des raisons politiques, les organismes extérieurs à l'assurance maladie ne sont pas autorisés à obtenir ce type de données détaillées au niveau médecin²². La seconde raison est que le dépistage du cancer est le résultat de nombreux déterminants également du côté de la demande qu'il est nécessaire de prendre en compte dans la modélisation du recours à la mammographie.

Le second choix méthodologique a été de s'intéresser à deux périodes d'analyses : la période couvrant la mise en place du CAPI, i.e. du 1^{er} juillet 2009 à la fin de l'année 2011 (soit 2 ans et demi au total), et la période antérieure à la mise en place du CAPI, de même

²² Par ailleurs, il n'est pas possible de les reconstituer à posteriori car les données de l'EGB sont représentatives au niveau patient, mais non au niveau médecin.

durée, soit du 1^{er} janvier 2007 au 30 Juin 2009. Le fait de considérer une période antérieure à la mise en place du CAPI avait pour objectif de tester l'hypothèse d'un phénomène d'auto sélection des médecins signataires du CAPI, potentiellement plus « productifs » en terme de dépistage du cancer.

Le troisième choix méthodologique a concerné la modélisation économétrique adoptée, c'est-à-dire un modèle de sélection bi-varié permettant de tenir compte de phénomènes de sélection des femmes ayant consulté au moins une fois un médecin généraliste traitant sur la période. Ce choix a été guidé par l'hypothèse selon laquelle les déterminants inobservés du recours au médecin généraliste traitant (MGT) et à la mammographie étaient corrélés positivement, rendant la variable explicative d'intérêt (avoir consulté un MGT signataire du CAPI) partiellement endogène. Le modèle de sélection bi-varié nous a permis d'estimer l'effet propre de la visite chez un médecin signataire du CAPI sur le recours à la mammographie, séparément de l'effet d'auto sélection des femmes ayant recouru à leur médecin traitant.

Toutefois, l'estimation de cet impact en période « post-CAPI » était conditionné au fait que le recours au MGT soit antérieur à la mammographie, ce qui ne pouvait être garanti par construction de la variable²³. Afin de s'assurer que la plupart des femmes avaient recouru au moins une fois au MGT avant la date de leur première mammographie, nous avons calculé, parmi les femmes ayant vu un MGT signataire du CAPI et ayant effectué une mammographie (14 446 femmes), le pourcentage de femmes ayant vu un MGT signataire du CAPI au moins 6 mois avant d'effectuer leur première mammographie (14 002 femmes), soit 97%. De fait, nous mesurons bien l'impact de la visite chez le médecin sur le recours à la mammographie.

Les résultats obtenus nous permettent de dégager un message clair. Avant la mise en place du CAPI, le fait d'avoir consulté un MGT futur signataire du contrat n'a pas eu d'impact sur le recours au dépistage du cancer du sein ($p=0.5357$). Nous pouvons ainsi formuler l'hypothèse qu'avant la mise en place du CAPI, les médecins (futurs) signataires n'avaient pas de pratiques de dépistage plus efficaces que les non signataires, et en déduire que l'effet de la visite chez le MGT signataire du CAPI identifié en 2^e période représente donc l'effet de l'incitatif uniquement. En revanche, durant la période de mise en place du CAPI et

²³ En effet, il aurait été nécessaire de construire la variable de recours au MGT comme prenant la valeur 1 si la femme avait recouru au MGT avant d'effectuer une mammographie, et 0 sinon. Dans ce cas, l'ensemble des femmes ayant recouru au MGT auraient aussi recouru au dépistage, et il aurait été impossible d'identifier l'effet du CAPI.

donc d'exposition aux incitatifs, le fait d'avoir consulté un MGT signataire du CAPI a eu un impact positif mais de faible ampleur, non significatif seuil de 10% ($p=0.0667$). Ce résultat apparaît cohérent au regard des hypothèses formulées et du contexte de mise en place du CAPI : le faible montant de bonus financier consacré à l'indicateur de dépistage du cancer du sein pourrait expliquer l'impact limité qu'il a eu sur les pratiques. Par ailleurs, ces résultats sont cohérents avec les indicateurs agrégés produits par l'Assurance Maladie à mi-parcours du dispositif, ne montrant pas de différence significative entre signataires et non signataires du CAPI en termes de réalisation des objectifs (Assurance Maladie 2010).

Limites

Les résultats de cette étude doivent néanmoins être nuancés par la présence probable de biais dans l'estimation des paramètres. En effet, ils peuvent être liés à la définition de la condition d'exclusion²⁴ nécessaire à l'estimation du modèle de sélection. Le choix de la variable d'ALD comme variable d'exclusion est justifié d'un point de vue théorique dans la mesure où, en tant que décision d'investissement, le recours à la prévention est a priori indépendant de l'état de santé (Kenkel 1994; Lairson et al 2005). Ainsi, dans une modélisation similaire, Carrieri et Bilger utilisent les variables d'état de santé comme variables d'exclusion, en les incluant uniquement dans l'équation de recours au médecin généraliste et non dans celle de recours au dépistage (Carrieri and Bilger 2013). Toutefois, empiriquement, nous avons obtenu une corrélation négative entre le fait d'avoir une affection de longue durée et la probabilité de recourir à la mammographie. Même si ce lien est d'ampleur limitée comparé à l'effet des autres variables indépendantes, nous ne pouvons exclure l'existence d'un biais de variable omise dans l'estimation du modèle de recours à la mammographie. Par ailleurs, en ce qui concerne l'estimation de l'impact de la variable CAPI plus précisément, il n'est pas possible d'exclure que des chocs exogènes puissent avoir affecté les pratiques des médecins en dehors de la mise en place du CAPI. Toutefois, si de tels chocs existent, nous pensons qu'ils ont eu un impact limité car, à notre connaissance, aucune politique incitative ou de promotion du cancer particulière n'a été mise en place concomitamment avec le dispositif CAPI. Notons par ailleurs que, malgré l'existence de polémiques concernant le dépistage de masse du cancer du sein (Elmore and Fletcher 2012),

²⁴ La condition d'exclusion fait référence au fait d'introduire une variable supplémentaire dans l'équation de recours au médecin traitant par rapport à l'équation de recours à la mammographie.

la médiatisation de ces polémiques est très récente et il est donc peu probable qu'elles aient affecté de façon significative le recours au dépistage sur les périodes considérées.

Par ailleurs, même si les données de l'EGB nous ont permis de définir de façon objective le recours au dépistage du cancer du sein, l'élaboration des codes associés à la mammographie ne permet pas d'exclure la comptabilisation de mammographies liées à des diagnostics plutôt que des dépistages. En effet, le code QEQQ004 associé à la mammographie de dépistage organisé est aussi utilisé pour le dépistage des femmes avec facteurs de risque. Par ailleurs, les mammographies de dépistage individuel partagent le même code CCAM que les mammographies réalisées dans le cadre d'un diagnostic ou d'un suivi. En revanche, le fait d'avoir restreint notre population aux femmes n'ayant pas d'ALD cancer et le fait que plus de 75% des mammographies soient codées QEQQ004 nous conduit à penser que la majeure partie des mammographies comptabilisées sont des mammographies de dépistage. Une dernière limite de l'EGB tient au manque de variables individuelles au niveau patient (statut socioéconomique, couverture complémentaire, état de santé). Or, comme nous l'avons montré dans la section précédente, ces variables constituent des déterminants significatifs du recours au dépistage du cancer du sein.

Conclusion

Le CAPI semble avoir eu un impact limité sur le recours au dépistage du cancer du sein. Ce résultat pourrait témoigner du fait que, dans le domaine particulier de la prévention/dépistage dans lequel l'exercice du médecin se heurte à de multiples freins (inégalités socioéconomiques, absence d'information quant à la réalisation du dépistage, débats actuels à propos de la pertinence des mammographies), les paiements à la performance ne constituent pas un levier suffisamment fort pour engendrer des pratiques de meilleure qualité. En termes de politiques publiques, il semble nécessaire de s'interroger sur les outils complémentaires aux P4P susceptibles d'être implémentés afin d'encourager les médecins généralistes à accroître leur offre de prévention et/ou l'efficacité de leurs pratiques.

Annexes

Annexe A. Tables et graphiques additonnels

Figure A.1. Formules pour calcul des montants de rémunération du CAPI

Encadré 2 : détermination de la rémunération du médecin par le CAPI
<p>A chaque date anniversaire du contrat, le médecin perçoit une rémunération dès lors que la moyenne des taux de réalisation des objectifs de chacun des champs "dépistage et pathologies chroniques" et "optimisation des prescriptions" est supérieure à 25%. Le taux de réalisation de chaque champ est une moyenne pondérée des taux de réalisation de chaque objectif le composant.</p>
<p>Calcul des taux de réalisation (TR)</p>
<p>Le TR de chaque objectif est calculé de manière différente selon que le niveau de l'indicateur constaté chaque année est inférieur ou supérieur à l'objectif intermédiaire, commun à tous les médecins.</p>
<p><i>Cas 1 : le niveau constaté de l'indicateur est inférieur à l'objectif intermédiaire</i></p>
<p>$TR = 50\% * (\text{niveau constaté} - \text{niveau initial}) / (\text{obj. intermédiaire} - \text{niveau initial})$. Le niveau initial est égal au niveau de l'indicateur constaté au trimestre précédent le contrat. Toutefois, s'il est supérieur à un seuil, défini pour tous les médecins (cf. tableau 1), le niveau initial est alors fixé au niveau de ce seuil. Exception : le niveau initial pour l'indicateur HTA est fixé à 30% pour l'ensemble des médecins.</p>
<p>Le taux de réalisation est au plus égal à 50%.</p>
<p><i>Cas 2 : le niveau de l'indicateur est supérieur à l'objectif intermédiaire</i></p>
<p>$TR = 50\% + 50\% * (\text{niveau constaté} - \text{obj. intermédiaire}) / (\text{obj. cible} - \text{obj. intermédiaire})$. Le taux de réalisation est plafonné à 100% même si le niveau de l'indicateur dépasse l'objectif cible.</p>
<p>Le taux de réalisation global est une moyenne pondérée à 60% des taux de réalisation du champ « dépistage et suivi des pathologies chroniques » et à 40% du champ « optimisation des prescriptions ».</p>
<p>Montant de la rémunération</p>
<p>Si la condition de 25% est réalisée, la rémunération est égale à 7 € par patient, multipliés par le taux de réalisation global. La rémunération est versée dans un délai de quatre mois suivant la date anniversaire du contrat.</p>

Encadré extrait du rapport de la commission des Comptes de Sécurité Sociale (Commission des comptes 2011).

Figure A.2. Classification des individus en différents profils d'alcoolisation selon l'AUDIT-C validé (extrait de (Com-Ruelle et al 2005))

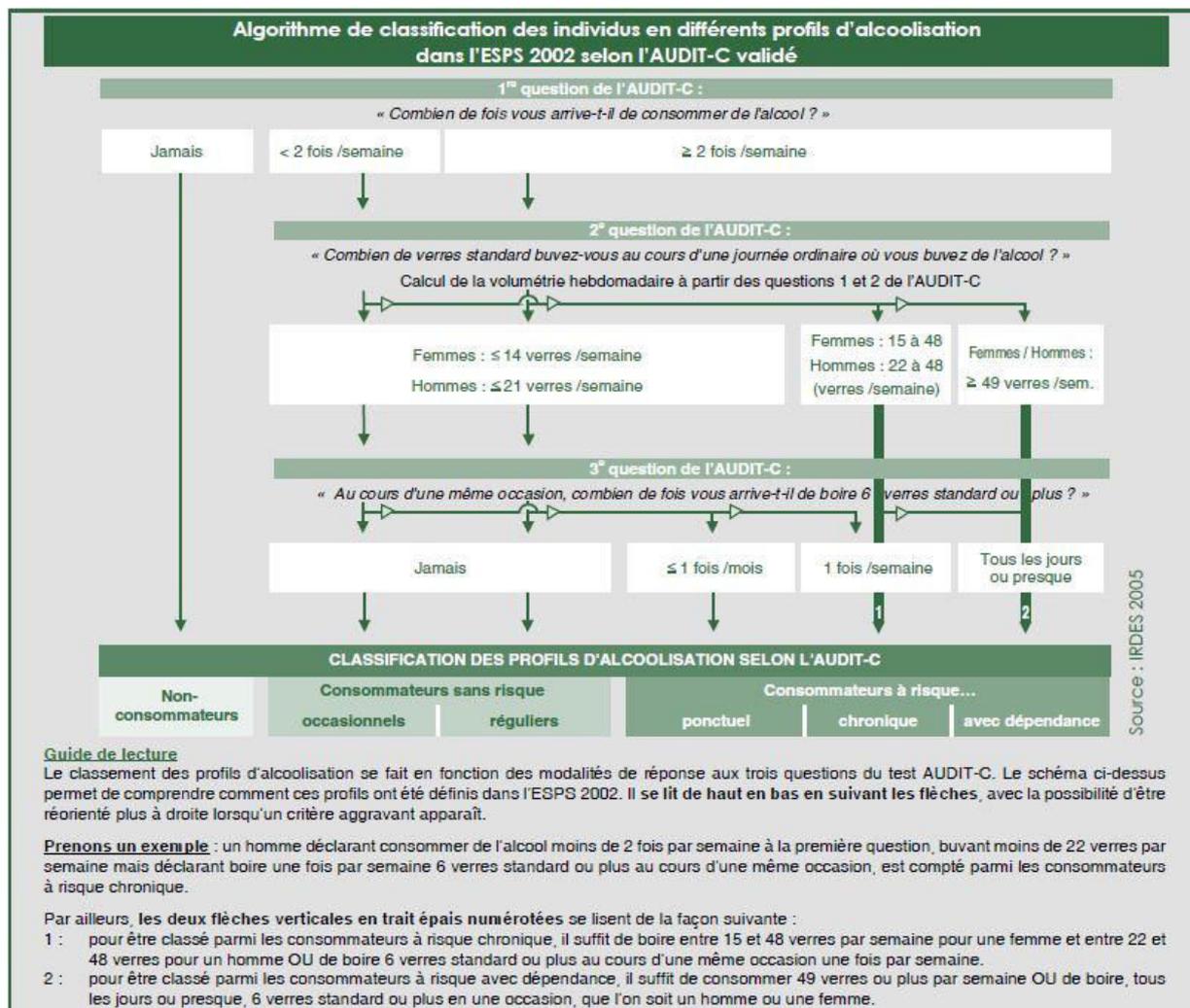


Figure A.3. Architecture des données du SNIR-AM

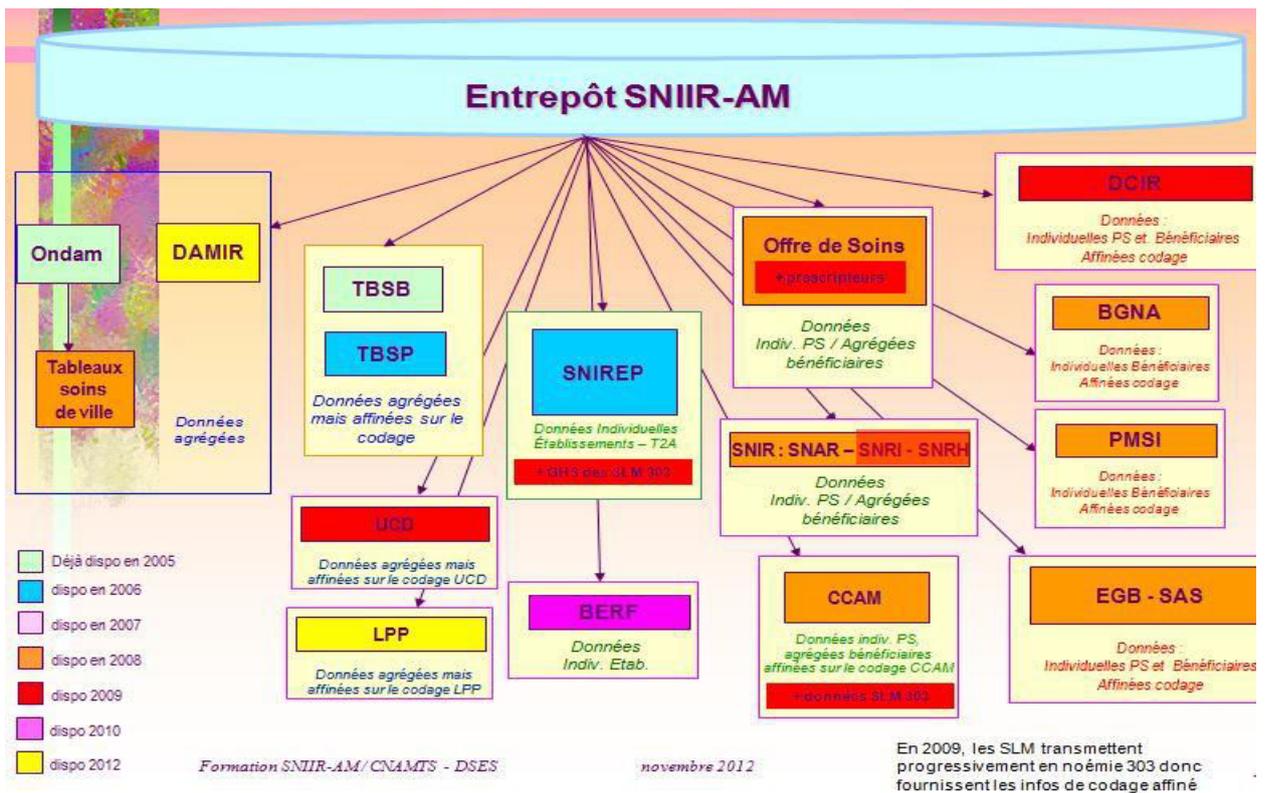


Figure A.4. Architecture des données de l'EGB

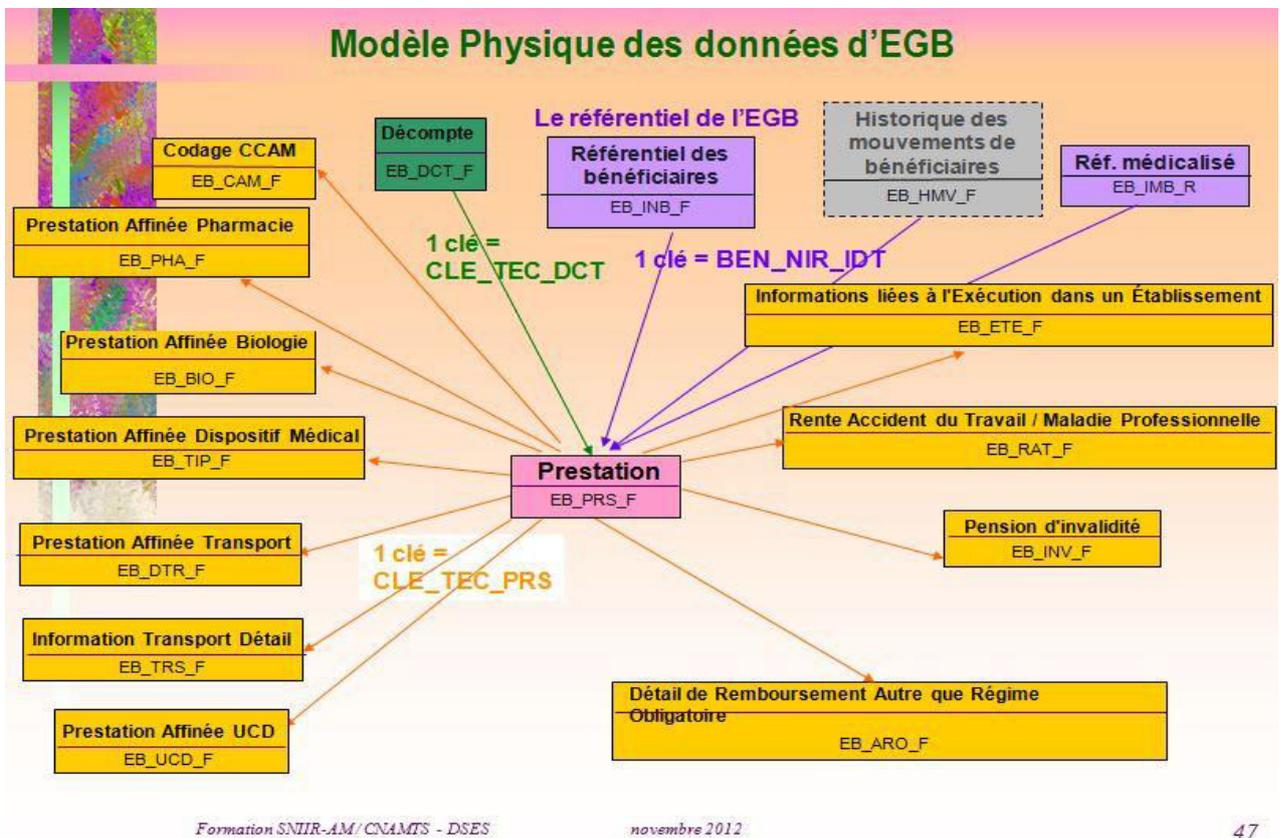


Tableau A.1. Résultats du modèle de recours au médecin généraliste traitant

	Période pre-CAPI			Période post-CAPI		
	Estimate	t-value	p-value	Estimate	t-value	p-value
Constante	1.2382	48.53	<.0001	1.3251	52.32	<.0001
Age	<i>ref = 65-74 ans</i>			<i>ref = 65-74 ans</i>		
50-54	-0.3591	-20.16	<.0001	-0.3438	-19.36	<.0001
55-59	-0.1232	-6.15	<.0001	-0.1081	-5.41	<.0001
60-64	-0.0819	-3.78	0.0002	-0.0438	-2.1	0.0354
Région	<i>ref = Rhones-Alpes</i>			<i>ref = Rhones-Alpes</i>		
Alsace	0.2178	4.62	<.0001	0.1547	3.34	0.0008
Aquitaine	0.0772	2.02	0.0429	0.0356	0.96	0.3391
Autre (ex: Outre-mer)	-0.2838	-6.88	<.0001	-0.3886	-10.24	<.0001
Auvergne	0.1338	2.61	0.0092	0.1143	2.23	0.0255
Basse Normandie	0.1135	2.32	0.0205	0.0732	1.51	0.1308
Bourgogne	0.0704	1.52	0.128	0.0381	0.83	0.4045
Bretagne	0.1229	3.17	0.0015	0.0638	1.68	0.0924
Centre	0.0930	2.27	0.0235	0.1052	2.57	0.0103
Champagne-Ardenne	0.3018	5.21	<.0001	0.1730	3.15	0.0016
Franche-Comté	0.0987	1.79	0.0731	0.1250	2.26	0.0241
Hte Normandie	0.2088	4.51	<.0001	0.1972	4.26	<.0001
Ile de France (hors Paris)	-0.1740	-6.36	<.0001	-0.2098	-7.82	<.0001
Inconnu	-0.2527	-5.41	<.0001	-0.2501	-4.94	<.0001
Languedoc-Roussillon	-0.0146	-0.37	0.7122	-0.0098	-0.25	0.7994
Limousin	0.1706	2.45	0.0145	0.0624	0.92	0.3563
Lorraine	0.2420	5.52	<.0001	0.2410	5.45	<.0001
Midi-Pyrénées	0.0637	1.61	0.107	0.1486	3.72	0.0002
Nord-Pas-De-Calais	0.3129	8.43	<.0001	0.2479	6.75	<.0001
PACA	-0.0244	-0.78	0.435	-0.0440	-1.44	0.1512
Paris	-0.6447	-16.98	<.0001	-0.6377	-17.44	<.0001
Pays de la Loire	0.1629	4.35	<.0001	0.1445	3.92	<.0001
Picardie	0.1737	3.77	0.0002	0.1699	3.71	0.0002
Poitou-Charente	0.1151	2.42	0.0156	0.1633	3.37	0.0008
CMU-C	<i>ref = non</i>			<i>ref = non</i>		
Oui	0.0085	0.31	0.7575	-0.0034	-0.13	0.8986
ALD	<i>ref = non</i>			<i>ref = non</i>		
Oui	0.2481	12.23	<.0001	0.3365	16.2	<.0001
Visite dentiste ou ophtalmologue*	<i>ref = non</i>			<i>ref = non</i>		
Oui	0.1921	12.7	<.0001	0.0891	6.21	<.0001

Chapitre 2

**Préférences des médecins généralistes pour
des dispositifs d'amélioration de leur
pratique : application au dépistage des cancers**

1. Introduction

1.1. Problématique et objectifs

Du fait de l'organisation de l'offre de soins et *a fortiori* depuis l'instauration du parcours de soins (2004), le médecin généraliste (MG) est un acteur clé dans la fourniture ciblée de dépistages et plus globalement dans la prévention et la prise en charge des facteurs de risque. Ce rôle du MG a été encore récemment souligné dans l'enquête barométrique INCa/BVA (Bungener et al 2010). Cependant, leur implication dans le dépistage des cancers reste très hétérogène, entre médecins eux-mêmes mais aussi selon le type de cancer. Les auteurs de l'enquête font alors l'hypothèse que les médecins n'ont pas les mêmes modes d'organisation, motivations et/ou habitudes de soins. Le constat de cette hétérogénéité pose la question de la nature des leviers susceptibles de favoriser les pratiques de dépistage du MG. En France, le dispositif de paiement à la performance (P4P) introduit au travers du CAPI semble avoir eu un effet positif mais très modeste sur le dépistage du cancer du sein (Chapitre 1). Ce résultat pourrait s'expliquer par le faible niveau de l'incitation financière mais aussi par le fait que, la prévention constituant un investissement de moyen et long terme, l'incitation économique pourrait ne pas être le dispositif le plus adéquat. En effet, les MGs pourraient être sensibles à d'autres types de dispositifs plus ou moins innovants, permettant d'accroître leur offre de dépistage des cancers.

L'objectif de notre étude est d'interroger le point de vue des médecins sur leurs préférences concernant les dispositifs ou mesures susceptibles d'accroître leur offre de dépistage et/ou l'efficacité de leur démarche. Ce type d'étude expérimentale met le MG « en situation », de manière à lui faire révéler les déterminants structurants de sa fonction d'offre de dépistage. Cette approche est de plus en plus mobilisée en économie de la santé particulièrement du côté des patients pour leur faire révéler leurs préférences et ou leur propension à payer pour un professionnel de santé, un traitement, un parcours de soins (Carrère et al 2000; Krucien et al 2014; Mühlbacher and Bethge 2014). Côté offre, peu de travaux ont été menés en France. Toutefois, utilisant la méthode de choix discrets (Discrete Choice Experiment, DCE) Ammi et Peyron (2010) ont montré l'importance des incitatifs non financiers dans la pratique de prévention des MGs. En adoptant une approche comparable, Berchi et al. (2006) ont montré qu'une amélioration de la sensibilité du test de dépistage combinée à un paiement ciblé permettrait de favoriser l'offre de dépistage du cancer

colorectal par les MGs français (Berchi et al 2006). Toutefois, à notre connaissance, aucune autre analyse n'a été menée dans le cadre précis du dépistage de plusieurs cancers et proposant un plus large éventail de dispositifs incitatifs possibles.

Nous nous intéresserons donc à trois cancers : le cancer du sein, du col de l'utérus et du côlon pour lesquels, d'une part, l'efficacité du dépistage est reconnue et, d'autre part, des recommandations nationales et internationales de bonne pratique sont disponibles.

Notre hypothèse de travail est la suivante : outre les incitatifs économiques, les médecins généralistes sont sensibles à d'autres types de dispositifs, non monétaires, permettant d'améliorer leur pratique de dépistage des cancers.

Dans la dynamique du plan cancer 2009-2013, de nombreuses et diverses initiatives sont et ont été conduites par les opérateurs locaux (associations de patients, ARS) afin de favoriser l'offre ou l'accès au dépistage des cancers. Toutefois, aucune étude n'a explicitement tenté de comprendre les interactions (complémentarité, substituabilité, redondance) entre différents dispositifs qui parfois, dans la réalité, se superposent. En ce sens, un apport majeur des méthodes de révélation de préférence est d'interroger les individus sur des dispositifs et des combinaisons de dispositifs qui n'existent pas forcément « sur le marché ». Ainsi, le premier objectif de notre travail consiste à révéler les dispositifs ou combinaisons de dispositifs les plus plébiscité(e)s par les médecins. Le deuxième objectif est de comparer le classement des dispositifs les plus valorisés selon les cancers parmi les trois étudiés. En effet, outre les différences liées aux déterminants de la réalisation d'une mammographie, d'un frottis et/ou d'un test hemocult étudiés dans le Chapitre 1, le dépistage des cancers du sein, du col de l'utérus et du côlon s'inscrivent dans des contextes différents : les deux premiers font l'objet d'un paiement à la performance, le premier et le dernier font l'objet d'un dépistage organisé. Ces différents contextes peuvent induire des préférences distinctes quant aux dispositifs à mettre en œuvre ou à renforcer. Le troisième objectif de notre étude est d'analyser l'hétérogénéité des préférences afin d'identifier des profils de médecins plus ou moins sensibles à certains types de dispositifs. Dans l'ensemble, les résultats de l'étude permettront de mieux définir les mesures à prioriser et/ou à associer aux schémas existants (dépistages organisés, P4P) selon les cancers et les objectifs ciblés afin de favoriser de façon plus efficiente l'offre de dépistage des cancers des MGs.

1.2. Revue de la littérature

Afin d'identifier les dispositifs à proposer aux médecins, nous nous sommes basés sur plusieurs revues de la littérature portant 1) sur les expériences étrangères d'évaluation de dispositifs incitatifs visant à l'amélioration des pratiques de prévention/dépistage et 2) sur le contexte français de dépistage des cancers du sein, du col de l'utérus et du côlon. Les dispositifs finalement retenus sont présentés dans la section suivante.

1.2.1. Expériences étrangères

Dans une revue de la littérature sur les interventions effectuées pour modifier les comportements des offreurs de soin, Grimshaw *et al.* (2001) ont montré que des interventions actives telles que des rappels ou des formations étaient efficaces pour changer les comportements, tandis que les interventions moins actives telles que participer à des conférences, lire les journaux médicaux ou envoyer par mail des recommandations médicales, n'étaient pas efficaces. Dans une autre revue de la littérature, les auteurs se sont basés sur 41 revues systématiques de la littérature d'interventions pour le dépistage des cancers du sein et du col de l'utérus et les comportements de prévention (alcool, tabac, alimentation) (Ellis *et al.* 2005). Selon eux, les interventions les plus efficaces incluaient l'utilisation de systèmes de rappel (« reminders and prompts ») et combinaient des incitatifs côté offre à destination des médecins et côté demande à destination des patients.

Dans une revue de la littérature plus récente, Sabatino *et al.* (2012) se sont intéressés plus précisément à l'efficacité d'interventions mises en œuvre pour augmenter le suivi des recommandations et la délivrance de dépistage des cancers du sein, du col de l'utérus et du côlon. A partir d'une dizaine d'études, les auteurs ont mis en évidence des résultats plus ou moins différents des précédents : les interventions d'évaluation et de retours sur les pratiques (« assessment and feedbacks ») ainsi que les formations pouvaient accroître le dépistage de ces cancers. En revanche, les études concernant l'impact des dispositifs financiers étaient trop limitées et de trop faible qualité pour pouvoir en tirer des conclusions robustes (Town *et al.* 2005b).

Les résultats obtenus par ces différentes revues de la littérature, bien que différents, ne sont pas forcément contradictoires. En effet, il est possible que certains dispositifs soient complémentaires ou substituables et donc que leur valorisation diffère selon qu'ils sont combinés ou non. Notons par ailleurs que le contexte dans lequel s'inscrit les différents

dépistages est susceptible de modifier la perception de l'utilité d'un dispositif. De fait, il est indispensable d'analyser plus précisément le contexte français de dépistage des cancers du sein, du col de l'utérus, et du côlon.

1.2.2. Contexte national

En France, un programme de dépistage organisé (DO) a été mis en place depuis 2004 pour le cancer du sein, coordonné par les structures de gestion au niveau départemental. Les femmes sont invitées par les structures de gestion qui, à la suite du dépistage, sont censées informer le médecin choisi par la femme dépistée (médecin traitant ou gynécologue) à la fois de la réalisation de la mammographie et des résultats. Toutefois, le circuit du dépistage semble encore souvent échapper au médecin traitant, puisque la femme est libre de choisir le médecin auquel sont envoyés les résultats. Le manque d'**information des MGs sur le fait que leurs patientes éligibles aient réalisé ou non une mammographie** constitue clairement une limite et un frein à leur engagement dans le dépistage de ce cancer (Liberalotto 2012). Ce frein n'est pas propre au dépistage du cancer du sein et peut aussi concerner notamment le dépistage du cancer du col de l'utérus. Parmi ses recommandations à l'évolution du DO, la HAS recommande la transmission systématique des résultats de la mammographie de DO au médecin traitant et au gynécologue²⁵.

D'après l'enquête BVA/INCa (Bungener et al 2010), 45% des MGs déclarent vérifier systématiquement auprès de leurs patientes concernées la réalisation d'un dépistage du cancer du col de l'utérus. Interrogés sur les raisons pour lesquelles ils réalisent le frottis, 52% répondent que « *réaliser cet acte fait partie de [leur] rôle* ». Depuis les années 90, quatre départements ont mis en place et maintenu des initiatives locales de dépistage organisé du cancer du col de l'utérus et ont été rejoint par 9 autres départements en 2009. Dans sa thèse, Liberalotto (2012) s'intéresse à l'impact de la mise en place de dépistage organisé dans certains départements. Elle montre que l'efficacité du dépistage organisé expérimental sur l'attitude des MGs est variable selon les départements pilotes. Par ailleurs, elle cite deux principaux dispositifs mis en place par les associations locales de dépistage de cancer (l'association Eve dans le Bas-Rhin) ou par les programmes départementaux de dépistage organisé: les **formations pratiques à la réalisation du dépistage** (frottis) destinées aux MGs et la mise à disposition de **dépliants informatifs sur le dépistage à destination des patients**.

²⁵ « La participation au dépistage du cancer du sein des femmes de 50 à 74 ans en France. Situation actuelle et perspectives d'évolution ». Haute Autorité de Santé. Synthèse et recommandations. Novembre 2011.

En France, des expérimentations de dépistage organisé (DO) du cancer colorectal ont été initié par les pouvoirs publics de 2002 et 2007 dans 23 départements pilotes. Le programme a été généralisé à l'ensemble du territoire français à partir de 2009 ; il est géré au niveau de chaque département par des structures de gestion qui se réfèrent à un cahier des charges national. Le médecin traitant occupe un rôle déterminant dans l'adhésion du patient au programme de DO : il évalue les situations d'exclusion en fonction des facteurs de risque du patient, et remet le test en expliquant ses modalités de réalisation et les conséquences en cas de positivité. Toutefois, d'après l'enquête INCa/BVA, les médecins déclarent être les initiateurs d'une discussion sur le dépistage du cancer colorectal dans moins de la moitié des cas (52%). Par ailleurs, ils déclarent ne le vérifier systématiquement que dans 34% des cas : ce dépistage n'apparaît donc pas encore bien inscrit dans les pratiques. Une des pistes d'explication pourrait être liée au temps nécessaire à sa proposition en consultation et à l'explication de la réalisation du test. En effet, 27% des généralistes disent que l'explication du test prend trop de temps (Bungener et al 2010). De fait, nous formulons l'hypothèse que les MGs pourraient être sensibles à **une aide en personnel compétent**, ou à une **rémunération additionnelle** permettant de compenser l'effort fourni.

2. Construction du questionnaire

Pour faire révéler aux MGs leurs préférences quant aux dispositifs susceptibles de modifier leurs pratiques de dépistage des cancers, nous avons construit un questionnaire de choix discret. La construction d'un tel questionnaire suit un processus rigoureux en 4 étapes :

- 1- La définition précise des dispositifs et de leurs modalités.
- 2- L'identification d'une méthode de révélation des préférences.
- 3- La sélection d'un design expérimental pour la construction des épreuves de choix.
- 4- Le choix d'une population d'étude et d'un mode de collecte des données.

La phase de collecte des données a été précédée par une étude pilote (auprès d'une quinzaine de MGs) permettant de tester le design du questionnaire, la compréhension des dispositifs et des modalités de réponses, la difficulté globale et la durée de réponse au questionnaire.

2.1. Définition des dispositifs et de leurs niveaux

2.1.1. Entretiens avec les experts

Afin d'identifier les dispositifs pertinents dans le contexte français, nous avons pris contact avec différents « experts » et responsables en charge du dépistage local. Nous avons participé à une première réunion le 26 Novembre 2013 à l'occasion de la rencontre entre l'ARS Ile de France et les associations départementales de dépistage des cancers sur le thème : « Recours aux dépistages des cancers dans le Val de Marne : état des lieux et préconisations ». Cette réunion a fait émerger un certain nombre de barrières : côté demande, les problématiques semblent essentiellement liées au manque d'information des patients concernant les différents dépistages, côté offre, deux barrières sont mises en évidence le manque de formation de certains médecins et les problèmes d'accès à l'information sur la réalisation ou non d'un dépistage par leurs patient(e)s éligibles. Parmi les initiatives mises en place dans le Val de Marne, des formations ainsi que de l'aide en personnel ont été proposées, mais aucune évaluation précise de ces dispositifs n'a été mise en œuvre.

Nous avons participé à une deuxième réunion le 16 Janvier 2014 au département de médecine générale de l'Université Paris 12 (UPEC) sur le thème : « Les programmes de formation pour le dépistage du cancer colorectal ». Cette réunion a permis de nourrir la réflexion sur le contenu précis des formations à proposer aux médecins pour le dépistage du cancer colorectal en particulier, et applicables en partie au dépistage des cancers féminins.

A la suite de ces réunions d'entretiens avec des experts de terrain et au regard des résultats de la revue de littérature, cinq dispositifs ont été choisis.

2.1.2. Présentation des dispositifs retenus

Les cinq dispositifs retenus sont les suivants :

- (1) Mise à disposition de documents informatifs sur le dépistage et le cancer à destination des patients ;
- (2) Financement d'une formation complémentaire pour le médecin ;
- (3) Accès pour le MG traitant aux résultats des tests qu'il en soit ou non le prescripteur ;
- (4) Aide en personnel présent ou non au cabinet ;
- (5) Offre de rémunération additionnelle.

Nous avons sollicité l'ARS Ile de France afin d'évaluer la pertinence et faisabilité de chacun de ces 5 dispositifs ainsi que leur contenu. Le détail de ces dispositifs pour le dépistage du cancer du sein est présenté dans le tableau 12. La description des dispositifs est similaire pour les cancers du col de l'utérus et du côlon. Les seules modifications concernent la terminologie employée pour décrire le dépistage (frottis ou test hemocult à la place de mammographie), la tranche d'âge pour le cancer du col de l'utérus (25-65 ans), et l'objectif cible pour percevoir le paiement à la performance pour le cancer colorectal (50%²⁶ au lieu de 80%).

Les trois premiers dispositifs ont pour objectif de promouvoir l'accès à l'information des médecins généralistes, le sentiment d'autonomie et de compétence rencontré dans sa relation avec le patient, permettant in fine d'augmenter ses motivations et son implication dans le dépistage. En sus, ces mesures complémentaires et/ou alternatives à des incitatifs financiers peuvent améliorer la transmission d'information vers le patient, améliorant éventuellement l'acceptabilité du dépistage par ce dernier et son observance. Le quatrième dispositif est pensé comme un support « matériel » aux médecins qui, par exemple, pourraient manquer de temps dans leur pratique quotidienne et/ou nécessiteraient une aide en personnel ou technique. Le dernier dispositif est un dispositif financier : dans la lignée des paiements à la performance, il propose différents niveaux de rémunération additionnelle sous forme d'une augmentation en pourcentage des honoraires de l'année. Nous avons veillé à choisir des modalités suffisamment différentes pour être sûrs que les médecins perçoivent la différence entre les niveaux proposés.

²⁶ Le taux de 50% correspond à l'objectif cible proposé par la Haute Autorité de Santé, permettant de réduire la mortalité par cancer de 15 à 20% : « Dépistage et prévention du cancer colorectal. Actualisation du référentiel de pratique de l'examen périodique de santé. HAS, Juin 2013 ».

Tableau 12. Dispositifs retenus (exemple dans le cadre du dépistage du cancer du sein)

Dispositifs	Modalités
1. Mise à disposition , au cabinet, de documents informatifs à destination de vos patientes sur le cancer du sein et sur la mammographie.	Oui/Non
2. Financement d'une formation sur le dépistage du cancer du sein pouvant inclure les points suivants : -Analyse des difficultés à la réalisation du dépistage, stratégies de conviction/application, coordination avec les autres spécialistes. -Bénéfices/risques de la mammographie, degré de traitements et qualité de vie en fonction du pronostic et du stade de détection du cancer.	Oui/Non
3. Résumé trimestriel (listing) de vos patientes de 50-74 ans ayant effectué une mammographie dans les 6 derniers mois.	Oui/Non
4. Aide ponctuelle en personnel qualifié pour le suivi du dépistage de vos patientes, pouvant prendre (au choix) les modalités suivantes : - Présence de personnel au cabinet une ½ journée toutes les trois semaines pour l'accueil/conseil des patientes dépistées. - Mise en relation avec une plateforme téléphonique pour relancer ou contacter les patientes non dépistées.	Oui/Non
5. Rémunération additionnelle sur objectif de santé publique conditionnée à l'atteinte et/ou l'amélioration d'un objectif cible de 80% de vos patientes âgées entre 50 et 74 ans ayant effectué une mammographie dans les 2 dernières années.	Non Augmentation de 1% du total de vos honoraires de l'année Augmentation de 3% du total de vos honoraires de l'année Augmentation de 5% du total de vos honoraires de l'année

2.1.3. Faisabilité des dispositifs

Les dispositifs (1) et (2) ne soulèvent pas de questions particulières quant à leur mise en oeuvre. En effet, l'INCa met déjà à disposition des MGs des brochures d'information à destination des patients concernant les dépistages organisés des cancers du sein et du côlon. Les MGs peuvent commander ces brochures sur le site internet de l'INCa²⁷. En ce qui concerne les formations, celles-ci sont généralement organisées par l'ordre des médecins ou

²⁷ Pour le dépistage du cancer du sein par exemple, ils peuvent le faire à l'adresse : <http://www.e-cancer.fr/publications/50-depistage-du-cancer-du-sein/706-brochure-dinformation-complete-sur-le-depistage-organise-du-cancer-du-sein>.

les sociétés savantes. Notons toutefois que le dispositif proposé prévoit le financement de telles formations.

Le dispositif (3) pose davantage de questions en termes de mise en œuvre. En ce qui concerne les dépistages des cancers du sein et du col de l'utérus, les mammographies ainsi que les frottis cervico-utérins font l'objet d'une prescription et l'information remonte donc dans les fichiers de l'Assurance Maladie (AM) et plus précisément du SNIIR-AM. La faisabilité du dispositif serait donc conditionnée à ce que l'AM accepte de transmettre aux médecins traitants le nom de leurs patient(e)s ayant effectué le dépistage. En ce qui concerne le dépistage du cancer colorectal, celui-ci est géré par les structures de gestion au niveau départemental. Aussi, l'AM n'a pas l'information concernant les tests hemocults effectués dans le cadre du dépistage organisé, mais elle connaît le nom du médecin traitant. En revanche, les structures de gestion n'ont pas le nom du médecin traitant des individus dépistés. La faisabilité du dispositif pour ce dépistage serait donc conditionnée à ce, par exemple, l'AM accepte de transmettre aux structures de gestion un fichier anonymisé des médecins traitants permettant d'effectuer une jointure entre les identifiants des patients dépistés et leur médecin traitant ou, au contraire, que les structures de gestion fournissent à l'AM le fichier des patients dépistés. Une fois le fichier constitué, les listings seraient envoyés aux médecins traitants directement par les structures de gestion ou par l'AM.

La mise en oeuvre du dispositif (4) ne semble pas poser de problèmes particulier : de telles initiatives ont déjà été mises en place par des associations au niveau départemental (Liberalotto 2012). Sa généralisation à une plus grande échelle pour les médecins souhaitant disposer d'une aide en personnel (physique ou virtuelle) serait principalement conditionnée à la formation de ces personnes.

Le dispositif (5) appliqué au dépistage des cancers du sein et du col de l'utérus est un prolongement des dispositifs de P4P existants et ne présente donc par de difficulté particulière en termes de mise en œuvre. Toutefois, en ce qui concerne le dépistage du cancer colorectal, la faisabilité serait conditionnée à ce que les structures de gestion en charge du dépistage organisé acceptent de transmettre à l'Assurance Maladie le nom des patients ayant effectué le dépistage. Ensuite, l'AM pourrait effectuer l'appariement avec le numéro du médecin traitant afin de calculer le niveau de l'indicateur et en déduire le montant du bonus financier correspondant.

2.2. Méthodes de révélation des préférences

Nous désignons par *attributs* les dispositifs proposés aux médecins, par *niveaux* les différentes modalités de ces attributs, par *scénario* toute association des attributs et de leurs niveaux selon un design expérimental, et par *épreuve de choix* les modalités de comparaison des attributs ou des scénarios entre eux.

2.2.1. Synthèses des méthodes de révélation des préférences

La méthode des choix discrets (DCE)

Différentes méthodes ont été développées pour révéler les préférences des individus dans le domaine de la santé. La méthode des choix discrets (Discrete choice Experiment, *DCE*) a connu un fort développement ces 10 dernières années et est très plébiscitée dans le domaine de la santé (Ryan and Gerard 2003; Berchi and Launoy 2007; Ammi and Peyron 2010; de Bekker-Grob et al 2012). Cette méthode consiste à présenter à chaque individu, différents scénarios hypothétiques, où chaque scénario est décrit par chacun des attributs auquel est associé un niveau particulier. Chaque individu doit ensuite choisir le scénario qu'il préfère. La répétition de ces « épreuves de choix » permet *in fine* d'observer les arbitrages que les individus effectuent, en situation de choix, entre les attributs des scénarios qui leurs sont présentés.

La méthode des choix discrets repose sur un certain nombre d'hypothèses dérivées de la théorie économique du consommateur (Lancaster K 1966). Elle permet donc d'estimer des utilités associées à un bien (un scénario dans notre cas), l'impact de chaque attribut sur l'utilité associée à un bien (scénario) ou encore des taux marginaux de substitution entre attributs (Berchi and Launoy 2007). Pour cela, le cadre théorique généralement utilisé est celui de la Théorie de l'Utilité Aléatoire (*Random Utility Theory, RUT*) développé par Thurstone (1927) et reposant sur trois principaux axiomes : (1) l'individu cherche à maximiser son utilité (ou sa satisfaction) et choisit donc le scénario assurant le niveau d'utilité le plus élevé ; (2) le chercheur n'est pas en mesure de mesurer l'utilité réelle procurée par le scénario sélectionné, ce qui nécessite la mise en œuvre d'une approche probabiliste du comportement de choix; (3) l'utilité réside dans les caractéristiques du scénario c'est-à-dire dans ses attributs. Le nombre nécessaire d'épreuves de choix ainsi que le contenu des scénarios sont généralement définis à partir de méthodes expérimentales, permettant d'identifier l'impact de chaque attribut de façon indépendante et efficiente (Street and

Burgess 2007). Le nombre de scénarios présentés dans les épreuves de choix dépend de la situation décisionnelle à présenter et de sa difficulté cognitive. Pour une décision médicale comme un traitement, il s'agit le plus souvent d'un scénario unique (ex : « si on vous proposait ce traitement, accepteriez-vous de le prendre ? » : oui/non) ou d'une paire de scénarios (ex : « préférez-vous le traitement A ou le traitement B ? ») (Krucien 2012).

Les méthodes de Best-Worst Scaling (BWS)

Récemment, d'autres méthodes de préférences révélées ont été développées. Les méthodes de classement du meilleur au pire (Best-Worst Scaling, **BWS**) ont pour objectif de révéler davantage d'information par épreuve et de mesurer de façon plus précise l'impact respectif de chaque attribut (Flynn et al 2007; Marley et al 2008). Ces méthodes ont en commun avec la DCE de présenter aux individus différentes épreuves de choix, mais diffèrent dans la conception de ces épreuves. Dans le cas de BWS, il est demandé aux individus de sélectionner l'attribut ou le scénario qu'ils préfèrent le plus (« best ») et celui qu'ils préfèrent le moins (« worst »). Il existe trois types de méthodes de classement du meilleur au pire (Lancsar et al 2013), dont seulement deux peuvent s'appliquer à notre contexte d'étude : la méthode de classement des attributs du meilleur au pire (Best-Worst Attribute Scaling, **BWAS**) d'une part, et la méthode de classement des scénarios du meilleur au pire (Best-Worst Discrete Choice Experiment, **BWDCE**) d'autre part²⁸.

La méthode BWAS consiste à présenter les scénarios un à un aux individus, chaque scénario étant décrit par chacun des attributs auquel est associé un niveau particulier. Au lieu d'évaluer et de comparer l'utilité de différents scénarios, le répondant évalue les utilités de chaque niveau d'attribut proposé en choisissant pour chaque scénario, l'attribut avec le niveau qui lui procure la plus grande utilité (« meilleur » ou « préféré ») et celui qui lui procure la plus faible utilité (« pire » ou « moins préféré »). De fait, la référence est un seul niveau d'attribut et non un scénario tout entier. Cette approche permet d'estimer l'effet des niveaux d'attributs sur une même échelle et donc l'utilité moyenne de l'attribut sur tous ses niveaux, soit l'impact de l'attribut à proprement parler (Flynn et al 2007). L'inconvénient de cette méthode par rapport à la DCE est que, dans la mesure où le répondant ne donne pas

²⁸ La méthode de classement du meilleur au pire omise à la méthode dite de type 1 ou « Best-Worst Object Scaling » (Lancsar et al 2013) qui permet de comparer les caractéristiques d'un bien entre elles (exemple : les caractéristiques d'un bien alimentaire (Finn and Louviere 1992) ou encore les caractéristiques de la relation médecin-patient (Krucien 2012).

d'information sur ses préférences pour le scénario dans son ensemble et donc pour des combinaisons de dispositifs, elle ne permet pas d'estimer des interactions entre dispositifs.

La méthode BWDCE est la méthode « alternative » la plus proche la DCE mais ayant le format de la BWS, permettant ainsi de collecter davantage d'information. En effet, elle consiste à présenter aux répondants plusieurs scénarios (au moins 3), pour lesquels l'individu choisit le scénario qu'il préfère et celui qu'il préfère le moins (Lancsar et al 2013). Il y a encore peu d'applications de la méthode BWDCE en santé, les propriétés mathématiques de cette méthode n'ont été établies que récemment (Marley and Islam 2012).

2.2.2. Choix d'une méthode

A priori, il ne semble pas y avoir de méthode qui « domine » les autres en termes de capacité à répondre à nos questions de recherche. Un critère de choix de méthode crucial est la difficulté cognitive attendue des épreuves, une trop grande difficulté pouvant engendrer des réponses inconsistantes. En effet, les arbitrages demandés aux médecins dans le cadre de notre problématique sont complexes : sélectionner des attributs/scénarios hypothétiques permettant d'améliorer la pratique peut être considéré comme difficile, notamment lorsqu'il existe déjà des interventions pour améliorer les pratiques. Cognitivement, choisir l'attribut préféré et moins préféré au sein d'un même scénario (épreuve BWAS) apparaît plus « facile » que choisir la combinaison d'attributs préférée et moins préférée parmi plusieurs scénarios distincts (épreuve BWDCE). Ces considérations nous conduisent à privilégier, parmi les méthodes de classement du meilleur au pire, la méthode BWAS par rapport à la méthode BWDCE.

Une solution intermédiaire proposée par certains auteurs (Flynn et al 2007) a été explorée dans le but de combiner les avantages des méthodes DCE et BWAS. Appliquée à notre contexte, cette solution consiste à interroger les médecins sur les attributs qu'ils préfèrent le plus et le moins pour chaque scénario/situation (épreuve BWAS), puis à leur demander s'ils « accepteraient » de changer le contexte de leur pratique de dépistage habituelle pour la situation proposée (épreuve DCE à scénario unique). Cette combinaison des deux méthodes a un double intérêt : elle permet (1) d'estimer un modèle de choix discret traditionnel dont les estimations peuvent être comparées à celles du best-worst scaling et (2) d'estimer des interactions entre attributs à partir du modèle de choix discrets.

2.3. Design expérimental

L'objectif de cette étape est d'identifier un design adapté permettant d'identifier les préférences pour les attributs/dispositifs de façon indépendante et éventuellement d'estimer des interactions entre attributs. Les scénarios sont généralement constitués à partir d'un sous-ensemble d'options de choix (plan factoriel fractionnel) au sein de l'ensemble des possibilités (plan factoriel complet). En effet, comme nous avons défini 4 dispositifs comportant 2 niveaux (oui/non) et un dispositif comportant 4 niveaux, le plan factoriel complet comporte $2^4 * 4 = 64$ scénarios et donc épreuves de choix à présenter aux médecins. Ce nombre est évidemment jugé beaucoup trop élevé pour être présenté dans un unique questionnaire. Afin de réduire le nombre de scénarios, nous avons appliqué les méthodes issues de la théorie des designs expérimentaux (Street and Burgess 2007) qui consistent à « créer » des scénarios de façon à équilibrer les fréquences d'apparition de chaque niveau d'attributs dans le design et à estimer les effets des attributs de façon indépendante (orthogonalité). Pour cela, nous avons utilisé le logiciel NGENE qui permet de construire des designs optimaux et efficaces (ChoiceMetrics, 2011). Dans ce but, il était nécessaire de déterminer dans une première étape:

- Le nombre scénarios minimal. Pour avoir un design équilibré, le nombre de scénarios doit être supérieur au nombre d'attributs (i.e. 5) et doit être un multiple du nombre de niveaux de chaque attribut (i.e. de 2 et de 4). La taille minimale du design est donc de 8 scénarios.
- Le nombre de scénarios maximal. Ce nombre a été fixé en fonction de la difficulté estimée du questionnaire et du temps envisagé pour le compléter. Etant donné que les répondants devaient compléter deux types d'épreuves de choix distinctes, et qu'il n'était pas envisageable de répartir les épreuves sur plusieurs versions du questionnaire pour un même cancer pour limiter le nombre de répondants requis, nous avons fixé le nombre maximal de scénarios à 12 (et nous avons confirmé la faisabilité à l'occasion de l'étude pilote).

Afin de choisir entre un design à 8 ou 12 scénarios, nous nous sommes basés sur la capacité du design à estimer des interactions entre dispositifs. Deux interactions étaient particulièrement intéressantes de notre point de vue. La première était l'interaction entre les attributs *documents* et *rémunération*, car nous souhaitions tester l'hypothèse selon laquelle les médecins valorisaient d'autant plus les documents à destination des patients qu'ils étaient accompagnés de rémunération compensant l'effort fourni. La deuxième était l'interaction

entre les attributs *listing* et *aide*, car nous souhaitions tester l’hypothèse selon laquelle le listing était davantage valorisé lorsqu’il était accompagné d’aide en personnel qualifié pour accompagner d’une part les personnes ayant été dépistées et les aiguiller dans le parcours de soins, d’autre part les individus n’ayant pas été dépistés et les sensibiliser. Seul le design à 12 scénarios permettait d’estimer des interactions indépendamment des effets principaux (cf. Tables B.1 et B.2 en Annexe B) et a donc été choisi.

Le design finalement retenu avait un critère de D-efficience de 92.9%. Une description du contenu des 12 scénarios est présentée dans le tableau 13.

Tableau 13. Description du design à 12 scénarios

	Documents	Formation	Listing	Aide	Remunération
Scénario 1	non	non	oui	non	5%
Scénario 2	non	oui	oui	non	0%
Scénario 3	oui	non	non	non	3%
Scénario 4	oui	oui	oui	oui	0%
Scénario 5	oui	non	non	oui	1%
Scénario 6	non	oui	non	oui	5%
Scénario 7	oui	oui	non	non	3%
Scénario 8	non	oui	non	non	1%
Scénario 9	oui	non	oui	non	1%
Scénario 10	non	non	oui	oui	3%
Scénario 11	non	non	non	oui	0%
Scénario 12	oui	oui	oui	oui	5%

Epreuve DCE: Accepteriez-vous de changer votre pratique habituelle pour la situation proposée ?

Epreuve BWAS : Sélectionnez: 1) Ce que vous préférez **le plus** dans la situation proposée,
2) ce que vous préférez **le moins** dans la situation proposée

2.4. Population d’étude et collecte des données

Nous avons fait appel à un institut de sondage pour construire la version web du questionnaire et recueillir les données, avec pour objectif de construire un échantillon représentatif des MGs libéraux français (âge, sexe, et département) et de s’assurer d’un taux de réponse le plus élevé possible. L’objectif cible était d’obtenir 400 questionnaires complétés, soit près de 133 questionnaires pour chaque type de dépistage (sein, col de l’utérus ou côlon)²⁹.

²⁹ Ce nombre peut apparaître faible en soit, mais est suffisant pour les analyses de choix discrets qui menées ultérieurement, les modèles étant estimés sur environ $133 \times 12 = 1593$ données individuelles.

Le questionnaire était constitué de deux parties : dans la partie 1, les médecins devaient compléter 24 épreuves de choix (12 DCE et 12 BWAS). Les figures B.1, B.2 et B.3 (Annexe B) présentent des extraits de la présentation et du déroulement des épreuves DCE et BWAS pour le questionnaire portant sur le dépistage du cancer du sein. Dans la partie 2, les MGs devaient répondre à 25 questions additionnelles permettant de collecter des données sociodémographiques, d'activité et de pratique de dépistage, destinées entre autres à l'analyse de l'hétérogénéité des préférences. Le temps moyen attendu de remplissage du questionnaire était de 15 minutes. De fait, les médecins répondants ont été rémunérés l'équivalent d'une consultation majorée, soit 28€.

Il a été choisi que le mode de passation du questionnaire serait en ligne. En effet, ce mode de collecte confère plusieurs avantages par rapport à un questionnaire papier : il permet de s'assurer de la représentativité des médecins (méthode des quotas) et permet de contrôler les réponses des individus : pour chaque question posée, les médecins étaient obligés de fournir une réponse pour pouvoir passer à la page suivante. De fait, les données sont caractérisées par aucune non-réponse. Par ailleurs, les médecins ayant accepté de participer à l'étude ont été répartis aléatoirement dans une des trois versions du questionnaire, chaque version s'intéressant soit au **dépistage du cancer du sein (version 1)**, au **dépistage du col de l'utérus (version 2)** ou au **dépistage du cancer colorectal (version 3)**. En outre, afin d'éviter tout biais de fatigue pouvant se caractériser par des réponses moins cohérentes pour les scénarios apparaissant à la fin de la partie 1, l'ordre d'apparition des scénarios dans le questionnaire a été randomisé.

Le questionnaire a fait l'objet d'une étude pilote auprès de 15 médecins généralistes qui avaient préalablement accepté de participer à l'étude. Cette phase pré-test a permis d'identifier les questions qui posaient problème aux médecins, ainsi que de collecter leurs avis concernant la formulation ou la présentation des dispositifs. A la suite de cette phase pilote certaines modifications ont été apportées dans la présentation des dispositifs³⁰, dans les catégories de réponses proposées pour les questions additionnelles, ainsi que l'ajout d'une barre de progression afin d'estimer l'avancement dans le questionnaire.

Le remplissage du questionnaire s'est déroulé entre le 25 Mars 2014 et le 18 Avril 2014, date à laquelle le nombre attendu de questionnaires complétés a été atteint. Durant cette

³⁰ Par exemple, concernant le dispositif n°3 « listing », la phrase de description suivante a été ajoutée : « Cette information vous permettra, par exemple, de créer des alertes ou d'actualiser vos données patients ».

période, environ 4000 liens vers des questionnaires en ligne ont été envoyés. Parmi ces 4000 liens, 685 ont été « touchés » et 402 ont aboutis à des questionnaires complétés, correspondant à un taux d'incidence de 58%.

2.5. Description des données de l'enquête

L'échantillon final comprend 402 médecins, 135 pour la version 1 (cancer sein), 133 pour la version 2 (cancer du col de l'utérus) et 134 pour la version 3 (cancer colorectal).

Les statistiques descriptives sur les trois versions du questionnaire ainsi que les tests d'égalités des proportions (tests du χ^2) sont présentés dans le tableau 14. Les trois échantillons ainsi constitués sont représentatifs en termes de genre (environ 72% d'hommes), d'âge (plus de 50% de moins de 54 ans), et de localisation géographique. Ils ont la même structure pour la quasi-totalité des variables excepté la variable de formation sur le dépistage : plus de 60% des médecins ont suivi une telle formation dans les versions 2 et 3, contre 50% dans la version 1 du questionnaire ($p=0.051$).

Dans les trois versions du questionnaire, les médecins sont caractérisés de la façon suivante : la majorité travaille en libéral uniquement (entre 75 et 80%), en cabinet individuel (entre 50 et 55%), suit des cycles de formation continue (entre 87 et 89%), exerce en secteur 1 (90%) et travaille plus de 10 demi-journées par semaine (entre 50 et 57%). Par ailleurs, les médecins travaillent beaucoup : entre 48 et 53% d'entre eux travaillent plus de 50 heures par semaine. En ce qui concerne les variables de patientèle, les médecins voient en majorité entre 100 et 150 patients par semaine, sont généralement médecins traitant de plus de 75% de leurs patients et ont une patientèle davantage féminine. La majorité d'entre eux déclare proposer « systématiquement » ou « très souvent » un dépistage du cancer du sein (environ 72%), du col de l'utérus (entre 56 et 67%) et du côlon (entre 69 et 76%)³¹. En revanche, ils sont moins nombreux à utiliser des brochures d'information à destination des patients dans leur pratique quotidienne (environ 39% pour le dépistage du cancer du sein, entre 31 et 37% pour le dépistage du cancer du col de l'utérus, et entre 46 et 54% pour le dépistage du cancer colorectal). Enfin, ils sont plus nombreux à avoir perçu une prime ROSP pour le dépistage du cancer du sein (42%) que pour le dépistage du cancer du col de l'utérus (26%).

³¹ Ces taux déclarés apparaissent plus élevés que les chiffres des enquêtes récentes effectuées sur le sujet (notamment le baromètre INCA/BVA de 2010), mais ils ne sont pas comparables car cette dernière enquête par exemple interrogeait sur le dépistage systématique, et obtenait 56% pour le cancer du sein, 45% pour le dépistage du col de l'utérus et 34% pour le dépistage du cancer colorectal.

Tableau 14. Statistiques descriptives sur les trois versions du questionnaire

	Cancer du sein		Cancer du col		Cancer colorectal		Chi-deux
	N	%	N	%	N	%	P-value
Nature de l'exercice							0.7707
Libéral uniquement	109	80.7	104	78.2	100	74.6	
Libéral + salarié hospitalier	9	6.7	12	9.0	12	9.0	
Libéral + vacations (médicales ou autres)	17	12.6	17	12.8	22	16.4	
Genre							0.9692
Homme	96	71.1	96	72.2	97	72.4	
Femme	39	28.9	37	27.8	37	27.6	
Age							0.3245
< 45 ans	21	15.6	28	21.1	32	23.9	
[45-54] ans	51	37.8	42	31.6	50	37.3	
>=55 ans	63	46.7	63	47.4	52	38.8	
Zone géographique							0.9713
Ile de France	22	16.3	24	18.1	26	19.4	
Ouest	28	20.7	28	21.1	26	19.4	
Nord	8	5.9	9	6.8	10	7.5	
Est	21	15.6	20	15.0	20	14.9	
Sud Ouest	18	13.3	16	12.0	16	11.9	
Centre Est	10	7.4	17	12.8	12	9.0	
Mediterranee	28	20.7	19	14.3	24	17.9	
Taille de le commune							0.3996
Moins de 10 000 habitants	44	32.6	44	33.1	51	38.1	
De 10 à 100 000 habitants	45	33.3	43	32.3	51	38.1	
Plus de 100 000 habitants	46	34.1	46	34.6	32	23.9	
Année d'installation pour la première fois							0.3294
Moins de 15 ans	32	23.7	32	24.1	37	27.6	
De 16 à 29 ans	62	45.9	59	44.4	69	51.5	
Plus de 30 ans	41	30.4	42	31.6	28	20.9	
Suivez-vous des cycles de formation continue ?							0.8347
Oui	121	89.6	118	88.7	117	87.3	
Avez-vous déjà suivi une formation sur le dépistage des cancers ?							0.0507
Oui	67	49.6	81	60.9	85	63.4	
Vous exercez dans un ...							0.1234
Cabinet individuel	75	55.6	69	51.9	66	49.3	
Cabinet de groupe de médecins	56	41.5	56	42.1	61	45.5	
Autre structure (cabinet pluridisciplinaire, etc)	4	3.0	8	6.0	7	5.2	
Secteur d'exercice							0.9723
Secteur 1	121	89.6	119	89.5	121	90.3	
Pour la prise de rendez-vous vous avez recours à ... (plusieurs réponses possibles)							0.5614
Un standard téléphonique de prise de RDV	28	20.7	37	27.8	31	23.1	
Un site internet (ex ameli.fr)	2	1.5	2	1.5	6	4.5	
Un(e) secrétaire salarié(e)	39	28.9	42	31.6	46	34.3	
Un(e) conjoint(e) « collaborateur »	7	5.2	12	9.0	6	4.5	
Vous prenez vous même vos rendez-vous	71	52.6	62	46.6	59	44.0	
Autre/ pas de prise de rendez-vous	6	4.4	5	3.8	4	3.0	
Combien d'heures travaillez-vous en moyenne par semaine ?							0.8201
<= 40 heures	20	14.8	25	18.8	23	17.2	
[41-50] heures	43	31.9	44	33.1	47	35.1	
[51-60] heures	50	37.0	42	31.6	38	28.4	
>60 heures	22	16.3	22	16.5	26	19.4	
Combien de demies-journées travaillez-vous en moyenne par semaine ?							0.4335
<10 demi-journées	58	43.0	61	45.9	68	50.8	
>=10 demi-journées	77	57.0	72	54.1	66	49.3	

* Chi-deux : tests d'indépendance mesurant l'association entre les réponses aux variables et le type de cancer.

Tableau 14. Statistiques descriptives sur les trois versions du questionnaire (suite)

	Cancer du sein		Cancer du col		Cancer colorectal		Chi-deux
	N	%	N	%	N	%	P-value
Sur une semaine de consultations, combien de patientes vous consultez pour un motif lié à la gynécologie* ?							0.9768
<= 15 patientes	83	61.5	82	61.6	-	-	
> 16 patientes	52	38.5	51	38.4	-	-	
Combien de patients voyez-vous en moyenne au cours d'une semaine normale ?							0.2961
<100 patients	34	25.2	28	21.1	27	20.2	
[100-149] patients	55	40.7	49	36.8	64	47.8	
>=150 patients	46	34.1	56	42.1	43	32.1	
Parmi les patients qui vous consultent au cours d'une semaine normale:							
% dont vous êtes le « médecin traitant » ?							0.1439
[0-49%] de vos patients	14	10.4	14	10.5	20	14.9	
[50-74%] de vos patients	29	21.5	44	33.1	32	23.9	
>=75% de vos patients	92	68.2	75	56.4	82	61.2	
% de femmes ?							0.8406
[0-49%] de vos patients	28	20.7	25	18.8	29	21.6	
[50-100%] de vos patients	107	79.3	108	81.2	105	78.4	
% de patients de 50 à 75 ans ?							0.2057
[0-49%] de vos patients	87	64.4	-	-	96	71.6	
[50-100%] de vos patients	48	35.6	-	-	38	28.4	
% de patientes de 25 à 65 ans ?							-
[0-49%] de vos patients	-	-	86	64.7	-	-	
[50-100%] de vos patients	-	-	47	35.3	-	-	
% de bénéficiaires de la CMUC ?							0.4962
<10 % de vos patients	88	65.2	79	59.4	88	65.7	
>=10% de vos patients	47	34.8	54	40.6	46	34.3	
Proposez-vous un dépistage du cancer du sein à vos patientes âgées entre 50 et 74 ans ?							0.6693
Systématiquement	55	40.7	51	38.4	43	32.1	
Très souvent	41	30.4	46	34.6	55	41.0	
Souvent	32	23.7	28	21.1	28	20.9	
Rarement, Jamais	7	5.2	8	6.0	8	6.0	
Proposez-vous un dépistage du cancer du col de l'utérus à vos patientes âgées entre 25 et 65 ans?							0.4913
Systématiquement	38	28.2	36	27.1	34	25.4	
Très souvent	52	38.5	46	34.6	41	30.6	
Souvent	31	23.0	39	29.3	38	28.4	
Rarement, Jamais	14	10.4	12	9.0	21	15.7	
Proposez-vous un dépistage du cancer colorectal à vos patient(e)s âgé(e)s entre 50 et 74 ans ?							0.8714
Systématiquement	43	31.9	46	34.6	39	29.1	
Très souvent	54	40.0	55	41.4	53	39.6	
Souvent	34	25.2	29	21.8	36	26.9	
Rarement, Jamais	4	3.0	3	2.3	6	4.5	
Dans votre pratique quotidienne, utilisez-vous des brochures d'information pour:							
<i>Le cancer du sein</i>							0.9529
Oui	52	38.5	53	39.9	54	40.3	
<i>Le cancer du col de l'utérus</i>							0.4435
Oui	42	31.1	41	30.8	50	37.3	
<i>Le cancer colorectal</i>							0.4187
Oui	62	45.9	64	48.1	72	53.7	
Avez-vous perçu en 2013 ou en 2012 une prime ROSP pour le dépistage du cancer du sein ?							-
Oui	57	42.2	-	-	-	-	
Avez-vous perçu en 2013 ou en 2012 une prime ROSP pour le dépistage du cancer du col de l'utérus							-
Oui	-	-	34	25.6	-	-	

* Y compris la contraception.

3. Analyse des données de choix discrets (DCE)

Dans cette section, nous présentons dans un premier temps les méthodes statistiques associées à l'analyse des réponses aux épreuves DCE : chaque médecin était confronté à 12 scénarios hypothétiques successifs et pour chacun d'eux, il devait indiquer s'il accepterait de changer le contexte de sa pratique habituelle pour la situation proposée (« oui », « non », « ne sait pas »). Dans un deuxième temps, nous présentons les résultats des différents modèles estimés. Ces derniers sont discutés dans la troisième sous section.

3.1. Méthodes

3.1.1. Des données au modèle économétrique sous-jacent

Notons Y_{ij} la variable d'intérêt représentant le choix effectué par le médecin j pour le scénario i . Cette variable est codée 1 si à la question : « changeriez-vous votre pratique de dépistage habituelle pour la situation proposée ? », le médecin a répondu « oui », et 0 s'il a répondu « non ». De fait, les réponses « ne sait pas » sont considérées comme des non réponses. Le modèle économique susceptible de décrire le comportement étudié repose sur un modèle de choix discret dichotomique, où Y_{ij} est reliée à une variable d'utilité latente U_{ij} par la relation suivante :

$$\begin{cases} Y_{ij} = 1 & \text{si } U_{ij} > 0 \\ Y_{ij} = 0 & \text{sinon} \end{cases}$$

Avec U_{ij} représentant l'utilité latente du scénario i pour le médecin j ³². Conformément à la théorie de l'utilité aléatoire (Thurstone 1927; McFadden 1974), nous faisons l'hypothèse que la différence d'utilité latente peut être décomposée en une composante systématique, $X_{ij}\beta$, et une composante aléatoire, ϵ_{ij} :

$$U_{ij} = X_{ij}\beta + \epsilon_{ij}$$

où X_{ij} est un vecteur d'attributs associés au scénario i , et β est un vecteur de coefficients reflétant la désirabilité des attributs. La composante aléatoire ϵ_{ij} représente l'ensemble des caractéristiques inobservables, telles que les erreurs de mesure et/ou

³² Ce modèle fait implicitement l'hypothèse que l'utilité associée à la situation de référence du médecin est nulle. Nous revenons sur cette hypothèse en discussion.

l'hétérogénéité des préférences dans l'échantillon. Ainsi, la probabilité que le médecin j « accepte » le scénario i est donnée par :

$$P_{ij} = P(Y_{ij} = 1 | X_{ij}) = P(X_{ij}\beta + \epsilon_{ij} > 0) = P(\epsilon_{ij} < U_{ij})$$

En supposant que les ϵ_{ij} sont indépendants et identiquement distribués selon une loi logistique, nous obtenons le modèle logit conditionnel (McFadden 1974) dans lequel la probabilité P_{ij} s'écrit :

$$P_{ij} = \frac{e^{\mu U_{ij}}}{1 + e^{\mu U_{ij}}}$$

où μ est un paramètre d'échelle qui est inversement proportionnel à la variance des termes d'erreurs : $\mu = \frac{\pi}{\sqrt{6\sigma_\epsilon^2}}$.

L'existence de ce paramètre d'échelle complexifie la comparaison des préférences entre cancers : comme le vecteur des paramètres β est estimé à un facteur d'échelle près, il n'est pas possible de comparer la magnitude des coefficients. Pour effectuer cette comparaison, il est nécessaire d'exprimer les coefficients sur échelle commune de rapport. Cette opération peut être réalisée en calculant des taux marginaux de substitution (TMS) entre attributs, en choisissant toujours un même attribut comme comparateur fixe pour chaque cancer. Il est intéressant de choisir l'attribut monétaire comme comparateur fixe s'il existe, car cela permet d'interpréter les résultats sous la forme de dispositions marginales à payer ou à renoncer. Dans notre cas par exemple, le TMS entre l'attribut x et la rémunération additionnelle s'interprète comme la disposition marginale à renoncer à un montant de paiement additionnel pour bénéficier (d'une unité supplémentaire) de l'attribut x , tout en conservant le même niveau de satisfaction.

Avant de présenter les formules pour le calcul des TMS, nous détaillons les différents modèles de choix discrets estimés, chaque modèle permettant de répondre à une question spécifique de la recherche.

3.1.2. Le modèle logit conditionnel « simple » à effets fixes (M1)

L'analyse des données de choix discret se fait dans un premier temps par un modèle logit conditionnel « simple » tenant compte de la structure répétée (de panel) des données.

Dans la mesure où chaque médecin effectue plusieurs choix (12 au total) et que ces choix sont potentiellement corrélés, l'hypothèse d'indépendance des termes d'erreurs ϵ_{ij} nécessaire à l'estimation du modèle logit conditionnel n'est pas vérifiée. Pour pallier ce défaut d'hypothèse, le modèle est estimé grâce à l'approche GEE (Generalized Estimating Equations) proposée dans le cadre des modèles longitudinaux (Liang and Zeger 1986). En prenant en compte la corrélation des termes d'erreur pour un même individu, cette approche produit des estimations robustes des écarts-types des coefficients.

Ce premier modèle a pour objectif de déterminer la préférence moyenne des médecins pour les différents attributs considérés, sans tenir compte des différences inter-médecins susceptibles de caractériser l'hétérogénéité des préférences. Dans le vocabulaire statistique, il s'agit d'un modèle dit « population-averaged » dont les estimations ont vocation à être généralisées à l'ensemble de la population, par opposition au modèle dit « unit-specific » qui prend en compte les différences inter-médecins et produit des estimations spécifiques à l'échantillon considéré (Hu et al 1998). Ces deux approches diffèrent donc dans les techniques statistiques d'estimation utilisées : l'approche « population-averaged » traite les effets individuels médecin comme des effets fixes susceptibles d'introduire du bruit et utilise une version de l'estimateur « within » adaptée aux variables discrètes pour éliminer ces effets. En revanche, l'approche « unit-specific » traite les effets individuels médecin comme des effets aléatoires, qui sont estimés en sus des autres paramètres du modèle. Elle sera donc privilégiée pour l'analyse ultérieure de l'hétérogénéité des préférences.

Dans le modèle 1, l'utilité latente U_{ij} du médecin j associée au scénario i s'écrit :

$$U_{ij} = \alpha_j + \beta_1 Doc^i + \beta_2 Form^i + \beta_3 List^i + \beta_4 Aide^i + \beta_5 Rem^i + \epsilon_{ij} \quad (M1)$$

où $Doc^i, \dots, Aide^i$ sont les valeurs des attributs définies pour le scénario i codés en variables indicatrices, c'est à dire, l'attribut est codé 1 si la modalité « oui » est présentée dans le scénario i , 0 sinon. Ainsi, le coefficient β_k indique l'effet, sur le choix du médecin, du fait bénéficié de la modalité « oui » de l'attribut k par rapport à ne pas en bénéficier, i.e. le niveau de référence. L'attribut monétaire Rem^i est lui codé de façon continue, prenant les valeurs 0, 1, 3, ou 5 (correspondant respectivement aux niveaux 0%, 1%, 3% et 5% d'augmentation des honoraires) selon le scénario proposé. De plus, β_1, \dots, β_5 sont les coefficients du modèle à estimer, reflétant la désirabilité moyenne des attributs (ou niveaux d'attributs). Nous faisons l'hypothèse que ces coefficients sont tous de signe positif, reflétant

une utilité marginale croissante des 5 dispositifs proposés. Les constantes α_j représentent les effets fixes « médecin », qui sont supprimés dans l'estimation du modèle logit conditionnel à effets fixes.

Hypothèses sur les termes d'erreur :

$\varepsilon_{ij} \sim \text{logitique}$

$Cov(\varepsilon_{ij}, \varepsilon_{kl}) = 0$ avec $i \neq k, j \neq l$

3.1.3. Le modèle logit conditionnel à effets fixes avec interactions (M2)

La deuxième question posée par cette étude est de savoir si les préférences des médecins pour les différents dispositifs proposés varient selon qu'un autre dispositif lui est associé ou non. Autrement dit, nous souhaitons étudier s'il existe des interactions entre dispositifs, ces interactions ayant été préalablement spécifiées dans la sous section 2.3 (Design experimental). Par ailleurs, l'attribut rémunération est introduit au carré car nous suspectons un effet non linéaire de cette variable.

Le modèle avec interaction entre dispositifs s'écrit de la façon suivante :

$$\begin{cases} U_{ij} = \alpha_j + \beta_1 Doc^i + \beta_2 Form^i + \beta_3 List^i + \beta_4 Aide^i + \beta_5 Rem^i \\ + \beta_6 List^i * Aide^i + \beta_7 Doc^i * Rem^i + \beta_8 Rem^i * Rem^i + \varepsilon_{ij} \end{cases} \quad (M2)$$

Avec les hypothèses précédentes sur les termes d'erreurs ε_{ij} .

Hypothèses sur le signe des coefficients d'interaction :

- $\beta_6 > 0$: l'aide en personnel et le listing sont des dispositifs complémentaires.
- $\beta_7 > 0$: les documents et la rémunération sont des dispositifs complémentaires.
- $\beta_5 > 0$ et $\beta_8 < 0$: l'utilité marginale de la rémunération est croissante à taux décroissant.

Les modèles (M1) et (M2) sont estimés à partir de la procédure *Clogit* du logiciel STATA13, qui permet d'estimer des modèles logistiques de panel à effets fixes, et en spécifiant une correction des écarts types des coefficients grâce à l'instruction « cluster-adjusted (robust) standard errors » pour tenir compte de la structure répétée des données par médecin. Seules les interactions significatives au seuil de 5% sont conservées dans le modèle

final, dont les estimations sont utilisées pour calculer les TMS entre les attributs non monétaires et la rémunération.

3.1.4. Estimation des taux marginaux de substitution (TMS)

Le TMS entre l'attribut non monétaire (x) et la rémunération (rem), noté $TMS_{rem \rightarrow x}$, se calcule comme le rapport des utilités marginales entre les attributs x et rem :

$$TMS_{rem \rightarrow x} = \frac{\partial U / \partial x}{\partial U / \partial rem}$$

- Cas n°1 : β_6, β_7 et β_8 ne sont pas significatifs: $TMS_{rem \rightarrow x} = \frac{\beta_x}{\beta_5}$
- Cas n°2 : β_6 et β_7 ne sont pas significatifs mais β_8 l'est : $TMS_{rem \rightarrow x} = \frac{\beta_x}{\beta_5 + 2\beta_8 rem}$
- Cas n°3 : β_6, β_7 et β_8 sont significatifs :

$$\left\{ \begin{array}{l} TMS_{rem \rightarrow doc} = \frac{\beta_1 + \beta_7 rem}{\beta_5 + 2\beta_8 rem} \\ TMS_{rem \rightarrow form} = \frac{\beta_2}{\beta_5 + 2\beta_8 rem} \\ TMS_{rem \rightarrow list} = \frac{\beta_3 + \beta_6 Aide}{\beta_5 + 2\beta_8 rem} \\ TMS_{rem \rightarrow aide} = \frac{\beta_4 + \beta_6 List}{\beta_5 + 2\beta_8 rem} \end{array} \right.$$

La valeur du $TMS_{rem \rightarrow x}$ obtenue s'interprète comme le pourcentage de paiement additionnel auquel le médecin est prêt à renoncer afin de bénéficier de l'attribut x . Afin d'interpréter les $TMS_{rem \rightarrow x}$ en valeur et non plus en pourcentage, le modèle final est estimé à partir de la variable rem recodée en unités monétaires, c'est-à-dire prenant les valeurs 0, 700, 2100 ou 3500, correspondant respectivement à 0%, 1%, 3% et 5% du revenu moyen des médecins généralistes français (environ 70 000€ en 2010³³).

3.1.5. Le modèle logit conditionnel à effets aléatoires (M3)

La troisième question posée par cette étude est celle de déterminer si les préférences varient selon les caractéristiques des médecins. Le modèle logit à effets aléatoires, aussi appelé modèle logit mixte (MIXL), fournit un cadre rigoureux pour répondre à la question de l'hétérogénéité des préférences (Stiratelli et al 1984; Wong and Mason 1985). Ce modèle comprend deux niveaux : un niveau « scénario » (niveau 1) et un niveau « médecin » (niveau 2), et se traduit par l'introduction de deux composantes de variabilités (inter et intra-médecin).

³³ Source : www.ecosante.fr.

Ainsi, ce modèle est dit « subject-specific » car il considère les effets individuels médecins comme aléatoires et utilise des méthodes d'estimation de type maximum de vraisemblance marginale (Hu et al 1998).

Dans ce modèle, nous relâchons l'hypothèse d'homogénéité des préférences en spécifiant les différents niveaux d'attributs comme aléatoires : c'est le modèle dit à *pentés aléatoires* (Snijders and Bosker 1999). Ce modèle permet de tester si certains dispositifs sont plus ou moins valorisés selon certains profils de médecins, et se traduit par l'estimation d'interactions entre attributs et variables médecins.

3.1.5.1. Ecriture du modèle

Le modèle logit mixte à pentés aléatoires s'écrit de la façon suivante (les coefficients soulignés sont aléatoires):

$$\left\{ \begin{array}{l} U_{ij} = \underline{\alpha}_{0j} + \sum_k \underline{\beta}_k x_k^i + \varepsilon_{ij} \quad (1) \\ \underline{\alpha}_{0j} = \alpha_{00} + u_{0j} \quad (2) \\ \underline{\beta}_k = \beta_{k0} + \sum_l \gamma_l MG_l^j + u_{kj} \quad (3) \end{array} \right. \quad (M3)$$

où x_k^i est la valeur de l'attribut k associée au scénario i , α_{0j} est une constante aléatoire estimée pour chaque médecin, MG_l^j est un vecteur de variables caractérisant le médecin j et γ_l représente le vecteur de coefficients à estimer correspondant. Les variables MG_l sont sélectionnées par une procédure stepwise avec un seuil de significativité de 10%³⁴. Par ailleurs, ε_{ij} est un terme d'erreur aléatoire variant d'un médecin à l'autre et d'un scénario à l'autre, représentant la variabilité intra-médecin ; α_{00} est la constante générale du modèle, et u_{0j} est un terme d'erreur aléatoire autour de cette constante, variant d'un médecin à l'autre, représentant la variabilité inter-médecin.

Dans ce modèle, le coefficient β_k représentant l'utilité marginale de l'attribut k est traité comme une variable aléatoire, qui est fonction de variables au niveau médecin (MG_l) et d'un terme d'erreur aléatoire u_{kj} variant d'un médecin à l'autre.

³⁴ Ce seuil de significativité plus élevé se justifie par le fait que nous travaillons sur des échantillons de taille réduite (133 médecins environ par version du questionnaire), se traduisant par une perte de puissance dans l'estimation des effets médecins par comparaison à l'estimation de l'impact des attributs.

Ecrit sous sa forme développée, en substituant les expressions de α_{0j} et β_k des équations (2) et (3) dans l'équation (1), ce modèle s'écrit de la façon suivante :

$$U_{ij} = \underbrace{\alpha_{00} + \sum_{m \neq k} (\beta_m x_m^i) + \beta_{k0} x_k + \sum_l \gamma_l (MG_l^j * x_k^i)}_{\text{partie fixe}} + \underbrace{u_{kj} x_k^i + u_{0j} + \varepsilon_{ij}}_{\text{partie aléatoire}}$$

Ce modèle revient donc à estimer l interactions entre l'attribut x_k et les variables médecin MG_l .

Hypothèses sur les termes d'erreur:

$\varepsilon_{ij} \sim \text{logistique}$

$Cov(\varepsilon_{ij}, \varepsilon_{kl}) = 0$ avec $i \neq k$ et $j \neq l$

$$\begin{pmatrix} u_{0j} \\ u_{kj} \end{pmatrix} \sim \mathcal{N} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u0}^2 & \tau_{ok} \\ \tau_{ok} & \sigma_{uk}^2 \end{pmatrix} \right)$$

Par ailleurs :

- β_{k0} représente la préférence moyenne des médecins pour l'attribut k
- u_{kj} représente l'effet aléatoire autour de cette préférence moyenne.

Par conséquent, si σ_{uk}^2 est significatif, cela indique une hétérogénéité des préférences individuelles autour de l'attribut k . L'introduction des variables MG_l dans l'équation (3) permet d'expliquer les composantes de cette variabilité et donc de caractériser l'hétérogénéité des préférences.

Le nombre de composantes de variabilité augmente rapidement en fonction du nombre d'effets aléatoires spécifiés dans le modèle. En pratique, ce nombre élevé de paramètres à estimer se heurte à des problèmes de convergence des modèles. Ainsi, l'analyse de l'hétérogénéité des préférences est effectuée en deux temps :

- Etape 1 : Plusieurs modèles de régression sont estimés en considérant les différents attributs supposés aléatoires l'un après l'autre. Cette première étape est utilisée pour identifier les coefficients/pentes (β_k) avec une variance significativement différente de 0 au seuil de 5%.

- Etape 2 : Pour les coefficients avec une variance significative, des variables au niveau 2 sont introduites afin de caractériser l'hétérogénéité des préférences.

Le modèle 3 est estimé à partir du logiciel HLM6.0, spécialisé dans l'analyse des modèles multi niveaux.

3.1.5.2. Hypothèses sur l'impact des caractéristiques médecins

Parmi les cinq attributs supposés aléatoire, trois avaient une variance significativement différente de 0 au seuil de 5% pour les trois cancers : le listing, l'aide et la rémunération. Les hypothèses faites sur les variables médecin susceptibles de caractériser l'hétérogénéité des préférences pour ces trois attributs/dispositifs sont synthétisées dans le tableau 15.

- *Listing* : Nous faisons l'hypothèse que les médecins en groupe, ayant une patientèle importante, une proportion plus élevée de patients dans les tranches de recommandation des différents dépistages et dépistant régulièrement les différents cancers, sont plus sensibles à ce dispositif.
- *Aide* : Nous faisons l'hypothèse que les médecins ayant une activité intense (en termes de nombre d'heures) et une patientèle importante sont plus sensibles à ce dispositif. Par ailleurs, nous supposons que la pratique en groupe ou le fait d'être aidé dans les prises de rendez-vous (standard, secrétaire) impactent la valorisation du dispositif.
- *Rémunération* : Nous faisons l'hypothèse que les médecins en secteur 1, ayant une proportion plus élevée de patients dans les tranches de recommandation des différents dépistages et dépistant régulièrement les différents cancers, sont plus sensibles à ce dispositif. Par ailleurs, nous supposons que les médecins ayant une activité intense (nombre d'heures et/ou de patients élevés) et ayant perçu un bonus ROSP en 2012 ou 2013 valorisent différemment ce dispositif.

Tableau 15. Variables médecins introduites dans les modèles mixtes et hypothèses

Attribut aléatoire	Variables médecin	Dépistage concerné	Signe attendu	Hypothèses
Tous	Variables de contrôles: âge, sexe, région	les trois	+/-	Pratiques de dépistage et/ou habitudes de soins différentes
Listing	Groupe	les trois	+	Possibilité de partage de dossiers et/ou discussion de cas concrets de patients
	% femmes % 25-65 ans % 50-75 ans	sein et col col sein et côlon	+	Proxy proportion patients concernés par le dépistage
	Fréquence dépistage	les trois	+	Proxy activité dépistage => intérêt pour un nouveau dispositif
Aide	Aide pour la prise de rendez-vous	les trois	+/-	Goût pour l'aide ou au contraire inutilité car déjà aidé
	Groupe	les trois		
	Nombre d'heures/sem Nombre de patients/sem	les trois les trois	+	Proxy activité/ charge de travail élevée Ferait gagner du temps pour les médecins ayant une forte activité
Rémunération	Secteur 2	les trois	-	Rémunération déjà plus élevée
	Nombre d'heures/ sem Nombre de patients/sem		-/+	Revenus déjà plus élevés (ex: revenu cible atteint) ou au contraire intérêt pour la rémunération
	% femmes % 25-65 ans % 50-75 ans	sein et col col sein et côlon	+	Proxy proportions patients concernés par le dépistage
	Rémunération ROSEP en 2012 ou 2013	sein et col	-/+	Déjà rémunéré à la performance ou au contraire "acoutumance" aux P4P
	Fréquence dépistage	les trois	+	Rémunération pour un travail qui prend du temps

3.1.6. Non réponse et non arbitrage

A la question posée dans l'épreuve DCE, les médecins avaient la possibilité de répondre « ne sait pas ». Ces réponses apparaissent dans les modèles économétriques comme des non réponses. Nous avons proposé cette option afin d'éviter la réalisation de choix forcés, et d'identifier :

- (1) Les médecins ayant de plus grandes difficultés à effectuer des choix.
- (2) Les scénarios potentiellement plus difficiles cognitivement que d'autres.

Ainsi, nous avons créé une première variable de comptage comptabilisant pour chaque médecin le nombre de fois où ce dernier avait répondu « ne sait pas » sur l'ensemble des 12 scénarios proposés. Puis, nous avons créé une seconde variable de comptage comptabilisant pour chaque scénario le nombre de réponses « ne sait pas », permettant de distinguer deux classes de scénarios : les scénarios dits « difficiles » comptabilisant un nombre de non

réponses supérieur strictement à la moyenne, et ceux dits « faciles » comptabilisant un nombre de non réponse inférieur à la moyenne.

Par ailleurs, l'estimation des modèles de choix discret repose sur l'hypothèse que les médecins effectuent des arbitrages entre les différents scénarios proposés. De fait, nous avons comptabilisé les médecins dits « non traders », c'est-à-dire répondant toujours « oui » ou « ne sait pas », ou toujours « non » ou « ne sait pas » aux épreuves DCE. Ces-derniers ont été par la suite exclus de l'analyse car ils ne vérifiaient pas les hypothèses de base du modèle³⁵.

³⁵ Cette exclusion est par ailleurs justifiée par le fait que, dans l'estimation des modèles logit conditionnels à effet fixes, il est impossible d'appliquer l'opérateur « within » aux individus n'ayant pas au minimum deux réponses différentes, ces derniers étant de fait exclus de l'analyse.

3.2. Résultats des modèles DCE

3.2.1. Statistiques descriptives

Le tableau 16 présente les résultats des réponses à la question : « Changeriez-vous votre pratique de dépistage habituelle pour la situation proposée ? ». Pour les trois cancers, les réponses varient en fonction du scénario proposé : les scénarios 1, 6, 7, 10 et 12 apparaissent les plus plébiscités (plus de 50% de « oui »), tandis que les scénarios 5 et 8 et 11 sont faiblement attractifs (plus de 50% de « non »). Les tests d'égalité des proportions montrent que les réponses sont assez similaires d'un cancer à l'autre, excepté pour les scénarios 3 ($p=0,0334$), 4 ($p=0,0032$) et 6 ($p=0,0806$). A première vue, les différences d'« acceptation » pourraient s'expliquer par le nombre de dispositifs proposé dans chaque scénario: dans la version du questionnaire portant sur le cancer du sein, seulement 14% des médecins « acceptent » le scénario 11 où un seul dispositif est proposé, alors qu'ils sont 80% à « accepter » le scénario 12 où tous les dispositifs sont proposés. Mais cette variabilité n'est pas uniquement liée à la quantité de dispositifs proposés : les scénarios 1 et 2 proposent chacun deux dispositifs, mais ils ne sont pas autant plébiscités : 57% « acceptent » le scénario 1 contre 36% pour le scénario 2 dans le cas du dépistage du cancer du sein. Il y a donc bien un arbitrage qui est effectué en fonction de la nature du dispositif proposé dans chaque scénario.

En ce qui concerne la difficulté des scénarios, calculée comme une proportion de réponses « ne sait pas » supérieure à la moyenne (10,2% pour le cancer du sein, 10,0 % pour le cancer du col de l'utérus et 9,3% pour le cancer colorectal), les scénarios 3 et 9 apparaissent les plus difficiles pour les trois cancers. Au contraire, les scénarios 1, 11 et 12 apparaissent plus « faciles » pour les trois cancers . Le taux de réponse « ne sait pas », plus élevé pour certains scénarios, pourrait s'expliquer par la proximité entre la situation hypothétique proposée et la situation habituelle des médecins. Par exemple, le scénario 3 présente une situation proche de celle de la plupart des médecins, qui ont à leur disposition des documents à destination des patients et une possibilité d'augmenter leurs honoraires (d'un peu moins de 1%) grâce à la rémunération sur objectif de santé publique (ROSP). En revanche, les scénarios 1, 11 et surtout 12 se démarquent davantage de leur situation habituelle.

Tableau 16. Réponses aux épreuves DCE selon le type de cancer

Scénarios hypothétiques	Cancer du sein		Cancer du col de l'utérus		Cancer colorectal		Test d'égalité des proportions*
	N	%	N	%	N	%	P-value
Scénario 1	<i>list + rem_5%</i>						0.4445
Oui	77	57.0	75	56.4	82	61.2	
Non	48	35.6	45	33.8	43	32.1	
Ne sais pas	10	7.4	13	9.8	9	6.7	
Scénario 2	<i>form + list</i>						0.1974
Oui	48	35.6	56	42.1	45	33.6	
Non	70	51.9	63	47.4	82	61.2	
Ne sais pas	17	12.6	14	10.5	7	5.2	
Scénario 3	<i>doc + rem_3%</i>						0.0334**
Oui	41	30.4	45	33.8	57	42.5	
Non	73	54.1	69	51.9	61	45.5	
Ne sais pas	21	15.6	19	14.3	16	11.9	
Scénario 4	<i>doc + form + list + aide</i>						0.0032***
Oui	61	45.2	76	57.1	49	36.6	
Non	59	43.7	42	31.6	75	56.0	
Ne sais pas	15	11.1	15	11.3	10	7.5	
Scénario 5	<i>doc + aide + rem_1%</i>						0.1747
Oui	40	29.6	47	35.3	33	24.6	
Non	82	60.7	71	53.4	79	59.0	
Ne sais pas	13	9.6	15	11.3	22	16.4	
Scénario 6	<i>form + aide + rem_5%</i>						0.0806*
Oui	75	55.6	85	63.9	81	60.5	
Non	47	34.8	37	27.8	38	28.4	
Ne sais pas	13	9.6	11	8.3	15	11.2	
Scénario 7	<i>doc + form + rem_3%</i>						0.288
Oui	72	53.3	68	51.1	69	51.5	
Non	50	37.0	46	34.6	50	37.3	
Ne sais pas	13	9.6	19	14.3	15	11.2	
Scénario 8	<i>form + rem_1%</i>						0.4043
Oui	46	34.1	44	33.1	47	35.1	
Non	68	50.4	77	57.9	76	56.7	
Ne sais pas	21	15.6	12	9.0	11	8.2	
Scénario 9	<i>doc + list + rem_1%</i>						0.6735
Oui	57	42.2	57	42.9	49	36.6	
Non	64	47.4	62	46.6	69	51.5	
Ne sais pas	14	10.4	14	10.5	16	11.9	
Scénario 10	<i>list + aide + rem_3%</i>						0.2578
Oui	73	54.1	76	57.1	70	52.2	
Non	50	37.0	43	32.3	47	35.1	
Ne sais pas	12	8.9	14	10.5	17	12.7	
Scénario 11	<i>aide</i>						0.5472
Oui	19	14.1	24	18.1	19	14.2	
Non	104	77.0	98	73.7	110	82.1	
Ne sais pas	12	8.9	11	8.3	5	3.7	
Scénario 12	<i>doc + form + list + aide + rem_5%</i>						0.6769
Oui	108	80.0	111	83.5	106	79.1	
Non	22	16.3	20	15.0	21	15.7	
Ne sais pas	5	3.7	2	1.5	7	5.2	
Total	1620	33.6	1596	33.1	1608	33.3	
Total NSP	166	10.2	159	10.0	150	9.3	

*Test du Chi-deux d'égalité des proportions en fonction du type de cancer.

Significativité : *** :1%, ** :5%, * :10%.

Par ailleurs, nous avons calculé le nombre de réponses « ne sait pas » pour chaque médecin. Entre 11 et 13% des médecins ont répondu « ne sait pas » à 4 reprises ou plus et entre 3 et 5% des médecins ont répondu « ne sait pas » à plus de 6 reprises. Seulement 1 médecin a répondu « ne sait pas » à plus de 8 reprises pour les versions portant sur le dépistage des cancers du sein (version A) et du col de l'utérus (version B), et 4 médecins pour la version portant sur le dépistage du cancer colorectal (version C). La Table B.3 en annexe présente les caractéristiques des médecins n'effectuant pas d'arbitrages, soit 25 médecins (18,5%) pour la version A du questionnaire, 20 médecins (15,0%) pour la version B et 19 médecins (14,2%) pour la version C. Aucun profil particulier de médecin « non-trader » ne se distingue pour l'ensemble des versions. En revanche, dans la version A, ces médecins sont surreprésentés dans la région Nord et Centre Est (50%) ainsi que parmi les MGs déclarant rarement ou jamais dépister le cancer du sein (43%) ; dans la version B, ils sont surreprésentés parmi les MGs déclarant avoir perçu une prime ROSP pour le dépistage du cancer du col de l'utérus (35%) et dans la version C, ils sont surreprésentés dans la région Est (35%). Ces résultats montrent que les biais de sélection semblent modérés, mais invitent tout de même à des précautions dans la généralisation des résultats.

3.2.2. Modèles à effets fixes (M1 et M2)

Dépistage du cancer du sein

Les résultats des modèles logit conditionnel à effets fixes pour le dépistage du cancer du sein sont présentés dans le tableau 17. Dans le modèle avec effets principaux uniquement (M1), tous les attributs à l'exception de l'aide en personnel sont significatifs au seuil de 5%. Le coefficient β_5 associé à la rémunération représente l'effet d'une augmentation de 1% de rémunération additionnelle sur le logit de la probabilité d'accepter le scénario proposé. Nous avons préalablement testé la linéarité de l'attribut rémunération en introduisant cette variable en classes : dans le tableau B.4 et sur la figure B.4. (Annexe B) nous remarquons que la probabilité d'accepter le scénario augmente avec le niveau de rémunération mais à taux décroissant, justifiant l'introduction d'un terme au carré dans le modèle 2. Dans le modèle avec interactions (M2), l'effet de la rémunération au carré est significatif ($p=0,009$) mais les interactions entre les attributs aide et listing ainsi que documents et rémunération ne sont pas significatives au seuil de 5% ($p=0,241$, $p=0,693$, respectivement) et ne sont donc pas inclus dans le modèle final. Dans le modèle final, le listing et la formation (financée) sont les

attributs les plus valorisés par les médecins, devant une augmentation de 1% de leur rémunération sur objectif de santé publique.

Tableau 17. Résultats des modèles logit conditionnels à effets fixes pour le dépistage du cancer du sein (N=110 MGs)

	Modèle 1			Modèle 2			Modèle final		
	Coeff	Std. Err.	P	Coeff	Std. Err.	P	Coeff	Std. Err.	P
Effets principaux									
Doc (β_1)	0,492	0,140	0,000	0,312	0,197	0,112	0,353	0,146	0,015
Form (β_2)	0,922	0,169	0,000	1,171	0,219	0,000	1,075	0,187	0,000
List (β_3)	1,044	0,189	0,000	1,435	0,316	0,000	1,193	0,206	0,000
Aide (β_4)	0,161	0,191	0,399	0,520	0,300	0,083	0,279	0,195	0,153
Rem (β_5)	0,410	0,052	0,000	0,960	0,238	0,000	0,834	0,177	0,000
Interactions									
List*Aide (β_6)	-	-	-	-0,373	0,318	0,241	-	-	-
Doc*Rem (β_7)	-	-	-	0,028	0,072	0,693	-	-	-
Rem*Rem (β_8)	-	-	-	-0,112	0,043	0,009	-0,085	0,032	0,008
Obs	1189			1189			1189		
LogL	-411,641			-408,414			-408,754		
Pseudo R2	0,2134			0,2196			0,2189		

Dépistage du cancer du col de l'utérus

Les résultats des modèles logit conditionnel à effets fixes pour le dépistage du cancer du col de l'utérus sont présentés dans le tableau 18. Dans le modèle (*M1*) avec effets principaux, tous les attributs sont significatifs au seuil de 5%. La non-linéarité de l'impact de l'attribut rémunération a été préalablement testée en introduisant cette variable en classes (cf. tableau B.4 et figure B.4). La probabilité d'accepter le scénario proposé augmente avec le niveau de rémunération proposé, mais de façon décroissante, justifiant à nouveau l'introduction d'un terme au carré dans le modèle suivant. Ainsi, dans le modèle (*M2*), l'effet de la rémunération au carré est significatif au seuil de 5% ($p=0,030$), mais les autres interactions ne le sont pas ($p=0,426$; $p=0,286$, respectivement). Dans le modèle final, le listing et la formation sont les attributs les plus valorisés par les médecins, devant l'augmentation de 1% de la rémunération, l'aide en personnel et les documents.

Tableau 18. Résultats des modèles logit conditionnels à effets fixes pour le dépistage du cancer du col de l'utérus (N=113 MGs)

	Modèle 1			Modèle 2			Modèle final		
	Coeff	Std. Err.	P	Coeff	Std. Err.	P	Coeff	Std. Err.	P
Effets principaux									
Doc (β_1)	0.588	0.142	0.000	0.624	0.182	0.001	0.487	0.145	0.001
Form (β_2)	0.903	0.171	0.000	1.100	0.194	0.000	0.997	0.180	0.000
List (β_3)	1.055	0.194	0.000	1.295	0.289	0.000	1.156	0.207	0.000
Aide (β_4)	0.578	0.173	0.001	0.798	0.258	0.002	0.664	0.168	0.000
Rem (β_5)	0.376	0.049	0.000	0.813	0.211	0.000	0.669	0.162	0.000
Interactions									
List*Aide (β_6)	-	-	-	-0.261	0.329	0.426	-	-	-
Doc*Rem (β_7)	-	-	-	-0.067	0.063	0.286	-	-	-
Rem*Rem (β_8)	-	-	-	-0.083	0.038	0.030	-0.059	0.029	0.040
Obs	1231			1231			1231		
LogL	-446.769			-444.709			-445.312		
Pseudo R2	0.2071			0.2108			0.2097		

Dépistage du cancer colorectal

Les résultats des modèles logit conditionnel à effets fixes pour le dépistage du cancer colorectal sont présentés dans le tableau 19. Dans le modèle 1, tous les attributs à l'exception de l'aide en personnel sont significatifs au seuil de 5%. Comme pour les autres dépistages, le test de linéarité de l'effet de l'attribut rémunération (cf. tableau B.4) justifie l'introduction d'un terme au carré dans (M2) : seule la rémunération au carré est significative au seuil de 5% ($p=0,050$). Dans le modèle final, le listing et la formation sont les attributs « non monétaires » les plus valorisés, alors que les documents et l'aide ne sont pas significatifs.

Tableau 19. Résultats des modèles logit conditionnels à effets fixes pour le dépistage du cancer colorectal (N= 115 MGs)

	Modèle 1			Modèle 2			Modèle final		
	Coeff	Std. Err.	P	Coeff	Std. Err.	P	Coeff	Std. Err.	P
Effets principaux									
Doc (β_1)	0.363	0.148	0.014	0.122	0.199	0.539	0.229	0.152	0.132
Form (β_2)	0.601	0.163	0.000	0.740	0.205	0.000	0.759	0.179	0.000
List (β_3)	0.670	0.168	0.000	0.853	0.297	0.004	0.830	0.182	0.000
Aide (β_4)	0.009	0.160	0.956	0.161	0.314	0.607	0.132	0.167	0.428
Rem (β_5)	0.541	0.054	0.000	0.952	0.238	0.000	0.980	0.171	0.000
Interactions									
List*Aide (β_6)	-	-	-	-0.016	0.393	0.968	-	-	-
Doc*Rem (β_7)	-	-	-	0.053	0.067	0.430	-	-	-
Rem*Rem (β_8)	-	-	-	-0.086	0.044	0.050	-0.088	0.030	0.003
Obs	1256			1256			1256		
LogL	-446.641			-443.235			-443.423		
Pseudo R2	0.2330			0.2389			0.2386		

3.2.3. Estimation des taux marginaux de substitution (TMS)

Pour les trois dépistages, la rémunération a un effet non linéaire (β_8 significatif) mais il n'y a pas d'interaction significative entre les dispositifs listing et aide d'une part et documents et rémunération d'autre part et ce, pour aucun des 3 cancers (β_6 et β_7 ne sont pas significatifs). Ainsi, l'estimation des $TMS_{rem \rightarrow x}$ est seulement calculée en fonction du niveau de rémunération additionnelle³⁶ (2^{ème} cas de figure).

Nous choisissons 4 niveaux de rémunération arbitraires mais susceptibles d'intéresser le décideur publique : 250€ (i.e. le niveau maximal que les MGs peuvent obtenir dans le cadre de la RO SP en 2013), 700€, 1400€ et 2100€ correspondant respectivement à 1%, 2% et 3% du revenu moyen des MGs Français. Seules les estimations et intervalles de confiance des TMS pour les attributs (autres que la rémunération) statistiquement significatifs dans les modèles finaux sont reportés dans le tableau 9.

Tableau 20. Estimation des $TMS_{rem \rightarrow x}$ en fonction du niveau de rémunération additionnelle et du type de dépistage

	P = 250€			P = 700€			P = 1400€			P = 2100€		
	Coeff [95% CI]			Coeff [95% CI]			Coeff [95% CI]			Coeff [95% CI]		
Cancer du sein												
TMS _{rem->doc}	320	16	624	373	34	711	502	91	912	768	222	1314
TMS _{rem->form}	974	601	1347	1135	726	1544	1528	993	2064	2338	1130	3547
TMS _{rem->list}	1081	654	1509	1260	796	1724	1696	1109	2284	2596	1307	3884
Cancer du col												
TMS _{rem->doc}	544	123	965	619	171	1067	789	292	1287	1089	471	1706
TMS _{rem->form}	1114	592	1635	1268	725	1810	1616	1002	2231	2229	1172	3287
TMS _{rem->list}	1292	689	1895	1471	836	2105	1875	1130	2621	2586	1275	3897
TMS _{rem->aide}	742	254	1231	845	322	1368	1077	473	1682	1486	615	2357
Cancer colorectal												
TMS _{rem->form}	580	326	834	662	377	946	848	479	1217	1180	563	1797
TMS _{rem->list}	634	360	907	723	419	1028	927	536	1318	1290	641	1939

Lecture : Pour le dépistage du cancer du sein et pour un niveau de rémunération additionnelle de 250€, les médecins seraient prêts à renoncer jusqu'à 974€ de paiement additionnel afin de bénéficier d'une formation financée, autrement dit, ils seraient prêts à payer « de leur poche » 724€ pour celle-ci. En revanche, pour un niveau de rémunération additionnelle de 2100€, ils ne seraient prêts à payer plus que 238€ pour bénéficier de cette formation financée.

Pour les trois cancers, conformément aux résultats des modèles estimés précédemment, les médecins ont de fortes préférences pour le listing et la formation (financée), notamment dans le cadre des cancers du sein et du col de l'utérus. Ainsi, afin de

³⁶ La formule employée est donc : $TMS_{rem \rightarrow x} = \frac{\beta_x}{\beta_5 + 2\beta_8 rem}$ où β_5 et β_8 sont les coefficients du modèle final estimé avec la variable rémunération recodée en unités monétaires.

bénéficiaire d'une formation financée pour le dépistage du cancer du sein (respectivement, d'un listing des patientes dépistées), ils seraient prêts à renoncer jusqu'à 974€ [601-1347€] (respectivement, 1081€ [601-1347€]) dans le scénario de rémunération additionnelle le plus faible, et jusqu'à 2338€ [1130-3547€] (respectivement, 2596€ [1307-3884€]) dans le scénario de rémunération additionnelle le plus élevé. Ces taux marginaux de substitution estimés sont encore plus élevés pour le cancer du col de l'utérus, mais ils sont environ deux fois plus faibles pour le dépistage du cancer colorectal, variant, par exemple, entre 580€ [326-834€] (scénario de paiement le plus faible) et 1180€ [563-1797] (scénario de paiement le plus élevé) pour la formation.

En ce qui concerne les documents et l'aide en personnel, les préférences varient encore plus selon le type de dépistage : les médecins n'y sont pas du tout sensibles pour dépister le cancer colorectal, ne sont sensibles qu'aux documents pour dépister le cancer du sein et sont sensibles aux deux dispositifs pour dépister le cancer du col de l'utérus. Plus précisément, ils ont des préférences plus fortes pour les documents dans le contexte du dépistage du cancer du col de l'utérus ($TMS_{rem \rightarrow doc} = 544\text{€}$) comparé au dépistage du cancer du sein ($TMS_{rem \rightarrow doc} = 320\text{€}$).

3.2.4. Modèles à effets aléatoires (M3)

Les résultats des modèles mixtes à pentes aléatoires (M3) pour le dépistage des cancers du sein, du col de l'utérus et du côlon sont présentés respectivement dans les tableaux 21, 22, et 23. Pour les trois cancers, le listing, l'aide et la rémunération ont une variance significative au seuil de 5% et sont donc inclus comme attributs aléatoires. Trois modèles sont donc estimés pour chaque cancer, considérant respectivement le listing, l'aide et la rémunération comme effets aléatoires. Tous ces modèles sont ajustés par le sexe du médecin, mais seules les autres variables significatives au seuil de 10% sont retenues dans les modèles afin de caractériser l'hétérogénéité des préférences. Rappelons que l'objectif de ces modèles est d'identifier des profils de médecins plus ou moins sensibles à certains dispositifs.

Dépistage du cancer du sein

En ce qui concerne le dépistage du cancer du sein (tableau 21), les préférences diffèrent selon la localisation géographique des médecins : les médecins exerçant dans la région Ile de France ont une préférence plus forte pour le listing des patients dépistés ($p=0,070$), pour l'aide en personnel ($p=0,001$) et pour la rémunération ($p=0,044$). En

revanche, les médecins exerçant dans la région Nord sont moins sensibles à l'aide en personnel ($p=0,012$). Conformément aux hypothèses, les médecins exerçant en groupe ont une préférence accrue pour le listing ($p=0,027$). Cela peut s'expliquer par le fait que leur pratique étant déjà organisée autour du partage des dossiers (probablement beaucoup plus fréquemment informatisés), le bénéfice du listing serait plus important pour eux que pour les médecins exerçant en cabinet solo. De plus, les médecins recourant à une secrétaire pour leur prise de rendez-vous sont plus sensibles à l'aide ($p=0,003$), ce qui pourrait s'expliquer par leur habitude de partage des tâches ou d'aide en personnel. Enfin, les médecins déclarant dépister « systématiquement » ou « très souvent » le cancer du sein ont une préférence plus forte pour la rémunération additionnelle ($p=0,007$, $p=0,035$, respectivement). Ce résultat est aussi cohérent avec les hypothèses émises révélant une attente de compensation financière liée à une pratique active de dépistage.

Tableau 21. Résultats des modèles mixtes pour le dépistage du cancer du sein

Attributs aléatoires	List			Aide			Rem		
	Coeff	Std. Err	p	Coeff	Std. Err	p	Coeff	Std. Err	p
Constante (α_0)	-2.435	0.307	0.000	-2.400	0.302	0.000	-2.499	0.309	0.000
Attributs									
Doc	0.361	0.143	0.012	0.349	0.140	0.013	0.379	0.142	0.008
Form	1.055	0.177	0.000	1.013	0.179	0.000	1.077	0.182	0.000
List	0.738	0.315	0.021	1.138	0.199	0.000	1.182	0.201	0.000
Aide	0.296	0.195	0.129	0.519	-0.509	0.287	0.291	0.197	0.140
Rem	0.837	0.176	0.000	0.809	0.177	0.000	0.597	0.181	0.002
Rem2	-0.085	0.033	0.010	-0.080	0.032	0.014	-0.077	0.032	0.016
Attributs*Covariables									
List*Homme	0.109	0.323	0.737						
List*IledeFrance	0.683	0.373	0.070						
List*Group	0.685	0.305	0.027						
Aide*Homme				0.355	0.277	0.204			
Aide*RdvSecr				0.873	0.278	0.003			
Aide*IledeFrance				0.694	0.206	0.001			
Aide*Nord				-1.376	0.534	0.012			
Rem*Homme							0.091	0.095	0.341
Rem*IledeFrance							0.208	0.102	0.044
Rem*FreqDep_Sys							0.219	0.079	0.007
Rem*FreqDep_TSv							0.212	0.099	0.035

Dépistage du cancer du col de l'utérus

Les préférences des médecins pour les dispositifs associés au dépistage du cancer du col de l'utérus varient aussi selon leur localisation géographique (tableau 22). Les médecins exerçant dans le Sud Ouest de la France sont moins sensibles au listing des patientes dépistées

($p=0,011$) et ceux exerçants dans le Sud Est (autour de la Méditerranée) sont davantage sensibles à la rémunération ($p=0,066$).

Les médecins travaillant en groupe accordent moins d'importance à l'aide en personnel pour le dépistage du cancer du col de l'utérus ($p=0,080$). Pour mieux interpréter ce résultat, il aurait été intéressant de pouvoir prendre en compte la nature de l'activité en groupe et notamment le genre et la spécialité du ou des collègues. Toutefois, cette information n'a pas été collectée dans nos données.

Concernant la rémunération, les médecins ayant une activité moyenne (i.e. entre 100 et 150 patients par semaine) ou élevée (plus de 150 patients par semaine) ont une préférence plus faible pour la rémunération par rapport à ceux ayant une activité plus faible (moins de 100 patients par semaine) ($p=0,001$, $p=0,019$, respectivement).

Tableau 22. Résultats des modèles mixtes pour le dépistage du cancer du col de l'utérus

Attributs aléatoires	List			Aide			Rem		
	Coeff	Std. Err	p	Coeff	Std. Err	p	Coeff	Std. Err	p
Constante (α_0)	-2.467	0.272	0.000	-2.331	0.259	0.000	-2.409	0.263	0.000
Attributs									
Doc	0.481	0.146	0.001	0.482	0.141	0.001	0.495	0.143	0.001
Form	1.042	0.181	0.000	0.995	0.173	0.000	1.036	0.178	0.000
List	1.544	0.375	0.000	1.156	0.199	0.000	1.166	0.201	0.000
Aide	0.675	0.172	0.000	0.705	0.290	0.017	0.666	0.169	0.000
Rem	0.717	0.165	0.000	0.663	0.154	0.000	0.965	0.202	0.000
Rem2	-0.066	0.030	0.027	-0.059	0.028	0.034	-0.052	0.029	0.069
Attributs*Covariables									
List*Homme	1.544	0.375	0.000						
List*SudOuest	-0.944	0.365	0.011						
Aide*Homme				0.251	0.271	0.357			
Aide*Group				-0.486	0.275	0.080			
Rem*Homme							-0.041	0.100	0.682
Rem*Mediterranée							0.158	0.085	0.066
Rem*NbPat_>100							-0.415	0.114	0.001
Rem*NbPat_>150							-0.274	0.115	0.019

Dépistage du cancer colorectal

En ce qui concerne le dépistage du cancer colorectal (tableau 23), les préférences des médecins varient aussi d'une région à l'autre : les médecins exerçant dans le Sud Ouest de la France ont une préférence plus faible pour le listing ($p=0,010$) alors que les médecins exerçant dans l'Est de la France sont moins sensibles à la rémunération ($p=0,007$). Par ailleurs, les médecins déclarant dépister « systématiquement », « très souvent » ou « souvent » le cancer colorectal ont une préférence plus faible pour le listing ($p=0,013$, $p=0,017$, $p=0,081$,

respectivement). Ce résultat peut s'expliquer par l'organisation du dépistage organisé du cancer colorectal : comme la transmission par les structures de gestion des résultats des tests hemocult aux MGs traitants est quasi-automatique pour ce cancer, les médecins dépistant régulièrement le cancer colorectal seraient de fait déjà bien informés.

Les médecins plus jeunes (25-44 ans) et les hommes sont plus sensibles à la rémunération additionnelle ($p=0,000$, $p=0,001$, respectivement). En outre, les médecins ayant une patientèle importante dans la tranche d'âge des recommandations du dépistage du cancer colorectal (plus de 50% de 50-74 ans) ont une préférence plus forte pour la rémunération additionnelle ($p=0.012$). Ce résultat pourrait s'expliquer par le fait qu'ils aient davantage de travail lié à leur profil de patientèle, nécessitant un travail d'éducation au dépistage et en particulier à la réalisation des tests hemocult ; ce temps dédié a un coût d'opportunité qu'il serait nécessaire de compenser financièrement.

Tableau 23. Résultats des modèles mixtes pour le dépistage du cancer colorectal

Attributs aléatoires	List			Aide			Rem		
	Coeff	Std. Err	p	Coeff	Std. Err	p	Coeff	Std. Err	p
Constante (α_0)	-2.488	0.299	0.000	-2.305	0.291	0.000	-2.400	0.301	0.000
Attributs									
Doc	0.239	0.150	0.111	0.239	0.145	0.100	0.245	0.152	0.107
Form	0.824	0.180	0.000	0.722	0.172	0.000	0.782	0.182	0.000
List	2.151	0.621	0.001	0.793	0.176	0.000	0.845	0.184	0.000
Aide	0.127	0.166	0.443	0.030	0.245	0.903	0.139	0.173	0.423
Rem	1.007	0.164	0.000	0.937	0.160	0.000	0.650	0.183	0.001
Rem2	-0.094	0.029	0.002	-0.083	0.029	0.004	-0.066	0.029	0.024
Attributs*Covariables									
List*Homme	-0.056	0.332	0.867						
List*SudOuest	-0.582	0.220	0.010						
List*FreqDep_Sys	-1.447	0.568	0.013						
List*FreqDep_TSv	-1.279	0.528	0.017						
List*FreqDep_Sv	-0.972	0.553	0.081						
Aide*Homme				0.058	0.265	0.827			
Rem*Homme							0.650	0.183	0.001
Rem*Age_25-44							0.563	0.119	0.000
Rem*Est							-0.276	0.099	0.007
Rem*Pct50-74ans>50%							0.205	0.080	0.012

3.3. Discussion

Notre premier objectif en mobilisant la méthode DCE était d'estimer les préférences moyennes des MGs pour les dispositifs et combinaisons de dispositifs susceptibles d'améliorer leur offre de dépistage des cancers. Tout d'abord, nous avons montré que les dispositifs plébiscités par les MGs dépendent de la nature des cancers : ainsi, il ne semble pas y avoir un dispositif qui domine tous les autres pour tous les cancers (y compris la rémunération additionnelle). Ce résultat est déjà intéressant car il semble ne pas avoir déjà été mis en évidence du moins en France, alors qu'il pourrait expliquer pour partie les résultats contrastés obtenus dans les revues de littérature internationale (Grimshaw et al 2001; Sabatino et al 2008). Nous avons également montré que, dans le cadre du dépistage des cancers féminins, les MGs sont particulièrement sensibles au fait de se voir financer une formation sur le dépistage du cancer du sein ou du col de l'utérus ainsi qu'au fait de disposer d'un listing pluriannuel des patientes dépistées, pour identifier les femmes éligibles non à jour de leur dépistage. Il est intéressant de noter que ces dispositifs seraient réalisables en pratique et ne seraient pas forcément aussi coûteux que des incitatifs financiers.

Pour le cancer du côlon, la rémunération additionnelle constitue un levier davantage plébiscité. Ce résultat pourrait s'expliquer par le fait que le MG est amené à jouer un rôle important pour convaincre les individus et/ou expliquer les modalités de réalisation du test, temps qui constitue un coût d'opportunité nécessitant une compensation financière. Un autre argument peut justifier ce résultat : le dépistage du cancer du côlon s'adresse à une population relativement nombreuse dans la patientèle des MGs (hommes et femmes entre 50 et 74 ans). Plus le nombre de patients éligibles est important dans la patientèle du MG, plus leur fournir de la prévention à tous (expliquer et justifier le dépistage) sera consommateur de temps et plus l'atteinte des objectifs ciblés sera potentiellement coûteux.

Notre second objectif permis par la méthode DCE était de comparer finement les préférences des MGs entre les différents dispositifs et selon le type de cancer par l'estimation de taux marginaux de substitution (TMS). Plus précisément, nous avons calculé des TMS entre les 4 attributs documents informatifs, listing de patients, formation professionnelle (transfert en nature), aide en personnel et le 5^e attribut (la rémunération additionnelle) recodé en montants monétaires (rémunérations maximales offertes en cas d'atteinte des objectifs). Ainsi, les résultats des TMS constituent une mesure des dispositions à payer (en renonçant à des montants en euros de P4P) pour bénéficier d'un autre dispositif en contrepartie. Nos

résultats confirment nos premiers constats puisque les dispositions à renoncer pour le listing et la formation sont très élevées pour les cancers féminins, dépassant parfois même le montant de P4P proposé. Pour le cancer du côlon, là encore les résultats sont un peu différents puisque les MGs ont des dispositions à payer pour ces dispositifs environ deux fois moins élevées par rapport à celles obtenues dans le cadre des cancers féminins.

Cependant, ces mesures de la propension à payer doivent être interprétées avec précaution car l'estimation du coefficient associé à la variable rémunération est imprécise pour les trois cancers. En effet, nous estimons un effet linéaire à partir d'une variable ayant un nombre limité de modalités. Bien qu'imprécis, ce procédé a été utilisé dans de nombreux travaux (Frew et al 2001; Wordsworth et al 2006; Berchi et al 2006; Louviere and Islam 2008). Par ailleurs, les montants de bonus de P4P que nous avons proposés comme (0%, 1%, 3% et 5%) ne sont pas espacés de la même façon, pouvant engendrer un effet non linéaire de la variable. Toutefois, cette non-linéarité a été prise en compte dans les modèles par l'introduction d'un terme au carré qui était significatif pour les trois cancers.

Le troisième objectif de ce travail était d'analyser l'hétérogénéité des préférences pour les dispositifs selon le type de cancer. Comme en attestent les résultats des modèles mixtes, les résultats sont assez contrastés et dépendent notamment de la région d'exercice du MG. Pour le dépistage du cancer colorectal, cela pourrait en partie s'expliquer par la montée en charge progressive du dépistage organisé du cancer du côlon qui est plus récent (2009) par rapport au cancer du sein (2004). Par ailleurs, les résultats concernant le dépistage du cancer du sein ont montré que les médecins exerçant en groupe avaient une préférence plus forte pour le listing. Ce résultat est intéressant étant donnée la tendance au regroupement notamment chez les jeunes médecins (Bourgueil et al 2007). Il faut toutefois noter qu'une des limites des modèles mixtes appliqués à l'analyse des données de choix discret est qu'il n'est pas possible en pratique de déterminer si cette hétérogénéité provient réellement de préférences distinctes ou de capacités cognitives différentes selon les profils de MG (Hole 2006; Fiebig et al 2010).

Limites

Une première limite associée à l'estimation des préférences et leur inférence à la population des MGs Français concerne la proportion relativement élevée de médecins n'effectuant pas d'arbitrage, soit respectivement 18% des médecins (cancer du sein), 15% (cancer du col de l'utérus) et 14% (cancer colorectal). Comme l'estimation des préférences

dans les modèles de choix discret repose sur la capacité des individus à arbitrer entre différentes options, ces MGs ont donc été exclus de l'analyse. Cette exclusion peut engendrer des biais de sélection dans le cas où les caractéristiques des MGs dits « non traders » sont différentes de celles des autres MGs. Ainsi, les caractéristiques des MGs « non traders » ont été comparées à celles des médecins s'étant prêtés à l'exercice, et les tests d'indépendance n'ont pas révélé de différences significatives entre les deux groupes en termes de caractéristiques sociodémographiques, d'activité ou de pratique de dépistage. Néanmoins, dans la version du questionnaire portant sur le dépistage du cancer du sein, les médecins exclus étaient sur-représentés dans la région Nord et Centre Est, et dans celle portant sur le dépistage du col de l'utérus, ils déclaraient plus fréquemment avoir perçu une prime ROSP. Toutefois, l'analyse de l'hétérogénéité des préférences étudiée dans les modèles mixtes n'a pas révélé d'impact significatif de ces caractéristiques dans la valorisation des attributs pour les cancers concernés. Ainsi, nous pouvons raisonnablement conclure que les biais de sélection ne sont pas de nature à remettre en question la généralisation des résultats de l'étude.

La méthode DCE repose sur plusieurs hypothèses quant au comportement des individus interrogés, parmi lesquelles, la stabilité des préférences dans le temps et l'absence de préférences lexicographiques, autrement dit, le fait que les individus ne fondent pas leurs choix par rapport à seul niveau d'attribut dans chaque scénario (Nguyen et al 2008). En pratique, le test de ces hypothèses nécessite de répéter certaines épreuves de choix (test de la stabilité des préférences) et d'interroger les individus en leur demandant s'ils ont pris en compte un attribut unique ou plusieurs attributs dans leurs choix successifs (test d'absence de préférences lexicographiques). Pour des raisons de taille de questionnaire, nous n'avons pas implémenté ces tests dans notre étude, et les résultats obtenus reposent donc sur l'hypothèse qu'elles sont vérifiées du moins pour la majorité des médecins répondants. Par ailleurs, nous pensons que les biais potentiels liés à des réponses inconsistents ou des phénomènes de fatigue sont d'ampleur limitée du fait des trois précautions prises en amont : proposition d'une réponse « ne sait pas » permettant de limiter les réponses incohérentes ou les choix forcés, répartition aléatoire des scénarios dans les questionnaires, et test de la capacité des médecins à « arbitrer » entre les différents scénarios.

Une seconde limite est liée à l'hypothèse sous jacente effectuée dans l'estimation des modèles DCE à scénario unique, c'est-à-dire, que l'utilité associée à la situation de référence (i.e. le contexte de la pratique habituelle de chaque MG) est nulle. Une alternative aurait été de se ramener à un modèle de choix discret avec deux options, l'option 1 consistant en une

reconstitution de la situation habituelle des médecins à partir des réponses aux questions additionnelles, et l'option 2 étant le scénario hypothétique proposé. Le modèle estimé aurait donc été un modèle de différence d'utilité « classique » (Berchi and Launoy 2007; Ammi and Peyron 2010). Les questions additionnelles, interrogeant les médecins sur le fait d'avoir suivi ou non une formation sur le dépistage des cancers (incluant la déclaration du ou des cancers suivis), d'utiliser ou non des brochures d'information à destination des patients, ainsi que le montant de P4P obtenu au travers de la ROSP pour le dépistage des cancers du sein et du col de l'utérus, visaient in fine à reconstruire le contexte habituel de la pratique des MGs. Toutefois, une reconstitution totale de la situation habituelle des médecins s'est avérée complexe du fait notamment d'un taux de réponse très faible à la dernière question. Il aurait été possible de reconstituer cette situation pour le dépistage du cancer colorectal, mais la comparaison des résultats avec les autres cancers aurait été complexifiée par l'utilisation d'une technique de modélisation différente.

Pour pallier ce défaut d'hypothèse, nous avons opté pour une solution plus simple à mettre en œuvre, c'est-à-dire d'introduire les caractéristiques de la situation habituelle dans les variables médecin des modèles mixtes, afin d'étudier si ces variables impactaient l'hétérogénéité des préférences. Pour les trois cancers, nous n'avons pas obtenu d'interactions significatives entre ces variables et les attributs : par exemple, les médecins ayant suivi une formation sur le cancer ne valorisaient pas significativement plus ou moins l'attribut formation. Ce résultat tend à montrer que les biais associés à l'hypothèse du modèle de choix discret à scénario unique semblent limités.

Conclusion

Malgré ces limites, les résultats obtenus en mobilisant la méthode DCE sont intéressants et innovants. Il apparaît non seulement que les dispositifs plébiscités sont différents selon les cancers et que les incitatifs financiers ne dominent pas les autres dispositifs ; il semble que les dispositifs que privilégieraient les MGs seraient plutôt des dispositifs qui améliorent la transmission et la qualité des informations entre les médecins et les patients, du moins dans le cadre du dépistage des cancers féminins. Dans la section suivante, nous mobilisons une méthodologie différente afin à la fois de confirmer et de préciser les résultats obtenus.

4. Analyse des données de best-worst scaling (BWS)

Dans cette section, nous mobilisons les méthodes de classement du meilleur au pire (BWS) avec l'objectif de mesurer de façon plus précise l'impact respectif de chaque attribut ou dispositif. Nous présentons dans un premier temps les méthodes statistiques associées à l'analyse des réponses aux épreuves de classement des attributs du meilleur au pire (BWAS). Au cours de ces épreuves, chaque médecin était confronté à 12 scénarios hypothétiques, pour lesquelles il devait indiquer l'item qu'il préférerait le plus et celui qu'il préférerait le moins (cf. figure B.3 en annexe B). Dans un deuxième temps, nous présentons les résultats des différents modèles estimés. Ces derniers sont discutés dans la troisième sous-section.

4.1. Méthodes

4.1.1. Analyse descriptive

Les résultats des épreuves de classement des attributs du meilleur au pire (BWAS) sont analysés dans un premier temps de façon descriptive par la méthode des scores. Pour chaque niveau d'attribut, une analyse au niveau agrégé est réalisée en calculant les fréquences de sélection comme « préféré » et « moins préféré » sur l'ensemble des répondants. Ces deux fréquences sont ensuite utilisées pour calculer le score d'importance « simple », permettant d'établir un classement des niveaux d'attributs sur une échelle d'intervalle (Cohen 2009; Krucien 2012). Toutefois, ce score « simple » est sensible à l'équilibre des niveaux. Comme les 4 niveaux de rémunération apparaissent moins souvent dans le design (3 fois chacun contre 6 pour les autres dispositifs), ce score peut conduire à une sous-estimation de l'importance des niveaux de rémunération. Le score « relatif » est utilisé de façon alternative car il permet de positionner les attributs sur une échelle de rapport (Marley and Louviere 2005). Par ailleurs, il est utile de standardiser les scores relatifs (par rapport au score le plus élevé) afin de faciliter la lecture des résultats (Cohen 2009; Krucien 2012). Les scores « simple » et « relatif » sont donc calculés à partir des formules suivantes :

$$\text{Score simple}_{kl} = \text{Best}_{kl} - \text{Worst}_{kl}$$

$$\text{Score relatif}_{kl} = \sqrt{\frac{\text{Best}_{kl}}{\text{Worst}_{kl}}}$$

avec $Best_{kl}$, la fréquence de sélection du niveau l de l'attribut k comme « préféré », $Worst_{kl}$, la fréquence de sélection du niveau l de l'attribut k comme « moins préféré ».

4.1.2. Analyse économétrique : approche par les paires versus par les attributs

Les résultats des analyses descriptives doivent être complétés par une modélisation économétrique au niveau individuel permettant d'une part, de tester l'importance de chaque niveau d'attribut et d'autre part, de modéliser l'hétérogénéité des préférences. Deux types d'approches peuvent être mobilisées pour analyser les données issues de l'épreuve BWAS.

- L'approche par les paires (*paired model*) utilise comme unité d'analyse les paires d'attributs « préféré / moins préféré » sélectionnées par les répondants. Le modèle cognitif sous-jacent est celui dit du « Maximum Difference Scaling » (MaxDiff) : l'individu évalue toutes les paires d'attributs au sein d'un scénario donné et sélectionne celle dont l'écart est maximal entre les deux attributs (Szeinbach et al 1999).
- L'approche par les attributs (*marginal model*), en revanche, utilise comme observations les attributs et leurs niveaux choisis comme préféré et moins préféré. Le processus cognitif sous-jacent est simultané : l'individu choisit simultanément et indépendamment l'attribut préféré et le moins préféré. Cette approche repose sur une hypothèse importante de réversibilité, selon laquelle la différence d'utilité entre deux niveaux d'attributs est la même quelque soit le positionnement de ces niveaux d'attributs dans les paires (Krucien 2012).

En pratique, les estimations issues des deux approches sont fortement liées, mais le modèle des paires produit généralement des erreurs standard moins élevées (Flynn et al 2008). Par ailleurs, si les hypothèses faites sur le processus psychologique ayant généré les données semblent très fortes, elles correspondent davantage au modèle statistiques sous-jacent à la méthode BWS, *i.e.* le modèle MaxDiff (Finn and Louviere 1992; Marley and Louviere 2005). Afin de choisir une approche qui convienne le mieux à nos données, nous estimons les deux types de modèles et comparons les résultats en termes de log-vraisemblance et de pseudo R^2 (McFadden 1974). Puis, nous sélectionnons l'approche la plus adaptée à nos données, c'est-à-dire minimisant la log-vraisemblance et maximisant le pseudo R^2 .

4.1.3. Approche par les paires

4.1.3.1. Modèle probabiliste des paires

Notons Y_i la variable d'intérêt indiquant la paire « préféré / moins préféré » choisie au sein du scénario i ³⁷. Le modèle statistique sous-jacent suppose que la probabilité relative de choix d'une paire de niveaux d'attribut est proportionnelle à la distance entre les deux niveaux d'attributs sur l'échelle d'utilité latente. Celle-ci peut être analysée par le modèle suivant :

$$D_{bw} = V_{bw} - \epsilon_{bw}$$

où D_{bw} représente la distance (inobservée) entre le niveau d'attribut préféré (best) et le moins préféré (worst). Cette distance est composée de deux parties : une partie systématique observable (V_{bw}) et une partie stochastique inobservable (ϵ_{bw}). La partie observable s'écrit comme la différence de deux positions mesurées sur une échelle d'utilité latente :

$$V_{bw} = L_b - L_w$$

La probabilité de sélectionner la paire bw dans le scénario i peut alors s'écrire:

$$P(Y_i = bw|i) = P(V_{bw} - \epsilon_{bw} > V_{lk} - \epsilon_{lk}) \text{ pour tout } (bw) \neq (lk)$$

En supposant que les termes d'erreur ϵ_{bw} sont indépendants et identiquement distribués selon une loi à valeur extrême de type 1 (ou loi Gumbel), cette probabilité peut être estimée à l'aide du modèle logit multinomial conditionnel:

$$P(Y_i = bw|i) = \frac{e^{\mu V_{bw}}}{\sum_{lk} e^{\mu V_{lk}}} = \frac{e^{\mu(L_b - L_w)}}{\sum_{lk} e^{\mu(L_l - L_k)}}$$

où μ est un paramètre d'échelle qui est inversement proportionnel à la variance du terme d'erreur : $\mu = \frac{\pi}{\sqrt{6\sigma_\epsilon^2}}$

Ce modèle fait l'hypothèse d'indépendance des alternatives non pertinentes (IIA). Cet axiome du modèle logit multinomial signifie que l'ajout ou le retrait d'une alternative (paire)

³⁷ Comme chaque scénario est caractérisé par 5 niveaux d'attributs et qu'un médecin ne peut pas choisir un niveau d'attribut à la fois comme « préféré » et « moins préféré », le nombre d'alternatives (i.e. de paire « meilleur-pire ») possibles par scénario est $5*4 = 20$.

ne modifie pas le rapport des préférences pour les alternatives (paires) restantes (Cheng and Long 2007).

4.1.3.2. Ecriture du modèle statistique

Le modèle des paires s'écrit de la façon suivante :

$$\left\{ \begin{array}{l} D_{bw}^{ij} = \beta_1 doc^i + \beta_2 form^i + \beta_3 list^i + \beta_5 rem^i + \beta_{11} doc_{oui}^i + \beta_{21} form_{oui}^i \\ + \beta_{31} list_{oui}^i + \beta_{41} aide_{oui}^i + \beta_{51} rem_{1\%}^i + \beta_{52} rem_{3\%}^i + \beta_{53} rem_{5\%}^i + \epsilon_{bw}^{ij} \end{array} \right.$$

où D_{bw}^{ij} est la distance entre le niveau d'attribut choisit comme « préféré » et le niveau d'attribut choisi comme « moins préféré » au sein du scénario i pour le médecin j , doc^i, \dots, rem^i représentent les attributs et $doc_{oui}^i, \dots, rem_{5\%}^i$, les niveaux d'attributs ; β_1, \dots, β_5 représente l'impact des attributs, $\beta_{11}, \dots, \beta_{53}$ représente l'utilité associée aux différents niveaux d'attributs (« level scale parameter », (Flynn et al 2008) et ϵ_{bw}^{ij} est le terme d'erreur du modèle. Notons que l'effet de l'attribut *Aide* est omis et sert de référence à l'estimation des effets des autres attributs.

4.1.3.3. Codage des attributs et de leurs niveaux

La méthode de codage en effets (*effect coding*) est particulièrement bien adaptée à la modélisation des données de BWS et est donc sélectionnée. En effet, elle conduit à centrer les attributs sur la moyenne de leurs effets et permet ainsi d'estimer l'impact des attributs dans leur ensemble séparément de l'utilité associée à un niveau particulier de l'attribut (Flynn et al 2008). Cette méthode consiste à coder « 1 » lorsque l'attribut (respectivement, la modalité de l'attribut) est choisi(e) comme préféré(e), « -1 » lorsque l'attribut (respectivement, la modalité de l'attribut) est choisi(e) comme moins préférée, et « 0 » sinon.

4.1.3.4. Construction et codage de la base de données

Préalablement à l'estimation du modèle, la base de donnée doit être manipulée : pour chaque médecin et pour chaque scénario, 20 lignes sont créées correspondants au nombre total de paires (alternatives) possibles par scénario. La variable dépendante est codée 1 pour la paire choisie, et 0 pour toutes les autres paires non choisies. Chaque médecin sélectionne au total 12 paires « préféré-moins préféré » ce qui signifie qu'il y a $12 \times 20 = 240$ observations par médecin.

Une version simplifiée de la base de données est présentée dans le tableau 24. Les 4 premières colonnes ont uniquement un caractère informatif. La variable dépendante est la variable « Choix ». Les variables indépendantes sont de deux sortes : les attributs (*Doc, Form, ..., Rem*) sont codés 1 dans les paires où ils apparaissent comme « préférés », -1 dans les paires où ils apparaissent comme « moins préférés », et 0 sinon ; les niveaux d'attributs (*Doc_oui, Form_oui, ..., Rem_5%*) sont codés 1 dans les paires où la modalité correspondante apparaît comme « préférée », -1 dans les paires où elle apparaît comme « moins préférée », et 0 sinon. Lorsque l'autre modalité de l'attribut apparaît dans la paire (ex : *doc_non*), les signes sont inversés. De fait, pour qu'il n'y ait pas de colinéarité entre variables explicatives, un niveau d'attribut est systématiquement omis de l'analyse (le niveau « non »), ainsi qu'un attribut qui servira de référence (généralement le moins valorisé, l'*Aide* dans notre cas). Le codage en effets rend les niveaux d'attributs centrés sur la moyenne (où la moyenne représente l'effet de l'attribut), de sorte que le niveau omis de l'attribut est calculé comme 1 moins la somme des estimations des niveaux restants.

Tableau 24. Représentation simplifiée de la base de données dans le modèle des paires

Id MG	Id Sc	Préféré	Moins préféré	Choix	Variables indépendantes								
					<i>Doc</i>	<i>Doc_oui</i>	<i>Form</i>	<i>Form_oui</i>	...	<i>Rem</i>	<i>Rem 1%</i>	<i>Rem 3%</i>	<i>Rem 5%</i>
1	1	<i>Doc_non</i>	<i>Form_non</i>	0	1	-1	-1	1	...	0	0	0	0
1	1	<i>Doc_non</i>	<i>List_oui</i>	0	1	-1	0	0	...	0	0	0	0
1	1	<i>Doc_non</i>	<i>Aide_non</i>	0	1	-1	0	0	...	0	0	0	0
1	1	<i>Doc_non</i>	<i>Rem_5%</i>	0	1	-1	0	0	...	-1	0	0	-1
1	1	<i>Form_non</i>	<i>Doc_non</i>	0	-1	1	1	-1	...	0	0	0	0
1	1	<i>Form_non</i>	<i>List_oui</i>	0	0	0	1	-1	...	0	0	0	0
1	1	<i>Form_non</i>	<i>Aide_non</i>	0	0	0	1	-1	...	0	0	0	0
1	1	<i>Form_non</i>	<i>Rem_5%</i>	0	0	0	1	-1	...	-1	0	0	-1
1	1	<i>List_oui</i>	<i>Doc_non</i>	0	-1	1	0	0	...	0	0	0	0
1	1	<i>List_oui</i>	<i>Form_non</i>	0	0	0	-1	1	...	0	0	0	0
1	1	<i>List_oui</i>	<i>Aide_non</i>	0	0	0	0	0	...	0	0	0	0
1	1	<i>List_oui</i>	<i>Rem_5%</i>	0	0	0	0	0	...	-1	0	0	-1
1	1	<i>Aide_non</i>	<i>Doc_non</i>	0	-1	1	0	0	...	0	0	0	0
1	1	<i>Aide_non</i>	<i>Form_non</i>	0	0	0	-1	1	...	0	0	0	0
1	1	<i>Aide_non</i>	<i>List_oui</i>	0	0	0	0	0	...	0	0	0	0
1	1	<i>Aide_non</i>	<i>Rem_5%</i>	0	0	0	0	0	...	-1	0	0	-1
1	1	<i>Rem_5%</i>	<i>Doc_non</i>	0	-1	1	0	0	...	1	0	0	1
1	1	<i>Rem_5%</i>	<i>Form_non</i>	1	0	0	-1	1	...	1	0	0	1
1	1	<i>Rem_5%</i>	<i>List_oui</i>	0	0	0	0	0	...	1	0	0	1
1	1	<i>Rem_5%</i>	<i>Aide_non</i>	0	0	0	0	0	...	1	0	0	1

Lecture : Le médecin 1 a choisi pour le scénario 1 la paire (*Rem_5%*, *Form_non*).

Le modèle des paires est estimé sous STATA13 grâce à la commande *Clogit*, en indiquant que les observations sont groupées par individu et par scénario (Flynn et al 2007), et en spécifiant une correction des écarts types des coefficients grâce à l'instruction « cluster-adjusted (robust) standard errors ».

4.1.4. Approche par les attributs

Nous présentons maintenant l'approche par les attributs, qui revient à faire l'hypothèse que l'individu (le médecin) choisit simultanément et indépendamment l'attribut préféré et le moins préféré au sein de chaque scénario.

4.1.4.1. Le modèle probabiliste marginal

Notons Y_i^{best} la variable indiquant le niveau d'attribut choisi comme « préféré » et Y_i^{worst} la variable indiquant le niveau d'attribut choisi comme « moins préféré » au sein du scénario i . Cette sélection étant supposée simultanée et indépendante, le nombre d'alternatives possibles pour chacun des choix « best » et « worst » est égale au nombre d'attributs, i.e. 5 dans notre cas.

Notons U_b l'utilité associée au choix du niveau d'attribut b comme préféré et U_w l'utilité associée au choix du niveau d'attribut w comme moins préféré :

$$\begin{cases} U_b = V_b + \epsilon_b \\ U_w = V_w + \epsilon_w \end{cases}$$

La probabilité de choisir le niveau d'attribut b comme « préféré » et le niveau d'attribut w comme « moins préféré » au sein du scénario i s'écrit :

$$\begin{cases} P(Y_i^{best} = b|i) = P(V_b - \epsilon_b > V_k - \epsilon_k) \text{ pour tout } k \neq b \\ P(Y_i^{worst} = w|i) = P(V_w - \epsilon_w > V_k - \epsilon_k) \text{ pour tout } k \neq w \end{cases}$$

En supposant que les termes d'erreur ϵ_b et ϵ_w sont indépendants et identiquement distribués selon une loi à valeur extrême de type 1 (ou loi Gumbel), et en notant Y_i la variable d'intérêt représentant le couplet $(Y_i^{best}, Y_i^{worst})$, cette probabilité peut être estimée à l'aide du modèle logit multinomial conditionnel:

$$P(Y_i = bw|i) = \frac{e^{\mu V_b}}{\sum_k e^{\mu V_k}} * \frac{e^{\mu V_w}}{\sum_k e^{\mu V_w}}$$

Avec μ le paramètre d'échelle qui est inversement proportionnel à la variance du terme d'erreur. Ce modèle revient à estimer un choix parmi 10 alternatives (5 « best », 5 « worst »), il fait donc aussi l'hypothèse d'indépendance des alternatives non pertinentes (IIA).

4.1.4.2. Ecriture du modèle statistique

Le modèle marginal s'écrit de la façon suivante :

$$\begin{cases} U_b^{ij} = BWindic + \beta_1 doc^i + \beta_2 form^i + \beta_3 list^i + \beta_5 rem^i + \beta_{11} doc_{oui}^i + \beta_{21} form_{oui}^i \\ \quad + \beta_{31} list_{oui}^i + \beta_{41} aide_{oui}^i + \beta_{51} rem_{1\%}^i + \beta_{52} rem_{3\%}^i + \beta_{53} rem_{5\%}^i + \epsilon_b^{ij} \\ U_w^{ij} = BWindic + \beta_1 \overline{doc}^i + \beta_2 \overline{form}^i + \beta_3 \overline{list}^i + \beta_5 \overline{rem}^i + \beta_{11} \overline{doc}_{oui}^i + \beta_{21} \overline{form}_{oui}^i \\ \quad + \beta_{31} \overline{list}_{oui}^i + \beta_{41} \overline{aide}_{oui}^i + \beta_{51} \overline{rem}_{1\%}^i + \beta_{52} \overline{rem}_{3\%}^i + \beta_{53} \overline{rem}_{5\%}^i + \epsilon_w^{ij} \end{cases}$$

où U_b^{ij} est l'utilité associée au choix du niveau d'attribut b comme préféré et U_w^{ij} est l'utilité associée au choix du niveau d'attribut w comme moins préféré au sein du scénario i pour le médecin j , doc^i, \dots, rem^i représentent les attributs et $doc_{oui}^i, \dots, rem_{5\%}^i$, les niveaux d'attributs codés en effets (*effect coding*) comme pour le modèle des paires. Par ailleurs, $BWindic$ est une indicatrice codée « 1 » dans l'équation de U_b^{ij} et « -1 » dans l'équation de U_w^{ij} , $\overline{doc}^i, \dots, \overline{rem}^i$ sont les attributs codés à l'inverse de doc^i, \dots, rem^i et $\overline{doc}_{oui}^i, \dots, \overline{rem}_{5\%}^i$ sont les niveaux d'attributs codés à l'inverse de $doc_{oui}^i, \dots, rem_{5\%}^i$. Notons que l'effet de l'attribut *Aide* est omis et sert de référence à l'estimation des effets des autres attributs.

4.1.4.3. Construction et codage de la base de données

Pour estimer le modèle marginal, la base de données est aussi manipulée en créant 10 lignes par scénarios et par médecins, les 5 premières lignes correspondant aux alternatives possibles pour le choix « préféré » et les 5 dernières correspondant aux alternatives possibles pour le choix « moins préféré ».

Une version simplifiée de la base de données correspondante est présentée dans le tableau 25. La variable dépendante est la variable Choix : chaque médecin effectue au total 12 choix « préféré / moins préféré » parmi 10 alternatives dans chaque épreuve, ce qui signifie qu'il y a $12 \times 10 = 120$ observations par médecin. Les variables indépendantes sont codées de façon similaire que pour le modèle des paires, et une variable supplémentaire est ajoutée, la variable $BWindic$, codée « 1 » pour l'ensemble de choix des niveaux d'attributs « préférés », et « -1 » pour l'ensemble de choix des niveaux d'attributs « moins préférés ».

Tableau 25. Représentation simplifiée de la base de données dans le modèle marginal.

Id MG	Id Sc	Préfééré	Moins préféré	Choix	Variables indépendantes									
					BWindic	Doc	Doc_oui	Form	Form_oui	...	Rem	Rem1%	Rem3%	Rem5%
1	1	Doc_non	-	0	1	1	-1	0	0	...	0	0	0	0
1	1	Form_non	-	0	1	0	0	1	-1	...	0	0	0	0
1	1	List_oui	-	0	1	0	0	0	0	...	0	0	0	0
1	1	Aide_non	-	0	1	0	0	0	0	...	0	0	0	0
1	1	Rem_5%	-	1	1	0	0	0	0	...	1	0	0	1
1	1	-	Doc_non	0	-1	-1	1	0	0	...	0	0	0	0
1	1	-	Form_non	1	-1	0	0	-1	1	...	0	0	0	0
1	1	-	List_oui	0	-1	0	0	0	0	...	0	0	0	0
1	1	-	Aide_non	1	-1	0	0	0	0	...	0	0	0	0
1	1	-	Rem_5%	0	-1	0	0	0	0	...	-1	0	0	-1

Lecture : Le médecin 1 a sélectionné dans le scénario 1 le niveau 5% de la rémunération comme préféré et le niveau *non* de la formation comme moins préféré.

Comme pour le modèle des paires, ce modèle est estimé sous STATA13 grâce à la commande *Clogit*, en indiquant que les observations sont groupées par individu et par scénario (Flynn et al 2007), et en spécifiant une correction des écarts types des coefficients grâce à l’instruction « cluster-adjusted (robust) standard errors ».

4.1.5. Analyse de l’hétérogénéité des préférences

Une fois le modèle le plus adéquat choisi (en termes de log-vraisemblance et R^2 de McFadden), nous analysons dans un second temps l’hétérogénéité des préférences en ajoutant des termes d’interactions entre les caractéristiques médecins et les attributs identifiés dans la section précédente (analyse DCE) comme étant valorisés différemment selon les médecins : le listing de patients dépistés et la rémunération additionnelle. En revanche, l’aide étant défini comme l’attribut de référence, son impact est omis et aucune interaction avec les variables médecin n’est estimée pour cet attribut.

4.2. Résultats des modèles BWS

4.2.1. Analyses descriptives

Contrairement aux données de choix discret, du fait du design de l'enquête qui imposait au répondant d'effectuer un choix pour poursuivre, il n'y a aucune non-réponse et donc toutes des données sont analysées. Le tableau 26 présente les fréquences de sélection de chaque niveau d'attribut comme « préféré » et « moins-préféré » selon le type de cancer, ainsi que les scores simples. Les résultats sont cohérents pour la quasi-totalité des niveaux d'attributs (à l'exception des documents pour le cancer du sein), c'est-à-dire que les scores simples associés aux niveaux positifs des attributs (« oui », « 1%, 3%, 5% ») sont supérieurs à 0, ce qui signifie qu'ils sont dans l'ensemble davantage choisis comme préférés que moins préférés. Parallèlement, les scores simples associés aux niveaux négatifs des attributs sont tous négatifs : ils sont davantage choisis comme moins préférés que préférés.

Tableau 26. Fréquence de sélection des niveaux d'attributs comme préférés/moins préférés

	Cancer du sein			Cancer du col de l'utérus			Cancer colorectal		
	<i>préféré</i>	<i>-préféré</i>	<i>diff*</i>	<i>préféré</i>	<i>-préféré</i>	<i>diff*</i>	<i>préféré</i>	<i>-préféré</i>	<i>diff*</i>
Doc_oui	122	133	-11	133	93	40	155	81	74
Form_oui	321	50	271	342	47	295	275	46	229
List_oui	268	49	219	269	48	221	270	56	214
Aide_oui	184	172	12	194	173	21	166	187	-21
Rem_1%	132	75	57	120	71	49	132	65	67
Rem_3%	206	35	171	202	30	172	224	23	201
Rem_5%	228	26	202	193	33	160	241	19	222
Doc_non	50	158	-108	39	148	-109	35	197	-162
Form_non	36	257	-221	36	281	-245	31	253	-222
List_non	24	316	-292	24	286	-262	28	321	-293
Aide_non	29	148	-119	25	172	-147	41	125	-84
Rem_non	20	201	-181	19	214	-195	10	235	-225
Total	1620	1620	-	1596	1596	-	1608	1608	-

**diff*: fréquence "préféré" - fréquence "moins préféré". Ce résultat correspond au score simple.

Le tableau 27 présente les résultats des scores relatifs (*Rscore*) et standardisés ainsi que le rang de classement des différents niveaux d'attribut. En ce qui concerne le dépistage du cancer du sein, les 4 niveaux d'attributs ayant le score relatif le plus élevé sont dans l'ordre le niveau 5% de rémunération (*Rscore*=2,96), le financement d'une formation (*Rscore*=2,53), le niveau 3% de rémunération (*Rscore*=2,43) et la présence d'un listing des patientes dépistées (*Rscore*=2,34). Ces quatre niveaux d'attributs sont aussi les mieux classés pour les deux autres cancers, même si le rang de classement diffère légèrement (notamment, le financement de la formation est le mieux classée pour le dépistage du cancer du col de l'utérus). Par

ailleurs, pour les trois dépistages, le niveau 1% de rémunération est toujours classé 5^e, devant la mise à disposition de documents ou l'aide en personnel qui sont classés en 6^e ou 7^e position pour les trois cancers.

Les résultats concernant les niveaux de dispositifs les moins bien classés sont similaires d'un cancer à l'autre : le non financement de formation, l'absence de rémunération et de listing des patients dépistés sont toujours classés en 10^e, 11^e ou 12^e position, respectivement. Il semble donc y avoir une symétrie dans les préférences des médecins, puisque les modalités « oui » de ces attributs sont les mieux classés, alors que les modalités « non » sont les moins bien classées. En revanche, cette symétrie n'est pas parfaite : par exemple, la présence de listing apparaît à la 4^e position pour les trois cancers mais son absence apparaît à la 12^e ou 11^e position, ce peut s'interpréter comme le fait que la désutilité associée à ne pas disposer de listing est plus forte que l'utilité associée au fait d'en disposer.

Tableau 27. Score relatif, score standardisé et rang des niveaux d'attribut

	Cancer du sein			Cancer du col de l'utérus			Cancer colorectal		
	Score relatif	Score Standardisé	Rang	Score relatif	Score Standardisé	Rang	Score relatif	Score Standardisé	Rang
Doc_oui	0.96	32%	7	1.20	44%	6	1.38	39%	6
Form_oui	2.53	86%	2	2.70	100%	1	2.45	69%	3
List_oui	2.34	79%	4	2.37	88%	4	2.20	62%	4
Aide_oui	1.03	35%	6	1.06	39%	7	0.94	26%	7
Rem_1%	1.33	45%	5	1.30	48%	5	1.43	40%	5
Rem_3%	2.43	82%	3	2.59	96%	2	3.12	88%	2
Rem_5%	2.96	100%	1	2.42	90%	3	3.56	100%	1
Doc_non	0.56	19%	8	0.51	19%	8	0.42	12%	9
Form_non	0.37	13%	10	0.36	13%	10	0.35	10%	10
List_non	0.28	9%	12	0.29	11%	12	0.30	8%	11
Aide_non	0.44	15%	9	0.38	14%	9	0.57	16%	8
Rem_non	0.32	11%	11	0.30	11%	11	0.21	6%	12

Lecture : un score relatif supérieur à 1 indique que le niveau d'attribut à une fréquence de sélection comme préféré supérieure à la fréquence de sélection comme moins préféré.

4.2.2. Modélisations économétriques

4.2.2.1. Choix d'un modèle

Le tableau 28 présente une comparaison des estimations du modèle des paires et du modèle marginal. Pour les trois cancers, le modèle des paires est la fois meilleur en termes de vraisemblance (log vraisemblance plus faible) et de qualité prédictive (pseudo R² plus élevé). Ce modèle est donc sélectionné pour les analyses économétriques et l'analyse de l'hétérogénéité des préférences.

Tableau 28. Comparaison des résultats des modèles des paires et marginaux

	Cancer du sein		Cancer du col de l'utérus		Cancer colorectal	
	<i>Paires</i>	<i>Marginal</i>	<i>Paires</i>	<i>Marginal</i>	<i>Paires</i>	<i>Marginal</i>
Obs	32400	16200	31920	15960	32160	16080
Log L	-4048.42	-5257.46	-3966.95	-5154.91	-3869.51	-5063.17
Pseudo-R2	0.1667	0.1484	0.1715	0.1527	0.1972	0.1733

4.2.2.2. Estimation des modèles logit conditionnels multinomiaux des paires

Les résultats des modèles des paires sont présentés dans le tableau 29 (à titre indicatif, les résultats du modèle marginal sont présentés dans le tableau B.5 de l'Annexe B).

L'aide en personnel qui est l'attribut dont l'impact est le plus faible pour deux des trois cancers (pour le cancer du sein il s'agit des documents) est omis dans l'analyse. Ainsi, les coefficients des quatre autres attributs s'interprètent relativement à l'impact de l'aide, qui correspond à la valeur 0 sur une échelle d'intervalle (Flynn et al 2008). Pour les trois cancers, la rémunération est clairement l'attribut le plus valorisé, suivi de la formation et du listing. Les documents sont préférés à l'aide en personnel pour le dépistage du cancer du col de l'utérus uniquement ($p=0,047$).

Les résultats des estimations concernant l'effet des niveaux d'attributs donnent une indication sur le degré d'éloignement des différentes modalités des attributs sur l'échelle d'utilité latente. Par exemple, pour le dépistage du cancer du sein, il y a une différence de 2,919 unités [1,243-(-1,676)] entre le niveau de rémunération le plus élevé (5%) et le niveau le plus faible (0%), versus 1,961 unités [2x0,980] entre les deux niveaux (oui / non) de l'attribut listing ou encore versus 1,893 unités [2x0,947] entre les deux niveaux (oui / non) de financement d'une formation. Ainsi, pour les trois cancers, les niveaux d'attributs les plus éloignées sur l'échelle d'utilité latente sont les niveaux 5% et 3% de rémunération vs 0% et les niveaux « oui » des attributs listing et formation vs « non ». Les deux niveaux d'attributs les moins éloignés sur l'échelle d'utilité latente sont les niveaux associés aux attributs aide et documents.

Tableau 29. Résultats des modèles logit conditionnels multinomiaux (modèles des paires)

Attribut	Cancer du sein			Cancer du col de l'utérus			Cancer colorectal		
	Coeff	Std. Err.	P	Coeff	Std. Err.	P	Coeff	Std. Err.	P
Doc	-0.011	0.053	0.831	0.108	0.054	0.047	0.030	0.057	0.593
Form	0.308	0.056	0.000	0.341	0.058	0.000	0.254	0.057	0.000
List	0.123	0.055	0.025	0.204	0.057	0.000	0.128	0.058	0.028
Aide	-	-	-	-	-	-	-	-	-
Rem	0.599	0.062	0.000	0.550	0.063	0.000	0.646	0.064	0.000
Niveau									
Doc_oui	0.243	0.042	0.000	0.364	0.040	0.000	0.570	0.044	0.000
Doc_non	-0.243	-	-	-0.364	-	-	-0.570	-	-
Form_oui	0.947	0.046	0.000	1.043	0.047	0.000	0.919	0.046	0.000
Form_non	-0.947	-	-	-1.043	-	-	-0.919	-	-
List_oui	0.980	0.045	0.000	0.950	0.045	0.000	1.023	0.048	0.000
List_non	-0.980	-	-	-0.950	-	-	-1.023	-	-
Aide_oui	0.233	0.044	0.000	0.313	0.046	0.000	0.092	0.047	0.050
Aide_non	-0.233	-	-	-0.313	-	-	-0.092	-	-
Rem_1%	-0.211	0.090	0.020	-0.144	0.091	0.112	-0.131	0.090	0.145
Rem_3%	0.644	0.089	0.000	0.808	0.088	0.000	0.930	0.089	0.000
Rem_5%	1.243	0.091	0.000	1.054	0.090	0.000	1.362	0.092	0.000
Rem_non	-1.676	-	-	-1.718	-	-	-2.161	-	-

4.2.2.3. Analyse de l'hétérogénéité des préférences

Les résultats des modèles avec interactions entre attributs et caractéristiques des médecins sont présentés dans le tableau 30. Nous incluons des interactions pour deux des trois dispositifs dont nous avons montré dans la section précédente qu'ils étaient caractérisés par une hétérogénéité des préférences : le listing et la rémunération. L'aide n'est pas considérée car il s'agit de l'attribut de référence dans les modèles. Les interactions sont incluses dans les modèles en suivant les hypothèses émises dans la section précédente (cf. tableau 15). L'interaction avec le sexe du médecin est systématiquement incluse dans les modèles et seules les autres interactions significatives au seuil de 10% sont retenues dans le modèle final.

En ce qui concerne le cancer du sein, les médecins d'Ile de France, qui sont médecins traitants de plus de 75% de leurs patients, déclarant dépister « systématiquement », « très souvent » ou « souvent » le cancer du sein et ayant perçu une rémunération sur objectif de santé publique (ROSP) en 2012 ou 2013, ont une préférence plus forte pour la rémunération. Ces résultats sont similaires à ceux obtenus dans le modèle de choix discrets mixte (cf. tableau 21), à l'exception de l'effet de la variable ROSP qui n'était pas significatif. Cette préférence pour la rémunération avec des montants plus élevés peut s'interpréter au regard des caractéristiques des médecins mises en évidence : les médecins ayant obtenu un bonus ROSP considèrent que ce montant est trop faible au regard des efforts à fournir pour atteindre les

objectifs de santé publique, *a fortiori* s'ils déclarent proposer « systématiquement » le dépistage. Pour tester cette hypothèse, nous avons inclus des interactions entre les niveaux 3% et 5% de rémunération additionnelle et la variable ROSP, et avons obtenu une interaction significative pour le niveau 3% (coeff = 0,476 ; p=0,005) mais non significative pour le niveau 5% (coeff = 0,180 ; p=0,288). Ce résultat montre que les médecins ayant déjà perçu des bonus liés au paiement à la performance seraient davantage sensibles à des montants plus élevés de P4P (mais raisonnables) par rapport aux médecins ne déclarant pas avoir perçu de bonus dans le cadre de la ROSP.

Concernant le dépistage du cancer du col de l'utérus, les médecins de sexe masculin exerçant dans la région Est ont une préférence plus faible pour le listing (p=0,089 et p=0,056 respectivement), à l'inverse des médecins d'Ile de France qui eux, ont une préférence plus forte pour le listing (p=0,009). Par ailleurs, les hommes et les médecins ayant une activité moyenne (entre 100 et 150 patients par semaine) sont moins sensibles à la rémunération (p=0,023 et p=0,000 respectivement), alors que les médecins plus jeunes y sont davantage sensibles. Enfin, similairement aux résultats obtenus pour le cancer du sein, les médecins déclarant dépister « très souvent » ou « souvent » le cancer du col de l'utérus ont une préférence plus forte pour la rémunération (p=0,004 et p= 0,000 respectivement).

En ce qui concerne le dépistage du cancer colorectal, les médecins jeunes (25-44 ans) ont une préférence plus faible pour le listing (p=0,024) mais plus forte pour la rémunération (p=0,030). Conformément aux résultats obtenus avec la méthode DCE (cf. tableau 23), les médecins ayant une patientèle importante dans la tranche d'âge des recommandations du dépistage du cancer colorectal (plus de 50% de 50-74 ans) ont une préférence plus forte pour la rémunération (p=0,065). Enfin, les médecins exerçant dans la région Est sont moins sensibles à la rémunération additionnelle (p=0,018).

Tableau 30. Résultats des modèles logit conditionnels multinomiaux avec interactions (modèles des paires)

Attribut	Cancer du sein			Cancer du col de l'utérus			Cancer colorectal		
	Coeff	Std. Err	P	Coeff	Std. Err	P	Coeff	Std. Err	P
Doc	-0.011	0.054	0.833	0.109	0.055	0.047	0.031	0.057	0.592
Form	0.311	0.057	0.000	0.345	0.059	0.000	0.255	0.058	0.000
List	0.042	0.090	0.638	0.320	0.092	0.000	0.310	0.123	0.012
Aide		<i>ref</i>			<i>ref</i>			<i>ref</i>	
Rem	-0.448	0.301	0.137	0.294	0.220	0.181	0.562	0.150	0.000
Attribut*Covariable									
List*Homme	0.115	0.098	0.239	-0.168	0.099	0.089	-0.041	0.106	0.698
List*Age_25-44		<i>ns</i>			<i>ns</i>		-0.277	0.123	0.024
List*Age_45-54		<i>ns</i>			<i>ns</i>		-0.230	0.106	0.031
List*IledeFrance		<i>ns</i>		0.221	0.085	0.009		<i>ns</i>	
List*Est		<i>ns</i>		-0.177	0.093	0.056		<i>ns</i>	
Rem*Homme	-0.072	0.119	0.544	-0.270	0.119	0.023	-0.129	0.121	0.285
Rem*Age_25-44		<i>ns</i>		0.283	0.136	0.037	0.319	0.147	0.030
Rem*Age_45-54		<i>ns</i>		0.443	0.124	0.000		<i>ns</i>	
Rem*IledeFrance	0.241	0.132	0.067		<i>ns</i>			<i>ns</i>	
Rem*Méditerranée	-0.310	0.118	0.009		<i>ns</i>			<i>ns</i>	
Rem*Est		<i>ns</i>		0.157	0.094	0.095	-0.353	0.150	0.018
Rem*PctMTT>75%	0.490	0.129	0.000		<i>ns</i>			<i>ns</i>	
Rem*Pct50-74ans>50%							0.227	0.123	0.065
Rem*Nbpat[100-150]				-0.512	0.137	0.000		<i>ns</i>	
Rem*FreqDep_Sys	0.765	0.291	0.009		<i>ns</i>			<i>ns</i>	
Rem*FreqDep_TSv	0.693	0.297	0.020	0.551	0.192	0.004		<i>ns</i>	
Rem*FreqDep_Sv	0.736	0.299	0.014	0.772	0.194	0.000		<i>ns</i>	
Rem*Rosp	0.252	0.107	0.019		<i>ns</i>		-	-	-
Niveau									
Doc_oui	0.245	0.042	0.000	0.368	0.041	0.000	0.573	0.044	0.000
Form_oui	0.954	0.046	0.000	1.054	0.047	0.000	0.924	0.046	0.000
List_oui	0.988	0.046	0.000	0.959	0.045	0.000	1.028	0.048	0.000
Aide_oui	0.235	0.044	0.000	0.318	0.047	0.000	0.092	0.047	0.049
Rem_1%	-0.213	0.089	0.016	-0.147	0.089	0.099	-0.134	0.090	0.136
Rem_3%	0.650	0.088	0.000	0.819	0.088	0.000	0.936	0.089	0.000
Rem_5%	1.256	0.092	0.000	1.068	0.090	0.000	1.373	0.092	0.000

Ns : non significatif. Cela signifie que l'interaction correspondante n'a pas été incluse dans le modèle final.

NB : les coefficients associés aux attributs qui sont impliqués dans des interactions (List et Rem) ne peuvent plus s'interpréter comme des effets principaux dans les modèles.

4.3. Discussion

Dans cette section, nous avons analysé les préférences de MGs pour les dispositifs d'amélioration de leur pratique de dépistage, en mobilisant la méthodologie du classement des attributs du meilleur au pire. La méthode BWAS permet de mesurer précisément l'impact respectif de chaque dispositif sur les préférences et donc sur l'utilité du MG. Cette méthode présente plusieurs intérêts par rapport à la méthode DCE. Tout d'abord, à partir de l'information collectée sur les niveaux d'attributs préférés et moins préférés, il est possible d'évaluer la symétrie des préférences des médecins ; autrement dit, il s'agit de vérifier si l'utilité associée au fait de bénéficier d'un niveau attractif d'un attribut est similaire à la désutilité associée au fait de ne pas en bénéficier. Cette symétrie a été étudiée au travers des scores standardisés, qui ont par exemple révélé que la désutilité associée à l'absence du listing des patients dépistés était plus forte que l'utilité associée à sa mise à disposition. D'une part, ce résultat est cohérent avec le constat selon lequel le manque d'information pour le MG constitue un frein dans son offre de dépistage ciblée (Liberalotto, 2012). D'autre part, le fait que l'utilité liée à ce dispositif soit faible (en tout cas plus faible que la désutilité de ne pas l'avoir) pourrait laisser penser que les MGs ne sont pas entièrement convaincus par la mise à disposition d'un tel listing ou du moins par les modalités de mises à disposition. Des analyses plus qualitatives seraient nécessaires pour comprendre la perception de ce dispositif et éventuellement s'il devrait être combiné à d'autres outils (aide à l'acquisition d'un logiciel de données patients, etc.).

Le deuxième intérêt de la BWAS par comparaison à la DCE est qu'elle permet d'estimer l'utilité des niveaux de chaque attribut séparément de l'impact de l'attribut dans son ensemble. Cela a permis de montrer, d'une part, que la rémunération apparaît comme l'un des dispositifs les plus plébiscités par les MGs pour les trois cancers et, d'autre part, que ceux-ci perçoivent des différences importantes entre les différents niveaux de rémunération proposés. Notons toutefois que l'effet mesuré est plus sensible aux différents niveaux proposés (1%, 3%, etc.), car il représente la moyenne des effets de chaque niveau. En revanche, dans la méthode DCE, nous avons estimé l'effet d'une augmentation de 1% des honoraires des MGs liés à l'atteinte des objectifs cibles, et il était apparu que cet effet était plus faible que, par exemple, le financement d'une formation pour le dépistage des cancers féminins.

Le troisième intérêt de la méthode est qu'elle permet de collecter davantage d'information par épreuve, augmentant de facto la capacité d'analyse de l'hétérogénéité des

préférences. Par ailleurs, le fait que cette méthode ne comprend aucune non réponse et que la totalité des médecins ait complété les épreuves a certainement engendré une plus grande puissance statistique dans l'analyse des réponses. Cela nous a notamment permis de montrer que les médecins ayant perçu un bonus ROSP pour le dépistage du cancer du sein sont davantage sensibles à l'attribut rémunération, ce résultat n'ayant pas été mis en évidence par la méthode DCE.

Les modèles estimés font l'hypothèse d'indépendance des alternatives non pertinentes (IIA). Cet axiome du modèle logit multinomial signifie que l'ajout ou le retrait d'une alternative (i.e. une paire préféré-moins préféré) ne modifie pas le rapport des préférences pour les alternatives (paires) restantes. La principale raison de violation de cette hypothèse peut provenir du fait que certains attributs sont des substituts proches (McFadden 1974). A notre connaissance, aucune étude n'a testé cette hypothèse dans le cadre des modèles de best-worst scaling, du moins appliqués au domaine de la santé (Flynn et al 2008; Flynn et al 2010). Un prolongement à ce travail consistera donc à tester cette hypothèse en comparant les résultats obtenus avec ceux d'un modèle logit emboîté (Cheng et al 2012a).

Les résultats des modèles issus de la méthode BWAS ont permis de confirmer et de compléter certains résultats obtenus dans la section précédente. Ainsi, en termes de préférences, les MGs seraient notamment sensibles à une formation financée et un listing des patients dépistés en particulier dans le cadre des dépistages des cancers du sein et du col de l'utérus. En revanche, la rémunération apparaît plus importante aux yeux des médecins dans le cadre du dépistage du cancer colorectal. En effet, le MG est l'un des acteurs les mieux placés pour dépister de ce cancer (Carcaise-Edinboro and Bradley 2008; Bungener et al 2010) mais il n'existe pas à ce jour de dispositifs incitatifs au niveau national permettant de compenser la charge de travail supplémentaire fournie. Cette absence de rémunération additionnelle pourrait donc expliquer le fait que les MGs auraient été davantage sensibilisés à ce dispositif. Une deuxième hypothèse permettant d'expliquer en partie ce résultat tiendrait au fait que l'objectif cible proposé afin de percevoir le bonus (50% de patients dépistés) était plus faible pour ce cancer que pour les deux autres (80% de patients dépistés), augmentant de fait l'espérance de gain associée au dispositif de rémunération.

La principale limite des méthodes DCE et BWAS, commune aux méthodes de révélation des préférences, est le fait que les choix sont révélés et non observés. De fait, elles ne constituent que la première étape à l'identification de dispositifs permettant d'améliorer les

pratiques de dépistage des cancers des MGs. Dans le prolongement de cette étude, la seconde étape consisterait à expérimenter (et évaluer les effets) des dispositifs mis en évidence c'est-à-dire, notamment, d'une formation financée et d'un listing des patientes dépistées prioritairement appliqué au dépistage des cancers féminins.

5. Conclusion

L'objectif de ce travail était de révéler les préférences des médecins pour des dispositifs susceptibles d'améliorer leur pratique de dépistage des cancers du sein, du col de l'utérus, et du côlon. Chacun de ces cancers fait déjà l'objet de différentes mesures incitatives, ciblant la demande d'une part (dépistages organisés, campagnes informatives), et l'offre d'autre part avec la généralisation depuis 2012 de la rémunération sur objectif de santé publique (ROSP). Sur la base d'une revue de la littérature des expériences internationales, d'une étude fine des dispositifs existants en France et d'entretiens avec des experts de terrain (médecins généralistes, Agences Régionales de Santé), nous avons identifié cinq dispositifs plus ou moins innovants, mais tous cohérents en termes de mise en œuvre, à proposer aux médecins. Ces dispositifs comprenaient des mesures de natures diverses : mise à disposition de documents à destination des patients, listing pluriannuel des patients dépistés, aide en personnel, financement d'une formation sur le dépistage et d'une mesure financière ayant le même format que la ROSP mais proposant des montants tous plus élevés. Pour révéler les préférences des médecins concernant ces dispositifs ou attributs, deux méthodologies distinctes ont été mises en œuvre : la méthodologie des choix discrets (DCE) d'une part, et la méthodologie du classement des attributs du meilleur au pire (BWAS) d'autre part. Ces deux méthodes avaient en commun, à partir d'un questionnaire construit de façon expérimentale, de proposer aux médecins plusieurs scénarios hypothétiques combinant différentes modalités des dispositifs. Puis, il leur était demandé d'effectuer des arbitrages entre ces différentes modalités, de façon à révéler *in fine* leurs préférences pour les dispositifs ou niveaux de dispositifs proposés.

Les résultats obtenus ont permis de valider l'hypothèse émise, c'est-à-dire, les MGs Français sont sensibles à des dispositifs non monétaires (ou des transferts en nature de type financement d'une formation) dans le cadre du dépistage des cancers. En outre, ils seraient prêts à renoncer à des montants conséquents de rémunération additionnelle pour en bénéficier. En particulier, pour le dépistage des cancers féminins, les MGs seraient prêts à payer « de leur poche » en plus de leur rémunération sur objectifs pour bénéficier d'une formation financée sur le dépistage ou d'un listing pluriannuel des patients dépistés. Ces résultats confirment donc les résultats de précédents travaux ayant montré l'importance des caractéristiques non monétaires de la pratique des médecins (Scott 2001; Janus et al 2007; Ammi and Peyron 2010). En outre, les dispositifs non monétaires que nous avons définis ont pour objectif de

promouvoir l'accès à l'information des médecins généralistes, le sentiment d'autonomie et de compétence rencontré dans sa relation avec le patient. Ces résultats semblent donc montrer que les motivations des médecins associées à des pratiques de qualité sont diverses et variées, qu'elles ne sont pas forcément (uniquement) d'ordre monétaires. En ce sens, il apparaît nécessaire, à la suite de nombreux travaux à la fois théoriques et empiriques, d'élargir certaines hypothèses de base de la théorie économique standard appliquée aux professionnels de santé (Ellis and McGuire 1986; Laffont and Martimort 2002) de façon à prendre en compte un registre plus large des motivations ainsi que leur interrelation (Frey 1997b; Gneezy and Rustichini 2000; Siciliani 2009; Janus 2010). Plus précisément, dans le contexte de renforcement des politiques incitatives au travers notamment des paiements à la performance, il s'agit d'analyser précisément les relations de complémentarité ou substituabilité entre les motivations propres des médecins, dites intrinsèques, et leurs motivations associées à des récompenses financières, dites extrinsèques (Deci 1971; Frey and Jegen 2001; Bénabou and Tirole 2003).

Annexes

Annexe B- Tableaux et figures additionnels

Table B.1. Matrice de corrélation entre les effets principaux et les interactions (design à 8 scénarios)

<i>Effets principaux</i>	Doc	Form	List	Aide	Rem
Doc	1				
Form	0	1			
List	0	0	1		
Aide	0	0	0	1	
Rem	0	0	0	0	1
<i>Interactions</i>					
Doc*Form	0	0	0	0	-0.45
Doc*List	0	0	0	-1	0
Doc*Aide	0	0	-1	0	0
Doc*Rem	0	-0.45	0	0	0
Form*List	0	0	0	0	-0.89
Form*Aide	0	0	0	0	0
Form*Rem	-0.45	0	-0.89	0	0
List*Aide	-1	0	0	0	0
List*Rem	0	-0.89	0	0	0
Aide*Rem	0	0	0	0	-0.8

Légende : les interactions d'intérêt sont indiquées en gras. Les coefficients de corrélations supérieurs à 0.4 en valeur absolue sont indiqués en rouge. Lorsque le seuil de 0.4 est dépassé, il est difficile d'estimer une interaction entre deux attributs indépendamment d'un effet principal. Dans ce design, la seule interaction qui pourrait être estimée est form*aide.

Table B.2. Matrice de corrélation entre les effets principaux et les interactions (design à 12 scénarios)

<i>Effets principaux</i>	Doc	Form	List	Aide	Rem
Doc	1				
Form	0	1			
List	0	0	1		
Aide	0	0	0	1	
Rem	0	0	0	0	1
<i>Interactions</i>					
Doc*Form	0	0	0.33	0.33	0.15
Doc*List	0	0.33	0	0.33	-0.15
Doc*Aide	0	0.33	0.33	0	-0.15
Doc*Rem	0	0.15	-0.15	-0.15	-0.27
Form*List	0.33	0	0	0.33	-0.45
Form*Aide	0.33	0	0.33	0	0.45
Form*Rem	0.15	0	-0.45	0.45	0.27
List*Aide	0.33	0.33	0	0	0.15
List*Rem	-0.15	-0.45	0	0.15	0.27
Aide*Rem	-0.15	0.45	0.15	0	0.27

Lecture : Dans ce design, aucun coefficient de corrélation entre les interactions d'intérêt doc*rem et list*aide et les effets principaux n'est supérieur à 0.4 en valeur absolue.

Figure B.1. Présentation du déroulement des épreuves (dépistage du cancer du sein)

Pour chacune de ces situations, nous vous demanderons :

- 1/ Dans un premier temps, d'indiquer si vous changeriez votre pratique habituelle pour la situation proposée.
- 2/ Dans un second temps, de sélectionner:
 1. Ce que **vous préférez le plus** dans la situation proposée.
 2. Ce que **vous préférez le moins** dans la situation proposée.

IMPORTANT

Pour répondre, rappelez-vous :

- 1/ Que cela concerne le dépistage du **cancer du sein pour vos patientes âgées entre 50 et 74 ans.**
- 2/ L'objectif de ce questionnaire est **d'identifier vos préférences** pour différents dispositifs (existants ou non) d'accompagnement de vos activités de dépistage du cancer du sein.
- 3/ L'aspect répétitif de l'exercice fait partie de la recherche. **Merci de rester vigilant jusqu'à la dernière situation.**

Figure B.2. Présentation d'une épreuve DCE

Voici la présentation de la situation hypothétique suivante

Situation hypothétique
Vous bénéficiez de brochures d'information
Vous bénéficiez d'une formation financée
Vous <u>ne</u> bénéficiez <u>pas</u> d'un résumé semestriel de vos patientes dépistées
Vous <u>ne</u> bénéficiez <u>pas</u> d'aide ponctuelle en personnel qualifié
Vous bénéficiez d'une augmentation de 3% de vos honoraires sur objectif

Cliquez [ici](#) pour revoir le descriptif des 5 dispositifs

S7B

Changeriez-vous votre pratique habituelle pour la situation proposée ?

- Oui
- Non
- Ne sait pas

Figure B.3. Présentation d'une épreuve BWAS

A présent, vous devez sélectionner :

Une seule réponse par colonne

Parmi les 5 propositions suivantes ...	Ce que vous préférez le plus	Ce que vous préférez le moins
Vous bénéficiez de brochures d'information	<input type="radio"/>	<input type="radio"/>
Vous bénéficiez d'une formation financée	<input checked="" type="radio"/>	<input type="radio"/>
Vous <u>ne</u> bénéficiez <u>pas</u> d'un résumé semestriel de vos patientes dépistées	<input type="radio"/>	<input type="radio"/>
Vous <u>ne</u> bénéficiez <u>pas</u> d'aide ponctuelle en personnel qualifié	<input type="radio"/>	<input checked="" type="radio"/>
Vous bénéficiez d'une augmentation de 3% de vos honoraires sur objectif	<input type="radio"/>	<input type="radio"/>

Cliquez [ici](#) pour revoir le descriptif des 5 dispositifs

Table B.3. Caractéristiques des médecins « non traders »

	Cancer du sein		Cancer du col		Cancer colorectal	
	N	row %	N	row %	N	row %
Genre	p = 0.3848†		p = 0.8134		p = 0.8915	
Homme	16	16.7	14	14.9	14	14.4
Femme	9	23.1	6	16.2	5	13.5
Age	p = 0.3338		p = 0.6667		p = 0.2813	
< 45 ans	3	14.3	4	14.3	6	18.8
[45-54] ans	7	13.7	8	19.1	4	8.0
>=55 ans	15	23.8	8	12.7	9	17.3
Zone géographique	p = 0.0181		p = 0.6567		p = 0.0149	
Ile de France	4	18.2	4	16.7	4	15.4
Ouest	4	14.3	6	21.4	6	23.1
Nord	4	50.0	1	11.1	0	0.0
Est	4	19.1	4	20.0	7	35.0
Sud Ouest	1	5.6	3	18.8	0	0.0
Centre Est	5	50.0	1	5.9	0	0.0
Mediterranee	3	10.7	1	5.3	2	8.3
Taille de la commune	p = 0.4734		p = 0.8822		p = 0.3060	
Moins de 10 000 habitants	10	22.7	7	15.9	7	13.7
De 10 à 100 000 habitants	9	20.0	8	18.6	5	9.8
Plus de 100 000 habitants	6	13.0	5	10.9	7	21.9
Année d'installation pour la première fois	p = 0.2067		p = 0.3338		p = 0.8155	
Moins de 15 ans	3	9.4	4	12.5	6	16.2
De 16 à 29 ans	15	24.2	9	15.3	10	14.5
Plus de 30 ans	7	17.1	7	16.7	3	10.7
Suivez-vous des cycles de formation continue ?	p = 0.7672		p = 0.1810		p = 0.2940	
Oui	22	18.2	16	13.6	18	15.4
Non	3	21.4	4	26.7	1	5.9
Avez-vous déjà suivi une formation sur le dépistage des cancers ?	p = 0.8567		p = 0.9285		p = 0.9786	
Oui	12	17.9	12	14.8	12	14.1
Non	13	19.2	8	15.4	7	14.3
Vous exercez dans un ...	p = 0.1654		p = 0.7619		p = 0.7506	
Cabinet individuel	17	22.7	11	15.9	10	15.2
Cabinet de groupe de médecins	8	13.3	9	14.1	9	13.2
Secteur d'exercice	p = 0.7672		p = 0.9337		p = 0.8957	
Secteur 1	22	18.2	18	15.1	14	14.1
Secteur 2	3	21.4	2	14.3	2	15.4
Combien d'heures travaillez-vous en moyenne par semaine ?	p = 0.3836		p = 0.2858		p = 0.9184	
<= 40 heures	6	30.0	3	12.0	4	17.4
[41-50] heures	8	18.6	5	11.4	6	12.8
[51-60] heures	9	18.0	10	23.8	6	15.8
>60 heures	2	9.1	2	9.1	3	11.5
Combien de demi-journées travaillez-vous en moyenne par semaine ?	p = 0.4359		p = 0.6872		p = 0.4161	
<10 demi-journées	9	15.5	10	16.4	8	11.8
>=10 demi-journées	16	20.8	10	13.9	11	16.7
Non traders : N (%)	25	(18.5)	20	(15.0)	19	(14.2)

† P-value du test d'indépendance du Khi-deux.

Lecture : Parmi les hommes, 16.7% sont catégorisés comme « non-traders » alors que ces derniers représentent 18.5% de l'échantillon des individus ayant répondu au questionnaire portant sur le dépistage du cancer du sein.

Table B.3. Caractéristiques des médecins « non traders » (suite)

	Cancer du sein		Cancer du col		Cancer colorectal	
	N	row %	N	row %	N	row %
Combien de patients voyez-vous en moyenne au cours d'une semaine normale ?	p = 0.3099		p = 0.6771		p = 0.7231	
<100 patients	9	26.5	3	10.7	5	18.5
[100-149] patients	10	18.2	7	14.3	9	14.1
>=150 patients	6	13.0	10	17.9	5	11.6
Proposez-vous un dépistage du cancer considéré (sein, col de l'utérus, ou côlon)	p = 0.1173		p = 0.5166		p = 0.6462	
Systématiquement	13	23.6	7	19.4	6	15.4
Très souvent	6	14.6	4	8.7	9	17.0
Souvent	3	9.4	7	18.0	4	11.1
Rarement, Jamais	3	42.9	2	16.7	0	0.0
Utilisez-vous des brochures d'information pour le cancer considéré (sein, col de l'utérus, ou côlon)	p = 0.8661		p = 0.5404		p = 0.8661	
Oui	10	19.2	5	12.2	12	16.7
Non	15	18.1	15	16.3	7	11.3
Avez-vous perçu une prime ROSP pour le dépistage du cancer du sein ?	p = 0.8032		-		-	
Oui	10	17.5	-	-	-	-
Non	15	19.2	-	-	-	-
Avez-vous perçu une prime ROSP pour le dépistage du cancer du col de l'utérus ?	-		p = 0.0001		-	
Oui	-	-	12	35.3	-	-
Non	-	-	8	8.1	-	-
Non traders N (%)	25	(18.5)	20	(15.0)	19	(14.2)

Table B.4. Résultats des modèles logit conditionnels à effets fixes avec la variable rémunération en classes

	Cancer du sein			Cancer du col			Cancer colorectal		
	Coeff	Std. Err.	P	Coeff	Std. Err.	P	Coeff	Std. Err.	P
Effets principaux									
Doc	0,320	0,149	0,032	0,472	0,145	0,001	0,222	0,152	0,143
Form	1,157	0,199	0,000	1,030	0,180	0,000	0,780	0,184	0,000
List	1,265	0,216	0,000	1,188	0,209	0,000	0,850	0,186	0,000
Aide	0,335	0,199	0,093	0,689	0,169	0,000	0,149	0,169	0,380
Rem					<i>ref = 0%</i>				
Rem_1%	1,194	0,245	0,000	0,817	0,209	0,000	1,031	0,244	0,000
Rem_3%	1,804	0,278	0,000	1,500	0,253	0,000	2,165	0,274	0,000
Rem_5%	2,240	0,281	0,000	1,964	0,269	0,000	2,757	0,293	0,000
Obs		1189			1231			1256	
LogL		-406,270			-444,658			-443,143	
Pseudo R2		0,2237			0,2108			0,2391	

Figure B.4. Impact du niveau de rémunération sur la probabilité d'accepter le scénario proposé (estimations de la Table B.4.)

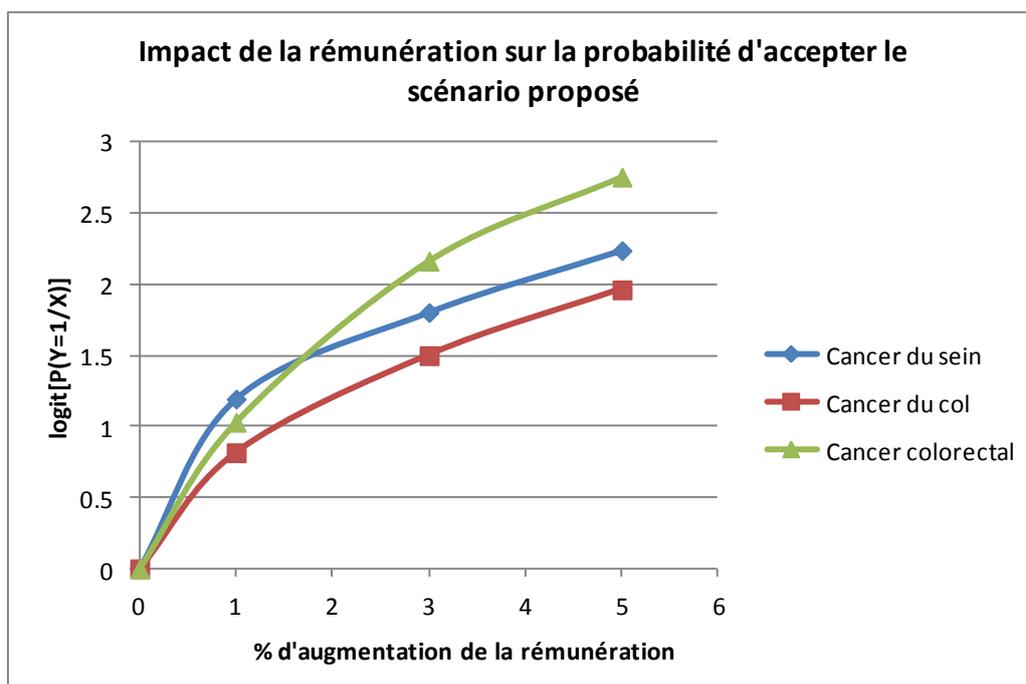


Table B.5. Résultats des modèles logit conditionnels multinomiaux marginaux

	Cancer du sein			Cancer du col de l'utérus			Cancer colorectal		
	<i>Coeff</i>	<i>Std. Err.</i>	<i>P</i>	<i>Coeff</i>	<i>Std. Err.</i>	<i>P</i>	<i>Coeff</i>	<i>Std. Err.</i>	<i>P</i>
bwindic	-0.224	0.049	0.000	-0.268	0.050	0.000	-0.232	0.051	0.000
Attribut									
Doc	-0.032	0.071	0.656	0.116	0.071	0.105	0.034	0.073	0.640
Form	0.333	0.072	0.000	0.378	0.075	0.000	0.259	0.073	0.000
List	0.087	0.070	0.213	0.193	0.074	0.009	0.083	0.074	0.261
Aide	-		-	-		-	-		-
Rem	0.715	0.080	0.000	0.641	0.081	0.000	0.762	0.081	0.000
Niveau									
Doc_oui	0.209	0.045	0.000	0.326	0.043	0.000	0.519	0.045	0.000
Doc_non	-0.209		-	-0.326		-	-0.519		-
Form_oui	0.988	0.046	0.000	1.098	0.047	0.000	0.947	0.046	0.000
Form_non	-0.988		-	-1.098		-	-0.947		-
List_oui	1.021	0.045	0.000	0.992	0.045	0.000	1.050	0.048	0.000
List_non	-1.021		-	-0.992		-	-1.050		-
Aide_oui	0.275	0.048	0.000	0.361	0.050	0.000	0.137	0.051	0.007
Aide_non	-0.275		-	-0.361		-	-0.137		-
Rem_1%	-0.044	0.094	0.637	0.023	0.093	0.808	0.003	0.092	0.971
Rem_3%	0.852	0.092	0.000	1.014	0.091	0.000	1.082	0.090	0.000
Rem_5%	1.106	0.091	0.000	0.916	0.091	0.000	1.262	0.090	0.000
Rem_non	-1.913		-	-1.952		-	-2.347		-

Chapitre 3

**Motivations intrinsèques et extrinsèques en
médecine générale: complémentarité ou
substituabilité ?**

1. Introduction

1.1. Problématique et objectifs

Les paiements à la performance essentiellement appliqués aux soins primaires ont été introduits dans de nombreux pays de l'OCDE (Lindenauer et al 2007). L'objectif principal de ces mécanismes incitatifs est de permettre une amélioration de la qualité et de l'efficacité des pratiques (Hughes 1993; Petersen et al 2006; Lazear and Shaw 2007; Mannion and Davies 2008). Les paiements à la performance (P4P) peuvent par ailleurs permettre d'atteindre plusieurs autres objectifs indirects ou secondaires tels que l'augmentation de la rémunération des médecins, l'amélioration de l'informatisation des cabinets et du retour d'information (Lanier et al 2003) ainsi qu'une réduction de la variabilité des pratiques (Dormont 2013). Pourtant, malgré leur place croissante dans la rémunération des médecins et leur adoption par de nombreux pays, des questions subsistent. Ces dispositifs de P4P, dits extrinsèques, n'ont pas clairement fait la preuve de leur capacité à améliorer la qualité ou l'efficacité des soins de premiers recours (Croxson et al 2001; Rosenthal and Frank 2006; Campbell et al 2009; Maynard 2012; Eijkenaar et al 2013; Li et al 2014). Dans le même temps, ce type de mécanisme incitatif peut également générer des effets indésirables : le risque de sélection de patients ou des pathologies les plus rémunérateurs/rémunératrices est parmi les effets indésirables les plus cités (Giuffrida et al 2000; Gosden et al 2001; Doran et al 2006; Mannion and Davies 2008), comme l'augmentation du stress au travail (Hadley et al 1999; Gené-Badia et al 2007) voire l'apparition de phénomènes de burn-out (Smith 2001; Gené-Badia et al 2007). Toutefois, les résultats attestant des réactions stratégiques des médecins concernant notamment la sélection des patients restent limités (Doran et al 2008).

Un autre effet indésirable potentiel des paiements de type P4P, étudié dans de nombreux travaux d'abord en psychologie puis en économie, est le risque d'éviction des motivations intrinsèques (**MI**) par les récompenses extrinsèques (Deci 1971; Frey 1997; Gibbons 1998; Deci et al 1999a; Bénabou and Tirole 2003; Sliwka 2006; Bowles 2008; Feng et al 2011; Schnedler and Vanberg 2014). Les motivations extrinsèques (**ME**) se réfèrent à des activités menées dans le but d'obtenir une récompense extérieure à l'activité comme par exemple, une rémunération ou un bonus. Selon certains auteurs, cet effet d'éviction des MI par les ME est d'autant plus plausible pour les MGs libéraux dans la mesure où la pratique libérale sous tend un profond désir d'autonomie au travail et que dans le même temps, la médecine est soumise à l'existence de normes professionnelles qui influencent et orientent les

pratiques (Videau et al 2010). Il y a deux principales conceptions des MI. La théorie psychologique d'une part, suppose que les individus retirent une satisfaction directe d'une activité considérée comme intrinsèquement motivante et gratifiante (Deci 1971). Les chercheurs en sciences sociales font l'hypothèse que les individus sont motivés à l'idée de jouer un rôle social dans un système organisé, leur permettant par exemple, d'accomplir une norme ou une éthique professionnelle. Il a été montré que les MI ont un effet bénéfique sur l'engagement au travail et la qualité des soins (Benware and Deci 1984; Janus 2010). D'autre part, certains travaux ont montré que les MI laissent plus de place à la créativité dans la résolution de tâches complexes, alors que les ME tendent à produire des solutions plus stéréotypées (Amabile et al 1976; Schwartz 1990). Enfin, les MI permettent la génération et la diffusion de connaissances dans des conditions où les ME ne se révéleraient pas efficaces (Osterloth and Frey 2000).

Dans le contexte de renforcement des schémas incitatifs appliqués aux soins de premiers recours, il apparaît indispensable de s'intéresser aux concepts de MI et ME ainsi qu'à leur interrelation. Tout d'abord, cette analyse est justifiée dans le contexte français du fait de l'organisation de la médecine libérale : historiquement, les médecins et les syndicats de médecins ont cherché à maintenir leur autonomie vis-à-vis du financeur public, et restent très attachés au paiement à l'acte (Samson 2009; Ammi and Peyron 2010). Ainsi les paiements à la performance sont perçus chez certains médecins comme une menace à la place dominante du paiement à l'acte en tant que système de rémunération de la médecine libérale (Saint-Lary et al 2011). Par ailleurs, les résultats d'études précédentes ont montré une relation complexe entre MI et ME des médecins généralistes (MGs) français. En effet, considérer que les médecins sont des acteurs purement rationnels cherchant à maximiser leurs revenus ne permet pas de comprendre pourquoi seulement un tiers d'entre eux ont accepté de signer le CAPI qui pourtant ne comportait que des bonus, et aucune sanction (Saint-Lary et al 2013). De plus, les résultats du Chapitre 1 ont montré que les MGs signataires du CAPI avaient en moyenne une durée de consultation plus faible. Ce résultat pourrait être la conséquence d'un arbitrage différent entre quantité et qualité, mais aussi entre ME et MI, puisque les MI sont souvent approximées par la durée pendant laquelle un individu effectue une activité (Deci 1971; Deci et al 1999a; Videau et al 2010). Enfin, les résultats de l'enquête de choix discrets effectuée dans le Chapitre 2 ont montré l'existence d'une grande hétérogénéité des préférences des MGs concernant les dispositifs à mettre en œuvre pour améliorer leurs pratiques. Ces résultats semblent indiquer qu'il existe des comportements ou une forme des

préférences qui contredisent ou du moins invitent à élargir les hypothèses de base des modèles standards appliqués aux médecins (Ellis and McGuire 1986; Rochaix 1989; Laffont and Martimort 2002). Ainsi, il semble nécessaire d'analyser précisément les relations entre MI et ME des médecins généralistes, de façon à mieux anticiper les effets indésirables potentiels d'un renforcement des ME par le biais d'incitations financières. L'objectif de ce chapitre est donc, dans un premier temps, de construire une mesure empirique de la MI des médecins généralistes français et, dans un second temps, de tester la relation de complémentarité ou substituabilité entre ce score et différents indicateurs de ME.

Dans un premier temps, nous effectuons une revue de la littérature du concept de MI et des théories de l'effet d'éviction des MI par les ME. Puis, nous présentons dans une deuxième section l'enquête réalisée et les questions permettant de mesurer les MI des médecins dans le domaine particulier de la médecine générale. Dans la troisième section, nous comparons différentes méthodes permettant de construire un score composite de la MI, et sélectionnons le modèle le plus adapté à nos données. Enfin, à partir du score obtenu de MI, nous étudions dans une dernière section les relations de complémentarité ou substituabilité entre MI et ME des médecins généralistes français.

1.2. Revue de la littérature

1.2.1. Définitions du concept de motivations intrinsèques

1.2.1.1. Les fondements psychologiques

Le concept de motivation intrinsèque (MI) est né des travaux en psychologie dans les années 60, avec le développement de deux théories : la « self-determination theory » (SDT) et la « cognitive evaluation theory » (CET). Selon la SDT, trois nécessités psychologiques sont essentiels pour comprendre le contenu et le processus de la poursuite d'un but : le besoin d'autonomie, le besoin de compétence et le besoin de lien ou d'échange (Deci and Ryan 2000). Les activités intrinsèquement motivantes sont définies comme des activités que les individus trouvent intéressantes et feraient en l'absence de conséquences extérieures séparables³⁸. Ainsi, Deci (1971) émet l'hypothèse que les comportements régis par les motivations intrinsèques sont issus de la nécessité pour les individus de se sentir compétents

³⁸ Selon la définition de Deci (1971), *'one is said to be intrinsically motivated to perform an activity when one receives no apparent reward except the activity in itself'*.

et auto-déterminés³⁹. L'hypothèse centrale de la SDT est que les motivations intrinsèques augmenteront dans les conditions permettant la réalisation de ces deux nécessités psychologiques ; revanche, elles seront évincées dans les conditions tendant à contrarier la réalisation de ces nécessités. Les analyses empiriques ont montré que le lien/l'échange ('relatedness') joue aussi un rôle sur le maintien des motivations intrinsèques, bien que ce lien soit plus ténu⁴⁰.

1.2.1.2. Les différentes interprétations des MI en sciences sociales et leur application dans de domaine de la santé

Pour comprendre le concept de MI, la psychologie organisationnelle considère que les individus dérivent une utilité de l'activité elle-même, qui leur permet par exemple d'accomplir une norme sociale et / ou de développer une culture professionnelle (Argyris 1964). Plus précisément, elle distingue deux types de motivations intrinsèques (Lindenberg 2001) : d'une part, la motivation basée sur le plaisir (« enjoyment-based motivation ») et d'autre part, la motivation basée sur l'engagement (« obligation-based motivation»). La motivation basée sur le plaisir que procure l'activité en elle-même et sa réalisation, se réfère à la vision psychologique traditionnelle synthétisée par Deci et Ryan (2000). Dans le domaine de la santé, elle peut faire référence à la motivation de guérir les patients qui est une source d'épanouissement et d'inspiration (Janus 2010). La motivation basée sur l'engagement correspond au fait de poursuivre des objectifs prédéfinis par soi-même (Loewenstein 1999) ou d'adhérer à des normes personnelles ou sociales tels que les codes de déontologie en médecine (March 1999). Des travaux empiriques mettent en évidence d'autres dimensions telles que la volonté de contribuer au bien commun d'une organisation ou d'une communauté (Frey 1997a).

Ces théories permettent notamment d'expliquer le fait que les individus acceptent de suivre des règles limitant leur intérêt personnel tant qu'ils en acceptent la légitimité (Simon

³⁹ Deci (1971) complète la définition des motivations intrinsèques en l'associant à la réalisation de ces deux nécessités psychologiques: '*Intrinsically motivated behaviors are those that are freely engaged out of interest without the necessity of separable consequences, and, to be maintained, they require satisfaction of the needs for autonomy and competence*'. Cette définition va dans le même sens que l'affirmation de deCharms (1968), pour qui la motivation première des individus est de ressentir qu'ils sont auteurs de leurs actions : « *people have a primary motivational propensity to feel like causal agents with respect to their own actions* ».

⁴⁰ Ainsi, l'étude d'Anderson et al (1976) a montré que lorsque des enfants travaillaient sur une activité intéressante en présence d'un adulte participant à l'expérience mais qui ignorait ses tentatives d'interagir, les enfants montraient un niveau très faible de motivations intrinsèques comparé au groupe où l'adulte interagissait (Anderson et al 1976).

1991; Osterloh et al 2002). Dans le domaine de la santé, cela permet notamment d'expliquer certains comportements de médecins comme par exemple les médecins qui auscultent des patients en dehors des heures de travail sans demander de paiement additionnel (Janus 2010). Malgré ces controverses, les ME sont parfois indispensables dans les situations où les MI n'existent pas en premier lieu ou sont insuffisantes, afin d'augmenter les performances (Lazear and Shaw 2007).

1.2.2. Les théories du crowding-out effect

1.2.2.1. Fondements psychologiques

C'est à la suite de deux expériences de laboratoire ayant montré l'effet contre-performant de l'instauration de récompenses monétaires sur les MI de collégiens à effectuer une activité ludique, que Deci (1971) a formulé la 'cognitive evaluation theory' (CET). Cette théorie postule que, lorsque des récompenses externes sont introduites pour effectuer une activité intéressante intrinsèquement, les individus tendent à se sentir « contrôlés » par les récompenses, entraînant une modification du motif ('locus') de causalité perçue, celui-ci passant d'interne à externe. Ce changement de perception de la finalité de l'action se traduit en retour par une baisse des motivations intrinsèques (DeCharms 1968; Deci and Ryan 2000). Plus précisément, la théorie prédit que les effets seront en moyenne différents selon le type de récompenses. Les récompenses non contingentes à la tâche effectuée n'auront pas d'effet négatif sur les MI car elles ne portent pas atteinte aux sentiments d'autonomie et de compétence perçus ; en revanche, des récompenses contingentes à la performance (incitatifs économiques, menaces, sanctions) auront plus de chance d'être perçues comme « contrôlantes », et diminueront les MI (Ledford et al 2013).

D'autres auteurs ont formulé une théorie concurrente à l'effet d'éviction des MI par les ME : 'l'overjustification effect' (Lepper et al 1973). Cette théorie utilise un cadre théorique appelé « attribution theory », qui postule que les individus interprètent leurs actions de la même façon que le ferait un acteur externe, c'est-à-dire, ils étudient de façon rétrospective, les raisons et les motifs pour lesquels ils agissent. Ainsi, en l'absence de compensation extérieure, l'individu interprétera son motif d'action comme issu de sa motivation intrinsèque. Si une compensation monétaire lui est proposée, il interprétera sa motivation comme étant le gain monétaire, et sa motivation intrinsèque sera plus faible que

s'il n'y avait pas de récompense. Cette théorie a reçu notamment beaucoup d'attention dans le domaine de l'éducation (Tang and Hall 1995)⁴¹.

1.2.2.2. Interprétations en économie

Frey est l'un des premiers auteurs à avoir formulé la théorie du 'motivational crowding-out effect' (MCO) dans un cadre économique (Frey 1994, 1997). Dans ce but, il pointe tout d'abord la divergence des visions économiques et psychologiques relatives à l'impact des incitatifs économiques sur le comportement humain. Le modèle économique du comportement humain est basé sur des incitatifs appliqués en dehors de la personne considérée, et les individus modifient leurs actions parce qu'ils y sont conduits par une intervention extérieure (Becker 1976; Frey 1999). Ainsi, les incitations économiques mises en œuvre sont sensées pallier les problèmes d'asymétrie d'information et d'aléa moral, qui résultent du fait que l'effort des agents est impossible à observer et que les coûts de contrôle a posteriori sont très élevés (Arrow 1963; Gibbons 1998). En ce sens, les incitatifs économiques permettent de compenser le coût de l'effort associé à la réalisation d'une tâche ou d'un travail, souvent considéré comme source de désutilité dans les modèles de type principal-agent (Laffont and Martimort 2002). Enfin, si ces modèles ne nient pas l'existence de MI, ils font du moins l'hypothèse que MI et ME sont indépendants.

Les psychologues, en revanche, mettent l'accent sur les motivations comportementales (« behavioural motives ») provenant de la personne elle-même. L'application de récompenses pour mener une activité a des conséquences négatives indirectes et comporte des coûts cachés, parmi lesquels, une baisse du sentiment d'auto-détermination et d'estime de soi (Greene and Lepper 1978). Dans cette approche, il y a donc interrelation entre les différents registres de motivation.

Les effets à attendre d'une intervention extérieure telle qu'une récompense monétaire peuvent être synthétisés dans le cadre d'un modèle simple de type principal-agent (Frey 1997b). Dans ce modèle, le principal (par exemple le financeur public) décide de déléguer une tâche à un agent (le médecin), celle-ci consistant à améliorer l'état de santé des patients.

⁴¹ La meta-analyse de Tang et Hall basée sur une revue de 50 articles qui confirme la théorie, mais son extrapolation reste limitée car ces études ont été effectuées sur des enfants ou des collégiens. Selon Deci et al (1999), les effets négatifs sur la motivation intrinsèques peuvent être plus marqués chez les enfants que chez les adultes, ces derniers ayant une plus grande capacité à distinguer les aspects informationnels des aspects contrôlant des récompenses.

Touefois, le principal ne détient pas toutes les informations concernant la réalisation de la tâche et le niveau d'effort fournis, et ne peut pas observer directement la qualité. L'agent effectue l'action en considérant les bénéfices (**B**) et les coûts associés (**C**). Les bénéfices et les coûts augmentent lorsque la performance (**P**) augmente avec des hypothèses standards, i.e. $\partial B/\partial P \equiv B_p > 0$ avec $B_{pp} < 0$ et $\partial C/\partial P \equiv C_p > 0$ avec $C_{pp} > 0$. On a :

$$B = B(P, E); B_p > 0, B_{pp} < 0 \quad (1)$$

$$C = C(P, E); C_p > 0, C_{pp} > 0 \quad (2)$$

Le principal veut utiliser une intervention extérieure (**E**) afin d'augmenter la performance de l'agent. L'agent rationnel cherche à maximiser son bénéfice net, conduisant à la condition de premier ordre :

$$B_p = C_p \quad (3)$$

Ainsi, un agent augmente sa performance jusqu'à égalisation du bénéfice marginal et du coût marginal. La performance optimale P^* de l'agent est modifiée lorsque le principal change le niveau de l'intervention extérieure E . Cette modification est exprimée mathématiquement en différenciant l'équation (3) par rapport à E :

$$B_{pE} + B_{pp} \frac{dP^*}{dE} = C_{pE} + C_{pp} \frac{dP^*}{dE}$$

$$\frac{dP^*}{dE} = \frac{B_{pE} - C_{pE}}{C_{pp} - B_{pp}}$$

Selon Frey (1997), trois cas peuvent être distingués :

- (a) Selon le modèle principal-agent standard, l'intervention extérieure augmente la performance, en diminuant le coût marginal (effet « prix ») et en augmentant le bénéfice marginal de l'effort⁴² : $C_{pE} < 0$ et $B_{pE} > 0$. De l'autre côté, l'effet d'éviction des motivations est négligé, i.e. l'intervention n'affecte pas négativement le bénéfice

⁴² Notons que les interprétations que nous faisons sont différentes de celles de Frey (1997), qui ne considère pas que l'intervention extérieure affecte le bénéfice marginal de l'agent ($B_{pE} = 0$). Nous considérons qu'une récompense monétaire peut quand même affecter le bénéfice extrinsèque retiré d'une action, mais pas son bénéfice intrinsèque. Toutefois, distinguer les deux types de bénéfices conduirait à des complexifications du modèle qui dépasse l'objet de cette thèse (James Jr. 2005; Siciliani 2009; Schnedler and Vanberg 2014).

marginal de la performance. Ainsi, l'intervention extérieure augmente toujours la performance : $\frac{dP^*}{dE} > 0$.

(b) Au contraire, si l'intervention est perçue comme contrôlante, elle va diminuer la motivation intrinsèque, ce qui va affecter en retour le bénéfice marginal de l'agent. Si la réduction du bénéfice intrinsèque de l'activité est plus forte que l'augmentation du bénéfice extrinsèque, alors $B_{PE} < 0$ (« Crowding-out effect »). Par ailleurs, si l'effet prix/disciplinant ne fonctionne pas ($C_{PE} = 0$), alors l'agent pourrait réduire son niveau de performance : $\frac{dP^*}{dE} < 0$

(c) En général, l'effet prix et l'effet d'éviction fonctionnent, de sorte que l'efficacité de l'intervention du principal dépendra de la taille relative de ces deux effets.

Dans un article de référence, Benabou et Tirole (2003) s'intéressent à la problématique de l'effet '*motivation crowding out*' (MCO) dans une modélisation économique ayant un double objectif : analyser les coûts cachés (« hidden costs ») des récompenses ou sanctions économiques et donner un contenu précis aux notions de MI et de ME. Les auteurs se placent aussi dans le cadre d'analyse de la relation d'agence : les auteurs formalisent l'incitation comme un signal perçu par l'agent de la difficulté de la tâche à accomplir, et d'un manque de confiance de la part du principal. Les auteurs en dérivent plusieurs propositions intéressantes : en situation d'asymétrie d'information, les récompenses (ME) permettent souvent de générer un équilibre de second rang, mais à court terme car ces mêmes ME diminuent l'appréciation que l'agent se fait de l'attractivité de la tâche à réaliser, et diminue sa confiance envers le principal.

De nombreux autres travaux en économie se sont intéressés à la théorie de l'effet MCO (Kreps 1997; Gneezy and Rustichini 2000; James Jr. 2005; Bowles 2008; Siciliani 2009). Plus particulièrement, certains économistes ont cherché une explication à la théorie de l'effet d'éviction des MI par les ME dans la théorie des contrats incomplets (Kreps 1997; Gneezy and Rustichini 2000; Bénabou and Tirole 2005). Cette théorie postule que la plupart des contrats ne peuvent définir toutes les conditions/clauses associées au travail d'un agent, laissant place à une part de compréhension et d'interprétation propre à ce dernier. Cette théorie est ainsi mobilisée afin d'expliquer les résultats d'une expérience auprès d'étudiants ayant montré les effets délétères de l'instauration d'incitatifs monétaires sur la performance telle que mesurée par le taux de réponses correctes à un test de QI (Gneezy and Rustichini

2000)⁴³. La relation entre motivation intrinsèque et incomplétude contractuelle a été par exemple synthétisée par Videau (2010, IV, p 103-106), qui montre comment les différents travaux économiques se rejoignent en identifiant un lien réciproque entre l'existence de contrat incomplet et la présence de motivation intrinsèque chez les agents (Kreps 1997; Falk et al 1999; Gneezy and Rustichini 2000; Fehr and Falk 2002). Ce type de contrat se retrouve d'ailleurs plus souvent dans les secteurs à but non lucratif où les valeurs autres que pécuniaires sont mises en évidence, pouvant parfois justifier des rémunérations plus faibles (Francois 2000; Heyes 2005; Narcy 2011).

Dans les analyses empiriques, pour valider et éventuellement quantifier cet effet de MCO, certains auteurs ont utilisé la fourniture d'un travail non payé comme un proxy de la motivation pour le service public (Bell and Hart 1999; Gregg et al 2011). Dans le domaine de la santé, Feng *et al* (2011), par exemple, utilisent l'offre de services additionnels non rémunérés dans le cadre du *Quality and Outcomes Framework* (P4P pour les médecins au Royaume-Uni) comme proxy de la motivation intrinsèque des médecins. Ils mettent en évidence un effet d'éviction des MI, c'est-à-dire, une augmentation du revenu lié au QOF d'une livre sterling réduit l'indicateur de MI calculé de 0.12%.

La nécessité de prendre en compte les différents registres de motivation des agents lors de la définition d'incitatifs visant à l'amélioration des pratiques semble donc justifiée d'un point de vue théorique comme empirique. Ainsi, dans la suite de ce chapitre, notre objectif est d'étudier empiriquement l'interrelation entre motivations intrinsèques (MI) et motivations extrinsèques (ME). Dans la suite de cette recherche, nous adoptons donc une démarche empirique afin d'une part de fournir une mesure de la MI des médecins généralistes et, d'autre part, de tester l'hypothèse de substituabilité entre MI et ME des MGs français.

⁴³ L'expérience a concerné 160 étudiants américains en premier cycle universitaire. Un paiement fixe était proposé initialement aux étudiants pour leur participation, puis ces derniers étaient séparés en trois groupes, le premier ne recevant pas de récompense, le second recevant 10 centimes par bonne réponse et le troisième recevant 1\$ par bonne réponse. Les résultats de l'expérience étaient les suivants : le premier groupe réussissait en moyenne mieux que le deuxième, et le troisième groupe mieux que le premier. L'interprétation faite par les auteurs de l'étude est que l'introduction des récompenses monétaires change la perception du contrat : le paiement fixe n'est plus donné pour participer à l'expérience mais pour « se montrer ». L'activité de réponse aux questions est maintenant payée « à la pièce », et ce taux devient le point de référence pour décider de l'intensité appropriée de l'effort à fournir : lorsque ce taux est faible l'effort fourni est plus faible que lorsque ce taux est élevé (Gneezy and Rustichini 2000b).

2. Données utilisées pour la mesure des motivations

2.1. Description de l'enquête

Une enquête *ad hoc* par questionnaire a été réalisée en 2010 auprès de médecins généralistes (MGs) exerçant en cabinet libéral dans le sud-ouest de la France (région Midi-Pyrénées) dans le cadre du partenariat de recherche sur l'organisation des soins de premiers recours (PROSPERE)⁴⁴. Le questionnaire a été construit avec l'objectif d'étudier l'organisation de la pratique des MGs, leurs activités et leur satisfaction au travail. Les questions ont été élaborées en collaboration avec les MGs de la Société Française de Médecine Générale (SFMG). L'échantillon comprenait environ 3000 médecins répertoriés dans la base de données de l'union régionale des médecins libéraux (URML) de la région Midi-Pyrénées. La taille de l'échantillon cible était de 450 répondants. L'enquête a été menée en collaboration avec l'union régionale de médecins libéraux de Midi-Pyrénées. Une lettre d'information présentant l'étude a été envoyée à la mi-mars 2010, deux semaines avant que le questionnaire soit envoyé par mail. Une lettre de rappel a été envoyée à tous les MGs trois semaines après l'envoi du questionnaire. Par ailleurs, un système d'anonymisation et de renvoi par enveloppe pré-timbrée avaient été mis en place et les médecins ont pu être compensés à hauteur de 33€ tout en préservant l'anonymat des questionnaires. A la fin du mois de juillet 2010, 438 questionnaires avaient été reçus, parmi lesquels 15 comportaient un nombre trop élevé de non réponses. Ainsi, 423 questionnaires ont pu être analysés. L'échantillon ainsi constitué était représentatif des MGs français en termes d'âge (la moyenne était de 52 ans vs. 50 ans), de pratique en groupe (58% vs. 54%) et de revenu net moyen (71 364€ vs. 71 690€) (Bellamy 2010).

Le questionnaire comporte 68 items divisés en 4 parties (cf. Annexe C.1 pour une présentation du questionnaire). La première partie concerne l'organisation de la pratique et comporte des questions concernant notamment la nature de l'exercice, la pratique en groupe ou la participation à la permanence des soins. La seconde partie s'intéresse au contenu de l'activité et aux caractéristiques de la patientèle (modes d'exercice particulier, activités de prévention, nombre de consultation, revenu, etc.). La troisième partie comprend des questions relatives à la satisfaction des médecins et leurs contraintes professionnelles. La dernière partie

⁴⁴ Ce projet a associé entre 2009 et 2013 trois entités distinctes et complémentaire : le CERMES3, l'IRDES et la SFMG autour de problématiques communes liées à la performance des organisations des soins de premiers recours.

est dédiée aux questions additionnelles d'ordre sociodémographiques. Les informations concernant le nombre de consultations dans l'année, le revenu ou les caractéristiques de la patientèle restent des données déclaratives, mais il était demandé aux MGs de référer au relevé individuel d'activité professionnelle (RIAP) de 2008, qui est fourni chaque année à tous les médecins par l'assurance maladie.

2.2. Sélection des variables de motivations intrinsèques

Des échelles de mesures de motivations professionnelles ont été définies (Amabile 1993) et certaines ont été utilisées dans un travail similaire à celui que nous souhaitons mettre en oeuvre (Janus 2013) mais, à notre connaissance, aucune échelle n'a été développée pour mesurer les MI des médecins généralistes en particulier. La profession médicale se définit par un ensemble de caractéristiques spécifiques incluant la fourniture d'un large éventail de tâches techniques nécessitant un haut niveau d'expertise et de responsabilité (diagnostic, choix d'un traitement, suivi des patients) ainsi que des compétences non techniques mises en oeuvre durant la relation médecin-patient (écoute du patient, éducation thérapeutique, compassion). De fait, il a été choisi d'élaborer un questionnaire spécifique pour mesurer les MI des médecins généralistes. Les items ont été construits puis sélectionnés à partir d'hypothèses concernant leur capacité à refléter un certain niveau de motivation intrinsèque, ainsi que sur la base de la littérature existante (Janus et al 2008; Videau et al 2010). De façon à harmoniser la construction des variables de MI, celles-ci sont codées 1 lorsque nous supposons que la réponse à la question/item témoigne de motivations intrinsèques élevées (réponse dite « positive »), et 0 sinon. Les statistiques descriptives sur les variables de MI sélectionnées sont présentées dans le tableau 31.

Un premier groupe d'items regroupe trois questions pour lesquelles nous supposons que les médecins répondent positivement lorsqu'ils sont caractérisés par de fortes MI : **la satisfaction de l'activité professionnelle** (question C.1), **le fait de considérer que la profession est gratifiante sur le plan humain** (question C.4) et **les échanges fréquents avec d'autres professionnels de santé** (question A.13).

Les questions C.1 et C.4 sont associées au concept de motivations basées sur le plaisir (« enjoyment-based motivation ») (Deci 1971). Théoriquement, deux facteurs peuvent influencer le sentiment de satisfaction au travail : les motivations intrinsèques et/ou extrinsèques (Goetz et al 2012). Afin de pouvoir correctement interpréter la satisfaction au travail comme reflétant essentiellement les MI (et non les ME), nous avons comparé les

revenus moyens des MGs déclarant être « souvent » satisfaits de leur activité professionnelle à ceux des autres MGs, et n'avons pas obtenu de relation significative ($p=0,1469$). Ce résultat tend à montrer que la satisfaction liée à l'activité n'est pas directement associée au niveau de revenu, ce qui constitue un argument pour l'hypothèse selon laquelle la satisfaction résulte davantage de MI. Les MGs devaient répondre à ces questions sur une échelle de Likert à quatre modalités (« souvent », « de temps en temps », « rarement », « jamais »). Comme la majorité des médecins déclarent être « souvent » satisfaits de leur activité professionnelle (58%) et la considèrent « souvent » gratifiante sur le plan humain (63%), nous avons codé ces variables de MI comme prenant la valeur 1 si le médecin répondait « souvent », et 0 sinon.

La question (A.13) sur la fréquence des échanges avec d'autres professionnels de santé est associée au concept de motivations basées sur l'engagement notamment à l'égard des résultats de la pratique médicale. Les médecins devaient répondre à la question sur une échelle de Likert à quatre modalités (« très souvent », « souvent », « de temps en temps », « rarement »). Afin de construire une variable suffisamment discriminante, la variable de MI est codée 1 lorsque les MGs répondent échanger « très souvent » autour de cas concrets de patients avec d'autres médecins généralistes ou spécialistes⁴⁵ (19%), et 0 sinon.

Un deuxième groupe d'items représente des activités pour lesquelles nous faisons l'hypothèse qu'elles indiquent un investissement du MG pour améliorer la qualité des soins et/ou à partager les connaissances : les **activités d'enseignement ou de maître de stage** (B.6 et B.7) et de **participation à la permanence des soins** (A.12). Ainsi, une première variable est codée 1 si le MG déclare être enseignant ou maître de stage (19% de MGs), 0 sinon ; une seconde variable est codée 1 si le MG déclare participer à la permanence des soins en tant que médecin de permanence (ce qui signifie qu'il peut être sujet à des astreintes la nuit ou le week-end) et de façon régulière (55% des MGs). Théoriquement, ces activités pourraient aussi être considérées comme « motivées extrinsèquement » du fait de rémunérations spécifiques. Toutefois, dans le contexte du paiement à l'acte, la rémunération additionnelle forfaitaire pour ces activités qui prennent du temps n'est pas très élevée. Nous faisons donc l'hypothèse qu'une réponse positive des MG à ces deux questions constitue un signal de MI liée à l'investissement du MG pour améliorer le suivi des patients et le partage des connaissances.

⁴⁵ Nous avons aussi considéré les réponses aux questions portant sur la fréquence des échanges avec les infirmiers et pharmaciens, mais ces items ne se sont pas révélés pertinents pour la mesure des MI, car non corrélés aux autres variables de MI.

En sus, le questionnaire s'intéresse aux activités de prévention fournies par le MG (alcool, tabac, obésité) et à leur fréquence (questions B8_e, B8_f et B8_g) en utilisant une échelle de Likert à quatre modalités (« très souvent », « souvent », « de temps en temps », « rarement »). En France, les activités de prévention ne sont pas rémunérées en tant que telles dans le cadre du paiement à l'acte ; nous faisons donc l'hypothèse, à la suite d'autres travaux (Videau et al 2010), qu'offrir des activités fréquentes reflète une motivation intrinsèque basée sur l'engagement (« obligation-based motivations »). Comme les réponses aux activités de prévention de tabac et d'alcool sont très liées ($p < 0.0001$), ces actions de prévention sont regroupées en une seule variable. Ainsi, une première variable de prévention est codée 1 si les MGs déclarent réaliser « très souvent » des **actions de prévention d'alcool ou du tabac** (15%), 0 sinon ; une seconde variable de prévention est codée 1 si les MGs déclarent réaliser « très souvent » des **actions de prévention contre l'obésité** (20%), 0 sinon.

Enfin, nous considérons deux items supplémentaires pouvant témoigner de MI liée à l'amélioration de la qualité des soins basée sur la confrontation avec les avis des pairs : le fait de participer à **l'évaluation des pratiques professionnelles** (42% des MGs) et **d'être impliqué dans un réseau de soins** (31%).

Le troisième et dernier groupe d'items fait référence aux **attitudes des médecins envers les patients mieux informés** et au **sentiment d'être contraint par le contrôle de l'assurance maladie**. La première variable construite de MI regroupe les réponses à deux questions : (C.6) : « êtes-vous confronté à des patients de plus en plus informés ? » (oui/non) et (C.7) : « si oui, considérez-vous que c'est une contrainte ? » (« souvent », « de temps en temps », « rarement », « jamais »). Parmi les MGs ayant répondu « oui » à la question C.6 (88% des MGs), nous supposons que seuls les MGs répondant « rarement » ou « jamais » à la question C.7 (40% des MGs) sont motivés intrinsèquement, car ces réponses peuvent témoigner de leur disponibilité envers le patient et leur inclination à l'éducation thérapeutique. La deuxième variable est construite à partir des réponses à la question C.11 : « vous sentez-vous contraint par le contrôle de l'assurance maladie ? » (oui/non). Nous faisons l'hypothèse qu'une réponse négative peut indiquer un niveau plus élevé de MI, car le sentiment d'autonomie joue un rôle important dans la construction et le maintien des MI (Deci and Ryan 2000). Ainsi, la variable est codée 1 si les MGs répondent « non » (34%) à cette questions, et 0 sinon.

Tableau 31. Description des variables de motivations intrinsèques

Variabiles (modalités) de motivations intrinsèques	N	%
1) Satisfaction de l'activité professionnelle (souvent)	245	57.9
2) La profession est gratifiante sur le plan humain (souvent)	266	62.9
3) Echanges avec d'autres professionnels de santé médicaux (très souvent)	82	19.4
4) Activités d'enseignement et/ou maître de stage (oui)	83	19.6
5) Participation à la permanence des soins (oui)	231	54.6
6) Prévention alcool-tabac (très souvent)	63	14.9
7) Prévention obésité (très souvent)	86	20.3
8) Participation à l'évaluation des pratiques professionnelles (oui)	176	41.6
9) Implication dans un réseau de soins (oui)	132	31.2
10) Sentiment d'être contraint par les patients mieux informés (rarement-jamais)	161	40.2
11) Sentiment d'être contraint par le contrôle de l'assurance maladie (oui)	139	33.5

3. Construction d'un indicateur composite de la motivation intrinsèque des médecins généralistes⁴⁶

3.1. Introduction

L'objectif de cette section est de présenter la méthodologie de construction d'un indicateur composite (IC) permettant de mesurer la motivation intrinsèque (MI) des médecins généralistes français. Pour ce faire, nous utilisons les variables de MI sélectionnées *a priori* (section 2). Un score ou indicateur composite est généralement construit à partir de l'agrégation de plusieurs variables en utilisant un algorithme adapté. De nombreuses méthodes existent (Nardo et al 2005) et l'objectif de ce travail n'est pas de présenter et de comparer l'ensemble des méthodes disponibles mais de sélectionner celles qui s'appliquent le mieux à notre contexte et surtout à nos données. Nous supposons que la méthode de construction de l'IC requiert les caractéristiques suivantes : premièrement, elle doit permettre la gestion de variables catégorielles utilisées pour décrire le concept de MI. Deuxièmement, elle doit permettre d'investiguer la validité structurelle de l'échelle de mesure des MI, c'est-à-dire, notamment, d'en analyser le nombre de dimensions. Enfin, elle doit permettre de prendre en compte les erreurs de mesures inhérentes aux données déclaratives.

⁴⁶ Ce travail a fait l'objet d'un article en commun avec Carine Franc et Marc Le Vaillant publié dans la revue *Int J Qual Health Care* (2014 Apr ; 26(2) :167-73) sous le titre : « Building a composite score of general practitioners' intrinsic motivation : a comparison of methods ».

Sur la base de ces pré-requis, trois méthodes de construction du score composite sont sélectionnées. La première est l'analyse des correspondances multiples (ACM), déjà utilisée par exemple pour la construction d'un score composite de la pauvreté (Asselin and Vu 2008). La seconde méthode est dérivée de l'analyse factorielle confirmatoire (CFA) adaptée aux variables catégorielles (Christoffersson 1975; Muthén 1978). La troisième méthode consiste en un modèle logistique à deux paramètres (Birnbaum 1968) issu de la théorie de réponse aux items (IRT). Cette dernière méthode est largement utilisée en psychométrie pour mesurer l'intelligence ou la qualité de vie (Sébille et al 2010). Les modèles de réponse à l'item tels que le modèle de Birnbaum sont des extensions des modèles linéaires mixtes généralisés et sont basés sur des techniques d'estimation similaires (Bartholomew and Knott 1999).

Avant de présenter ces différentes techniques statistiques et leurs résultats, nous examinons la validité structurelle des 11 variables de MI sélectionnés *a priori* (tableau 31). Une fois validées, les variables sont appelées « indicateurs » de MI et seront utilisées pour construire les indicateurs composites de motivations intrinsèques (**ICMI**).

3.2. Validation des variables de motivations intrinsèques

3.2.1. L'analyse des correspondances multiples

Nous utilisons la méthode de l'analyse des correspondances multiples (**ACM**) – technique d'analyse factorielle exploratoire adaptée aux données catégorielles (Greenacre 1984) - afin d'examiner la validité structurelle des variables de MI sélectionnées *a priori* et d'en extraire le nombre de dimensions. En effet, en se basant sur une décomposition de l'inertie projetée des données sur des axes factoriels, l'ACM permet dans un premier temps l'identification d'axes ou de dimensions latentes du concept de MI. Cette étape est essentielle afin de déterminer s'il existe une seule dimension de MI et quelles sont les variables qui la caractérisent. Puis, en se basant sur le calcul de la contribution des variables et des modalités des variables à l'axe ou aux axes factoriels mis en évidence, elle permet de sélectionner une majorité de variables corrélées à ces axes qui constituent autant d'indicateurs validés du concept de MI.

3.2.2. *Sélection des indicateurs de motivations intrinsèques*

Les 11 variables de MI (tableau 31) sont définies comme variables actives de l'ACM, c'est-à-dire qu'elles participent toutes à la construction des axes. Puis, l'ACM est réalisée à partir du logiciel SPAD (version 7.4) ; les résultats sont présentés dans le tableau 32. Selon la correction de Benzécri pour les valeurs propres (Benzécri 1979), le premier axe conserve 71% de l'inertie totale des données et les deux premiers axes conservent conjointement 95% de cette inertie. De fait, nous limitons l'interprétation des relations entre les variables de MI et leurs modalités à ces deux premiers axes. Tout d'abord, toutes les variables contribuent de façon significative à l'axe 1 (le plus important en termes de pouvoir explicatif), à l'exception de la variable « participation à la permanence des soins » et « sentiment d'être contraint par l'assurance maladie ». En outre, toutes les modalités codées 1 des 9 autres variables ont la même coordonnée sur l'axe 1, indiquant qu'elles sont positivement corrélées. De fait, ces 9 variables constituent bien des indicateurs de MI et sont sélectionnées pour construire les ICMI.

En ce qui concerne l'axe 2, il apparaît que les variables de prévention (alcool/tabac et obésité) sont moins corrélées avec les indicateurs de MI, en particulier avec « la satisfaction » et « le sentiment de gratification humaine » de l'activité. Si ces résultats ne sont pas contradictoires avec la construction d'un score reposant sur une logique exploratoire, ils peuvent poser problème lorsqu'il s'agit de valider une échelle de mesure. De fait, différentes ACM sont menées à partir des 9 indicateurs de MI présélectionnés et en omettant soit une des variables de prévention soit les deux, de façon à maximiser l'inertie conservée sur le premier axe de l'ACM. La solution permettant d'y parvenir consiste à retirer la variable de prévention contre l'obésité de l'échelle de mesure, ramenant cette échelle à **8 indicateurs** de MI, permettant de conserver plus de 92% de l'inertie des données sur l'axe 1.

Tableau 32. Résultats de l'Analyse des Correspondances Multiples

Variables et modalités de motivation intrinsèque	Axe 1		Axe 2	
	Contr.	Coord.	Contr.	Coord.
1) Satisfaction de l'activité professionnelle				
De temps en temps - rarement - jamais	12.63	-0.85	5.47	-0.50
Souvent	7.25	0.49	3.14	0.28
2) La profession est gratifiante sur le plan humain				
De temps en temps - rarement - jamais	11.60	-0.76	5.97	-0.48
Souvent	8.33	0.54	4.29	0.35
3) Echanges avec d'autres professionnels de santé médicaux				
Souvent - de temps en temps - rarement	1.56	-0.20	0.00	0.00
Très souvent	6.41	0.82	0.00	0.00
4) Activités d'enseignement/maître de stage				
Non	2.22	-0.24	0.00	0.00
Oui	8.85	0.96	0.00	0.01
5) Participation à la permanence des soins				
Non	0.44	-0.14	0.80	0.17
Oui	0.36	0.12	0.66	-0.14
6) Prévention alcool-tabac				
Souvent - de temps en temps - rarement	1.08	-0.16	5.00	0.31
Très souvent	6.00	0.90	27.86	-1.73
7) Prévention obésité				
Souvent - de temps en temps - rarement	1.36	-0.19	6.95	0.38
Très souvent	5.25	0.73	26.90	-1.46
8) Participation à l'évaluation des pratiques professionnelles				
Non	3.72	-0.36	0.00	0.01
Oui	5.13	0.50	0.00	-0.01
9) Implication dans un réseau de soins				
Non	2.72	-0.29	0.19	0.07
Oui	5.94	0.63	0.43	-0.15
10) Sentiment d'être contraint par les patients mieux informés				
Rarement - jamais	4.71	0.51	0.31	0.12
Souvent - de temps en temps	2.75	-0.34	0.56	-0.14
11) Sentiment d'être contraint par le contrôle de l'assurance maladie				
Non	0.77	0.23	6.62	0.57
Oui	0.38	-0.11	3.31	-0.28

Lecture : les modalités de variable en gras indiquent une contribution à l'axe supérieure à la contribution moyenne (CTR_{moy}= 4,5). Les modalités ayant la même coordonnée sur la l'axe sont corrélées positivement.

3.3. Méthodes pour la construction des ICMI

Nous faisons l'hypothèse qu'un médecin répondant positivement à un ou à plusieurs indicateurs de MI parmi les 8 validés est intrinsèquement motivé jusqu'à un certain point que nous souhaitons mesurer. Nous présentons les trois méthodes permettant de construire des indicateurs composites (IC) afin de classer les médecins sur un continuum latent de MI.

3.3.1. L'analyse des correspondances multiples

L'ICMI n°1 est construit à partir des résultats de l'ACM effectuée en intégrant les 8 indicateurs sélectionnés de MI comme variables actives, chaque indicateur ayant deux catégories (0 ou 1). Les contributions des catégories codées 1 (et ainsi associées positivement au concept de MI) sont utilisées comme pondérations pour construire l'IC. Ces poids sont standardisés de façon à ce que leur somme soit égale à 1 permettant ainsi d'obtenir un IC théoriquement compris entre 0 et 1, 0 indiquant le niveau le plus faible de MI et 1, le niveau le plus fort. Ainsi, l'ICMI1 est calculé en utilisant la formule suivante :

$$ICMI_1 = \sum_{i=1}^p \frac{c_i}{T} I_i$$

Où $c_i = \frac{n_i(a_i)^2}{np\mu}$ est la contribution de la catégorie i au premier axe de l'ACM : n_i est l'effectif des MGs appartenant à la catégorie i ; a_i est la coordonnée de la catégorie i sur le premier axe de l'ACM, n est le nombre total de MGs, p est le nombre total de catégories et μ est l'inertie associée au premier axe de l'ACM. T est la somme des contributions au premier axe de l'ACM, des catégories codées 1, et I_i est une indicatrice prenant la valeur 1 si le MG appartient à la catégorie i , et prenant la valeur 0 sinon.

Gestion des données manquantes :

Afin de calculer l'ICMI1, il est nécessaire d'imputer une valeur aux non répondants, égale à 0 ou à 1. Trois indicateurs contenaient un nombre de valeurs manquantes supérieur à 2%⁴⁷ : la satisfaction de l'activité professionnelle, le sentiment de gratification professionnelle de l'activité, et le sentiment d'être rarement contraint par les patients plus informés. Afin d'inférer le niveau moyen de MI des MGs non répondants, nous avons créé une catégorie spécifique de non réponses pour ces trois indicateurs qui a été incluse comme modalité active dans l'ACM. Puis, nous avons observé que les coordonnées de ces trois modalités étaient toutes négatives, indiquant un degré plus faible de MI des MGs non répondants. Ainsi, nous avons imputé la valeur 0 à ces catégories de non réponse. Nous avons aussi imputé par défaut

⁴⁷ Ce seuil de 2% est généralement utilisé (par défaut) dans les ACM concernant la gestion des valeurs manquantes associées à une variable: lorsque ce nombre est inférieur à 2%, les valeurs manquantes sont réparties aléatoirement dans les différentes catégories de la variable ; lorsqu'il est supérieur à 2%, il est possible de construire une catégorie spécifique de non réponse (comprenant un nombre suffisant d'individus) qui sera utilisée comme modalité active de l'ACM.

la valeur 0 aux catégories de non réponse des autres variables (comptabilisant moins de 2% de non réponse), qui comprenaient un nombre trop faible d'individus (moins de 4) pour être incluses dans l'ACM.

3.3.2. L'analyse factorielle confirmatoire pour indicateurs catégoriels

L'analyse factorielle confirmatoire (CFA) est une technique très utilisée en psychométrie et en mesure de qualité de vie ; elle permet notamment de valider des échelles de mesures et/ou de construire des variables latentes (Mulaik 1988; Watt et al 2014). Appliquée à notre contexte unidimensionnel, cette analyse consiste à estimer le système d'équations suivant :

$$\begin{cases} y_1 = a_1 + b_1\theta + e_1 \\ \dots \\ y_p = a_p + b_p\theta + e_p \end{cases}$$

Où y_i dénote la réponse à l'item i ($i = 1, \dots, p$), a_i est une constante, b_i est un facteur de charge (« factor loading ») représentant la proportion de la variance totale de y_i expliquée par la variable latente θ de MI, et e_i est un terme d'erreur supposé suivre une loi normale centrée réduite.

Comme nous ne disposons pas de variables continues, l'hypothèse de normalité des termes d'erreur ne tient pas et il est nécessaire d'adapter ce modèle au cas d'indicateurs binaires (Muthén 1978). Notons y , le vecteur composé de p indicateurs dichotomiques. Nous supposons que la valeur de y_i est déterminée par une variable latente θ_i^* , telle que :

$$y_i = \begin{cases} 1, si \theta_i^* \geq \tau_i \\ 0, si \theta_i^* < \tau_i \end{cases}$$

Où τ_i peut être interprété comme une valeur seuil ('threshold') pour θ_i^* . Nous supposons par ailleurs que :

$$\theta^* = \Lambda \theta + \varepsilon$$

Avec Λ un vecteur de « factor loadings » de dimension p , θ la variable latente de MI, et ε un vecteur de résidus de dimension p non corrélé à θ , supposé normalement distribué et centré en 0, de matrice de variance-covariance Ψ supposée diagonale (Muthén 1978).

Ainsi, le modèle contient trois types de vecteurs/matrices de paramètres à estimer : τ , Λ and Ψ . Le vecteur de valeurs seuils τ est estimé en fonction de la proportion de réponses positives aux p indicateurs sélectionnés (ici, $p=8$), et dénote le degré de sélectivité des indicateurs. Le vecteur Λ et la matrice Ψ sont estimés en utilisant une technique de moindres carrés robustes permettant de prendre en compte les corrélations tétrachoriques, i.e. les corrélations entre indicateurs binaires (Muthén et al 1997). Ces différents paramètres ainsi que l'ICMI2 déduit du modèle sont estimés à partir du logiciel *Mplus* (6.12) spécialisé dans l'inférence de modèles à variable(s) latente(s) avec indicateurs dichotomiques.

Par ailleurs, différents indicateurs permettent de mesurer la qualité d'ajustement des données au modèle à une seule dimension/variable latente. Tout d'abord la statistique du χ^2 permet de tester l'hypothèse nulle selon laquelle la matrice de variance covariance estimée Ψ par le modèle est égale à la matrice de variance-covariance de l'échantillon. Le non rejet de l'hypothèse nulle signifie que le modèle est correctement spécifié. Toutefois, cette statistique augmente avec la taille de l'échantillon et la déviation par rapport à l'hypothèse de normalité des termes d'erreur ε ; de fait, cette statistique est souvent complétée en pratique par d'autres indicateurs moins exigeants (Baumgartner and Homburg 1996; Iacobucci 2010). Ainsi, l'erreur moyenne d'approximation ou « root mean square of approximation » (RMSEA) constitue un second indicateur de qualité d'ajustement, qui doit être inférieur à 0.08 pour que le modèle soit considéré comme correctement spécifié (Browne and Cudeck 1993). Enfin, l'indice de Tucker Lewis (TLI) permet d'évaluer la qualité d'ajustement du modèle à un facteur par comparaison au modèle 'nul' c'est-à-dire où tous les indicateurs de MI sont supposés indépendants. Cet indicateur est compris entre 0 et 1 : par exemple une valeur de 0.9 indique que le modèle spécifié augmente la qualité d'ajustement de 90% par rapport au modèle nul. Il est conseillé de retenir un modèle dont l'indice TLI est supérieur à 0.9 (Browne and Cudeck 1993).

3.3.3. *Le modèle logistique à deux paramètres (modèle de Birnbaum)*

Le modèle logistique à deux paramètres, aussi appelé **modèle de Birnbaum**, s'écrit de la façon suivante:

$$\text{logit } p_{ij} = \alpha_i \theta_j - \delta_i + e_{ij}$$

Où p_{ij} dénote la probabilité que le médecin j réponde positivement à l'item i , e_{ij} et θ_j représentent deux termes d'erreurs aléatoires : e_{ij} est un terme d'erreur qui varie d'un item à l'autre et d'un médecin à l'autre (variabilité intra-médecin) et θ_j est un terme d'erreur spécifique au médecin j (variabilité inter-médecin). Ainsi, θ_j (ICMI3) est la réalisation d'une variable latente θ de MI, supposée normalement distribuée et centrée en 0. Plus précisément, afin d'obtenir un IC compris entre 0 et 1, l'ICMI3 pour le médecin j est calculé de la façon suivante :

$$\widehat{\text{ICMI3}}_j = \frac{\exp(\hat{\theta}_j)}{1 + \exp(\hat{\theta}_j)}$$

Les paramètres δ_i et α_i représentent respectivement les paramètres de difficulté et de discrimination de l'item i . Selon la théorie de réponse aux items, la difficulté d'un item est associée à la proportion de réponses négatives à cet item : plus ce taux est élevé, plus l'item est dit difficile. Le paramètre de discrimination indique quant à lui dans quelle mesure les réponses à l'item i permettent de discriminer entre des individus ayant un niveau élevé ou faible de MI : plus il est élevé, plus l'item est pertinent pour mesurer les MI. Les paramètres $\hat{\delta}_i$, $\hat{\alpha}_i$ et $\hat{\theta}_j$ sont estimés par la méthode du maximum de vraisemblance marginale (Bock and Aitkin 1981) en utilisant la procédure Nlmixed de SAS (version 9.3).

3.3.4. Méthodes pour la comparaison des trois indicateurs composites

Les résultats des méthodes statistiques de construction des ICMI sont comparés de trois façons différentes.

Représentation graphique des scores :

Dans un premier temps, nous représentons sur un graphique les valeurs des trois ICMI calculés (pour les 30 premiers MGs) de manière à inspecter des différences inter-MGs et inter-méthodes.

Coefficients de Kappa pondérés :

Dans un second temps, afin d'évaluer la concordance dans le classement des MGs selon la méthode utilisée, nous classons les médecins dans 4 groupes de score basés sur les quartiles de la distribution des ICMI. Puis, nous calculons des coefficients de Kappa pondérés

(K_w) permettant de donner une mesure de la discordance entre deux méthodes de classement (Cohen 1968). Ces coefficients sont calculés de la façon suivante :

$$K_w = \frac{P_{0(w)} - P_{e(w)}}{1 - P_{e(w)}}$$

Avec $P_{0(w)} = \sum_{i=1}^4 \sum_{j=1}^4 w_{ij} p_{ij}$, où w_{ij} est un coefficient de pondération, p_{ij} représente la proportion d'individus classés respectivement dans les groupes i et j par les deux méthodes de scoring : $P_{0(w)}$ s'interprète donc comme la proportion d'accords observés entre les deux méthodes, qui est maximale si les groupes i et j de classement sont identiques pour chaque individu.

Par ailleurs, $P_{e(w)} = \sum_{i=1}^4 \sum_{j=1}^4 w_{ij} p_{i.} p_{.j}$, où $p_{i.}$ est la proportion marginale colonne et $p_{.j}$, la proportion marginale ligne du tableau de contingence : $P_{e(w)}$ s'interprète donc comme la proportion d'accords attendue sous l'hypothèse d'indépendance des deux méthodes de scoring.

Différentes techniques de pondération existent. Nous utilisons la technique de Fleiss-Cohen (Cohen 1968), aussi appelée pondération quadratique. Cette pondération permet de prendre en compte le degré de concordance entre deux méthodes : plus les catégories de classement sont éloignées (i est éloigné de j), plus la pondération est faible. De façon extrême, si une méthode classe un médecin dans le 1^{er} groupe et une autre méthode le classe dans le 4^e groupe, alors la pondération sera nulle. Les poids w_{ij} sont calculés de la façon suivante :

$$w_{ij} = 1 - \left(\frac{i-j}{4-1} \right)^2$$

Comparaison des effectifs des MGs les moins/plus motivés :

De façon complémentaire au calcul des coefficients Kappa, nous comparons les proportions de MGs qui, pour chaque méthode, sont tous classés dans le premier groupe (le quartile des médecins les moins motivés intrinsèquement) et tous classés dans le quatrième groupe (le quartile des médecins les plus motivés intrinsèquement).

3.4. Estimation et comparaison des trois ICMI

3.4.1. Pondération des indicateurs

Le tableau 33 présente une synthèse des 8 indicateurs de MI sélectionnés ainsi que leur poids respectif standardisé calculé à partir de leur contribution sur le 1^{er} axe de l'ACM. Les indicateurs les plus contributeurs sont l'activité d'enseignement/maître de stage ($w=0,19$), la satisfaction de l'activité professionnelle ($w=0,19$) et le sentiment que l'activité est gratifiante sur le plan humain ($w=0,16$).

Tableau 33. Synthèse des indicateurs de MI et estimation des poids basés sur l'ACM

Indicateurs de motivations intrinsèques	Modalité associée à la MI	Poids standardisé sur l'ACM
1) Satisfaction de l'activité professionnelle	Souvent	0.19
2) La profession est gratifiante sur le plan humain	Souvent	0.16
3) Echanges avec d'autres professionnels de santé médicaux	Très souvent	0.13
4) Activités d'enseignement et/ou maître de stage	Oui	0.19
5) Prévention alcool-tabac	Très souvent	0.04
6) Participation à l'évaluation des pratiques professionnelles	Oui	0.10
7) Implication dans un réseau de soins	Oui	0.11
8) Sentiment d'être contraint par les patients mieux informés	Rarement jamais	0.08

Les résultats des paramètres associés au calcul de l'ICMI2 (paramètres de seuil, factor loadings et écarts-types des coefficients) sont présentés dans le tableau 34. Les indicateurs les plus corrélés à la variable de MI mesurée (factor loadings élevés) sont la satisfaction de l'activité professionnelle et le sentiment que l'activité est gratifiante sur le plan humain ; ils sont suivis par les activités d'enseignant/maître de stage et les échanges avec les autres professionnels de santé. Les indicateurs de qualité d'ajustement du modèle révèlent une adéquation acceptable du modèle à nos données : l'indice de Tucker-Lewis (TFI) est égal à 0,92 et l'erreur moyenne d'approximation (RMSEA) est égale à 0,074. En revanche, la statistique du χ^2 est statistiquement significative ($\chi^2 = 66,4$; $p=0,001$) témoignant du fait que ce modèle à un facteur ne permet pas d'expliquer toute la variabilité des données.

Les résultats des paramètres associés au calcul de l'ICMI3 (paramètres de difficulté, de discrimination et écarts-types) sont présentés dans le tableau 34. Les paramètres de difficulté sont comparables aux paramètres de seuil estimés par la CFA : les 8 indicateurs de MI occupent la même position sur l'échelle de difficulté. Par ailleurs, les indicateurs les plus discriminants sont les mêmes que ceux obtenus par la CFA : la satisfaction de l'activité

professionnelle et le sentiment que l'activité est gratifiante sur le plan humain. Même si certains indicateurs semblent peu discriminants, ils sont tout de même significatifs au seuil de 5%.

Tableau 34. Résultats des estimations des modèles CFA et Birnbaum

Indicateurs de motivations intrinsèques	CFA		Modèle de Birnbaum	
	Threshold (SE)	Factor loadings (SE)	Difficulté (SE)	Discrimination (SE)
1) Satisfaction de l'activité professionnelle	-0.24 (0.06)	2.35 (0.53)	-0.19 (0.07)	4.45 (1.50)
2) La profession est gratifiante sur le plan humain	-0.35 (0.06)	2.34 (0.53)	-0.27 (0.09)	4.51 (1.42)
3) Echanges avec d'autres professionnels de santé médicaux	0.86 (0.07)	1.00 (.)	1.56 (0.15)	1.00 (.)
4) Activités d'enseignement et/ou maître de stage	0.86 (0.07)	1.15 (0.31)	1.40 (0.46)	1.13 (0.40)
5) Prévention alcool-tabac	1.04 (0.08)	0.51 (0.26)	3.58 (2.00)	0.50 (0.29)
6) Participation à l'évaluation des pratiques professionnelles	0.21 (0.06)	0.90 (0.26)	0.47 (0.21)	0.76 (0.28)
7) Implication dans un réseau de soins	0.49 (0.06)	0.82 (0.26)	1.20 (0.47)	0.69 (0.27)
8) Sentiment d'être contraint par les patients mieux informés	0.16 (0.07)	0.84 (0.27)	0.36 (0.20)	0.77 (0.30)

* SE : standard errors.

3.4.2. Distribution des ICMI

3.4.2.1. Statistiques descriptives

Les statistiques descriptives concernant la distribution des trois ICMI sont présentées dans le tableau 35. Les ICMI1 et ICMI3 sont théoriquement distribués entre 0 et 1. L'ICMI1 est compris entre 0 pour les MGs les moins motivés intrinsèquement et 0,99 pour les MGs les plus motivés intrinsèquement. La moitié des MGs ont un score supérieur à 0,38. L'ICMI3 a une variabilité plus faible ($sd=0,11$ vs $sd=0,24$ pour l'IMCS1) : il est compris entre 0,32 et 0,77 et la médiane est égale à 0,49. L'ICMI2 n'est pas compris entre 0 et 1, et peut contenir des valeurs négatives, rendant la comparaison moins aisée avec les autres scores. Les trois indicateurs composites ont une distribution légèrement plate (coefficients de kurtosis négatifs). Par ailleurs, la distribution de l'ICMI1 est légèrement décalée sur la droite ($skewness = 0,23$) indiquant des scores de MI relativement plus élevés par rapport aux autres ICMI.

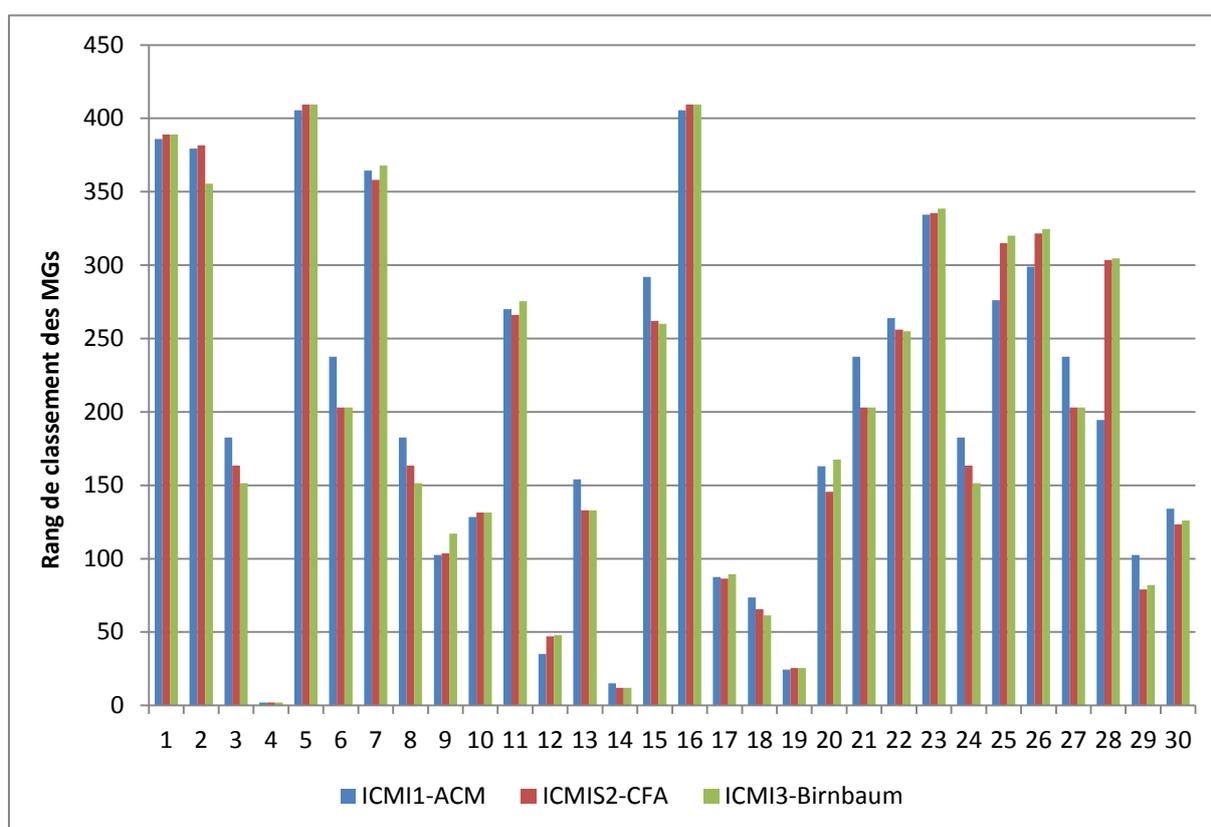
Tableau 35. Statistiques descriptives sur la distribution des ICMI

	<i>Min-Max</i>	<i>Mean</i>	<i>Median (Q1;Q4)</i>	<i>SD</i>	<i>Skewness</i>	<i>Kurtosis</i>
ICMI1-ACM	0 - 0.99	0.38	0.38 (0.18;0.56)	0.24	0.216	-0.649
ICMI2-CFA	-0.43 - 0.64	-0.01	0.06 (-0.27;0.18)	0.26	0.002	-0.925
ICMI3-Birnbaum	0.31 - 0.77	0.49	0.53 (0.38;0.58)	0.11	0.016	-1.020

3.4.2.2. Comparaison du classement des MGs

Le classement des MGs selon chaque méthode est présenté sur la figure 6. Notons que le classement semble similaire d'une méthode à l'autre. En particulier, nous observons une concordance forte entre la CFA et le modèle de Birnbaum.

Figure 6. Histogramme du rang de classement de 30 MGs en fonction de l'ICMI calculé



Les résultats du calcul des coefficients de Kappa pondérés sont présentés dans le tableau 36. Tous les coefficients sont supérieurs à 0,80, indiquant un niveau de concordance très bon voire excellent entre les méthodes. Les coefficients les plus élevés sont obtenus pour les ICMI2 et ICMI3 ($K_w=0,96$), confirmant que la CFA et le modèle de Birnbaum produisent les résultats les plus cohérents. Par ailleurs, parmi les MGs classés dans le premier groupe (les médecins ayant le niveau de MI le plus faible) 84% des MGs le sont par les trois méthodes. A

l'opposé, parmi les MGs classés dans le 4^e groupe (les médecins ayant le niveau de MI le plus élevé), 86% le sont les trois méthodes.

Tableau 36. Coefficients de Kappa pondérés entre les trois ICMI

	ICMI1-ACM	ICMI2_CFA	ICMI3_Birnbaum
ICMI1-ACM	1	0.86	0.89
ICMI2_CFA		1	0.96
ICMI3_Birnbaum			1

3.5. Discussion

L'objectif de cette étude était d'identifier et de valider différents indicateurs de la motivation intrinsèque (ICMI) des médecins généralistes français et de sélectionner la méthode la plus adaptée à nos données. Parmi onze variables de MI sélectionnées *a priori* sur la base d'hypothèses et d'une revue de la littérature, nous avons identifié huit indicateurs corrélés entre eux et représentant le concept de MI à travers une unique dimension. Par ailleurs, les trois méthodes utilisées pour la construction des ICMI ont produits des résultats similaires concernant à la fois la pondération des indicateurs et le classement des MGs sur l'échelle de MI. En effet, les indicateurs les plus importants étaient la satisfaction de l'activité professionnelle et le sentiment d'exercer une profession gratifiante sur le plan humain. La comparaison des estimations obtenues (coefficients de Kappa pondérés) a indiqué une forte corrélation entre les rangs de classement des MGs pour chaque méthode, et particulièrement entre la CFA et le modèle de Birnbaum, témoignant de la robustesse du score final obtenu. Ces différents résultats suggèrent que les ICMI construits représentent une mesure valide du concept de MI en médecine générale.

Théoriquement, l'ACM et la CFA utilisées dans ce travail sont des méthodes d'analyse factorielle au sens large, mais elles diffèrent dans leur traitement des motifs de corrélations entre variables non continues. Dans l'ACM, une variable ou modalité de variable active contribue à la construction d'un axe factoriel au travers de son inertie par rapport au centre de gravité des données ; les axes factoriels sont construits comme des combinaisons linéaires de modalités de variables associées entre elles. La CFA se base elle sur une décomposition de la matrice de variance-covariance des données, permettant de dégager des axes factoriels au moyen de techniques inférentielles de type maximum de vraisemblance (Jöreskog 1969). Lorsque les indicateurs utilisés pour mesurer le concept latent sont binaires,

la matrice de variance-covariance de Pearson est « remplacée » par la matrice de corrélation tétrachorique, i.e. la matrice des corrélations entre les variables latentes déterminants les réponses à chaque indicateur binaire. Ceci permet d'obtenir des résultats plus robustes par rapport à la CFA « traditionnelle » (Muthén et al 1997).

La CFA mise en œuvre dans cette étude et le modèle de Birnbaum diffèrent dans leur façon de modéliser la relation entre un indicateur binaire et une variable latente continue. La CFA utilise un modèle linéaire de la relation entre l'indicateur observé et la variable latente, et les coefficients de charge (factor loadings) ainsi que l'ICMI sont calculés par des techniques de moindres carrés robustes (Muthén et al 1997). Dans le modèle de Birbaum, les indicateurs observés sont reliés à la variable latente par une fonction de lien logit, et les paramètres du modèle ainsi que l'ICMI sont estimés par maximum de vraisemblance marginale.

Comparée à la CFA et au modèle de Birnbaum, la faiblesse de l'ACM est qu'elle ne repose pas sur un modèle spécifié et ne tient pas compte des erreurs de mesures inhérentes aux données déclaratives. De plus, l'ACM est moins efficace concernant la gestion des valeurs manquantes, car il est nécessaire d'imputer ces valeurs, contrairement à la CFA et au modèle de Birnbaum (et, plus généralement, de tout modèle basé sur l'IRT).

Différentes considérations nous conduisent à penser que le modèle de Birnbaum est probablement la méthode la plus adaptée pour construire l'ICMI. Tout d'abord, ce modèle est basé sur des techniques d'estimation qui permettent une gestion efficace des données manquantes (Sébille et al 2010; Brodin et al 2010). Deuxièmement ce type de modèle a été de plus en plus utilisé pour mesurer des variables latentes basées sur des indicateurs catégoriels (Hays et al 2000). Enfin, empiriquement, il s'agit de la méthode la plus conciliante au sens des coefficients de Kappa.

L'ACM initiale exécutée afin d'identifier les indicateurs de MI présélectionnés fournit un argument fort en faveur de l'estimation d'une unique variable latente, celle-ci préservant plus de 90% de l'inertie des données. Ces résultats sont confirmés par les indices de qualité d'ajustement calculés grâce à la CFA, montrant un modèle à un facteur acceptable (CFI = 0,92 ; RMSEA = 0,074). Cependant, la valeur significative du χ^2 met en garde concernant la capacité du modèle à rendre compte de l'ensemble des structures de corrélation du modèle. Ce résultat peut révéler que certains des indicateurs construits manquent de précision. De plus, une limite des méthodes est que nous ne pouvons pas déterminer précisément quelle

proportion du concept de motivation intrinsèque est captée par nos indicateurs ; nos données tout comme la littérature ne nous permettent pas de répondre à cette question.

Sur la base de ces résultats, d'autres études devraient examiner la notion de motivation intrinsèque par la construction d'indicateurs plus précis. Par exemple, en ce qui concerne la fourniture d'actes de prévention, il serait intéressant de recueillir des informations sur les actions précises effectuées pour améliorer l'état de santé des patients et/ou sur le temps consacré à l'exécution de ces activités. En effet, en tant que tels, ces indicateurs n'ont pas semblé les plus pertinents pour mesurer la MI des médecins généralistes.

4. Relation entre motivation intrinsèque et extrinsèque en médecine générale. Une étude exploratoire⁴⁸

4.1. Objectifs

Une fois la méthode de construction du score de MI établie, il s'agit alors d'analyser les relations de complémentarité ou de substituabilité entre les MI et les ME des médecins généralistes français. En effet, si la complémentarité des MI et ME semble constituer un levier d'amélioration de l'efficacité des politiques d'incitation, la mise en évidence d'une relation inverse, soit une relation de substitution, poserait de réelles questions. En réalité, la mise en évidence d'une relation de substitution conforterait l'hypothèse d'effet d'éviction des MI par les ME ce qui pourrait constituer une remise en cause de l'efficacité des politiques de promotion d'incitatifs extrinsèques tels que, par exemple, les P4P, et un questionnement sur les effets à long terme de telles politiques. Pour étudier la nature des liens entre MI et ME, nous mobilisons les indicateurs de MI validés précédemment (section 3), et développons par ailleurs des indicateurs de motivations extrinsèques (ME).

A partir de la méthode de construction du score de MI des MGs validé, nous étudions l'impact des caractéristiques des médecins sur sa variabilité. Il s'agit de construire un modèle qui permette certes l'estimation du score de MI mais aussi et surtout d'étudier sa variabilité entre médecins. Dans un second temps, nous testons la relation entre l'ICMI obtenu et plusieurs indicateurs identifiés de ME.

4.2. Données et méthodes

4.2.1. Caractéristiques des médecins

Les variables médecins sélectionnées pour l'analyse de la variabilité du score de MI sont des variables individuelles : l'âge, le sexe, l'état de santé perçu (question D.13), le fait de travailler en groupe (question A.3) et le fait d'avoir une activité salariée en dehors de l'activité libérale (question A.1). La variable d'état de santé est codée 1 lorsque le médecin

⁴⁸ Ce travail a fait l'objet d'un article en commun avec Carine Franc et Marc le Vaillant publié dans la revue *Health Policy* (2012 Dec ; 108(2-3) :140-8) sous le titre : « Intrinsic and extrinsic motivations in primary care : an explanatory study among French general practitioners ».

reporte un « excellent » ou « très bon » état de santé, et 0 sinon. La variable « exercice en groupe » est indépendante de la spécialité des autres médecins (soit des médecins généralistes soit des médecins de différentes spécialités). Plus précisément, cette variable est codée 1 si le médecin déclare partager les dossiers des patients (question A.4), et 0 sinon. Enfin, la variable d'activité salariée est codée 1 si le médecin déclare être salarié hospitalier ou effectuer des vacations, et 0 sinon.

Nous considérons aussi des caractéristiques relatives au contexte de l'activité : la zone rurale ou urbaine déterminée par le nombre d'habitants ($\leq 10\,000$ / $> 10\,000$ habitants) ainsi qu'une variable indiquant si les MGs rencontrent des difficultés particulières liées au profil de leur patientèle ('oui'/'non'). Toutes ces variables sont susceptibles d'avoir un impact sur le niveau de MI. Le tableau 37 présente des statistiques descriptives associées à ces variables, et récapitule les variables de MI sélectionnées dans la section précédente.

L'âge moyen dans l'échantillon est de 52 ans et il y a une majorité d'hommes (72%). Par ailleurs, 39% des médecins déclarent un excellent ou très bon état de santé, 50% d'entre eux sont en groupe et partagent les dossiers, et 21% d'entre eux ont une activité salariée en dehors de leur activité libérale. Enfin, 60% des médecins vivent en zone urbaine et 24% d'entre eux déclarent rencontrer certaines difficultés dans leur pratique liée à leur profil de patientèle.

4.2.2. *Variables de motivations extrinsèques*

Contrairement aux MI, les ME se réfèrent à des motivations 'externes' à la personne, lorsque l'activité est réalisée pour obtenir une récompense ou une gratification (Frey 1997a; Bénabou and Tirole 2003). Cinq variables sont sélectionnées comme proxies des ME. Trois variables concernent les motivations monétaires : **(a) le chiffre d'affaire en 2008** extrait du RIAP, **(b) le chiffre d'affaire divisé par le nombre d'actes réalisés en 2008** et **(c) la satisfaction de la rémunération** ('oui'/'non'). Nous supposons que l'indicateur (b) peut être plus informatif que l'indicateur (a) car il peut constituer un proxy du rendement moyen de la consultation. En ce qui concerne la variable (c), nous faisons l'hypothèse que les MGs déclarant ne pas être satisfaits de leur revenu pourraient avoir des motivations extrinsèques plus grandes. Ainsi, la variable (c) de ME est donc codée 1 si le MG répond « oui » à la question, et 0 sinon.

Les autres variables de ME sélectionnées illustrent le sentiment d'être contraint par les exigences des patients concernant : **(d) les horaires d'ouverture du cabinet** (question C.3A)

ou (e) la **durée de la consultation** (question C.3B). En effet, ces deux variables pourraient refléter un arbitrage des médecins entre travail et loisir. Nous faisons l'hypothèse que les MGs déclarant être « souvent » ou « de temps en temps » contraints par les exigences des patients pourraient ne pas vouloir perdre de temps, autrement dit, de l'argent ou du temps de loisir, et donc être davantage motivés extrinsèquement. Les variables (d) et (e) sont donc codées 1 si le MG répond « souvent » ou « de temps en temps » à la question posée, et 0 sinon. Les statistiques descriptives des cinq indicateurs de ME ainsi construits sont présentés dans le tableau 37.

Tableau 37. Statistiques descriptives (MI, ME et variables médecin)

Variables	%
1-Motivations intrinsèques	
Satisfaction de l'activité professionnelle (souvent)	57.9
La profession est gratifiante sur le plan humain (souvent)	62.9
Echanges avec d'autres professionnels de santé médicaux (très souvent)	19.4
Activités d'enseignement et/ou maître de stage (oui)	19.6
Prévention alcool-tabac (très souvent)	14.9
Prévention obésité (très souvent)	20.3
Participation à l'évaluation des pratiques professionnelles (oui)	41.6
Implication dans un réseau de soins (oui)	31.2
Sentiment d'être contraint par les patients mieux informés (rarement-jamais)	40.2
2-Motivations extrinsèques	
Chiffre d'affaire (moyenne \pm S.D.)	125139 \pm 55451
Chiffre d'affaire/nb actes (moyenne \pm S.D.)	26.6 \pm 5.1
Satisfaction du revenu (non)	63.1
Sentiment d'être contraint par les exigences des patients concernant ...	
... les horaires d'ouverture du cabinet (de temps en temps-souvent)	54.8
... la durée de la consultation (de temps en temps-souvent)	55.3
3-Caractéristiques des médecins	
Age (moyenne \pm S.D.)	52.1 \pm 7.6
Sexe (homme)	71.8
Etat de santé (excellent-très bon)	39.4
Groupe avec partage des dossiers	49.7
Activité salariée	20.6
Zone urbaine	59.6
Difficultés liées au profil de patientèle (oui)	24.1

Le chiffre d'affaire moyen dans l'échantillon est de 125 139€, et le chiffre d'affaire moyen par acte est de 27€. Par ailleurs, 63% des MGs déclarent ne pas être satisfaits par le niveau de leur rémunération, et environ 55% d'entre eux déclarent être « de temps en temps »

ou « souvent » contraints par les exigences des patients concernant les horaires d'ouverture du cabinet ou la durée de la consultation.

4.2.1. Construction et analyse de la variabilité du score de MI

Le modèle sélectionné pour construire l'ICMI et analyser sa variabilité est un modèle logistique à constante aléatoire (Rice et Jones, 1997), du même type que le modèle de Birnbaum sélectionné dans la section précédente. L'avantage de ce type de modèle est qu'il permet de distinguer les composantes de la variabilité du score de MI liées à la variabilité intra- ou inter-médecin (Rice and Jones 1997; Snijders and Bosker 1999). Ainsi, le modèle multi-niveau permet une approche structurelle du problème : au sein d'un même modèle, nous mesurons la variable latente de MI et étudions sa variabilité par l'introduction de variables caractérisant le médecin. Plus précisément, ce modèle constitue une simplification du modèle logistique à deux paramètres (modèle de Birnbaum), permettant l'introduction de covariables médecin.

Nous supposons qu'il existe une variable latente de MI, notée θ , avec comme réalisation θ_j , qui détermine les réponses du médecin j aux différents indicateurs de MI sélectionnés. De plus, nous supposons que ces réponses sont indépendantes pour chaque médecin, conditionnellement à θ_j . La variable latente θ est supposée normalement distribuée, de moyenne nulle et de variance τ_π .

Le modèle multi-niveau à constante aléatoire s'écrit de la façon suivante :

Modèle 1 : Le modèle de mesure

$$\text{logit } p_{ij} = \pi_{0j} + e_{ij} \quad (\text{niveau 1})$$

$$\pi_{0j} = \beta_{00} + \theta_j, \quad \theta_j \rightarrow N(0, \tau_\pi) \quad (\text{niveau 2})$$

où p_{ij} représente la probabilité que le médecin j réponde positivement à l'indicateur i . Nous rappelons qu'une réponse positive à un indicateur correspond à une réponse témoignant de MI. De plus, π_{0j} est un coefficient aléatoire variant d'un médecin à l'autre, représentant la proportion de réponses positives pour le médecin j ; β_{00} est la constante générale du modèle représentant la proportion de réponses positives l'échantillon. Enfin, e_{ij} et θ_j sont deux

termes d'erreurs aléatoires : e_{ij} est un terme d'erreur qui varie d'un indicateur à l'autre et d'un médecin à l'autre (variabilité intra-médecin) et θ_j est un terme d'erreur spécifique au médecin j (variabilité inter-médecin).

Afin d'obtenir un IC compris entre 0 et 1, l'ICMI pour le médecin j est calculé de la façon suivante :

$$\widehat{\text{ICMI}}_j = \frac{\exp(\widehat{\theta}_j)}{1 + \exp(\widehat{\theta}_j)}$$

Modèle 2: Le modèle complet avec covariables

Avec d'expliquer la variabilité de l'ICMI représenté, pour le médecin j , par θ_j , les variables médecin sont introduites au 2^e niveau. Le modèle s'écrit de la façon suivante :

$$\text{logit } p_{ij} = \pi_{0j} + e_{ij} \quad (\text{niveau 1})$$

$$\pi_{0j} = \beta_{00} + \sum_{k=1}^K \beta_k * x_k + \theta_j, \quad \theta_j \rightarrow N(0, \tau_\pi) \quad (\text{niveau 2})$$

Où x_k représente la k -ième variable médecin et β_k est le coefficient associé.

Les modèles 1 et 2 sont estimés par la procédure Nlmixed de SAS (9.3).

4.3. Résultats

4.3.1. Distribution de l'indicateur composite de MI

Basé sur les résultats de l'ACM de la section 3, les neufs indicateurs de MI corrélés sur un unique axe latent sont sélectionnés et donc inclus dans le modèle 1 pour construire l'ICMI (cf. tableau 37). La distribution ainsi que la distribution cumulée de l'ICMI obtenue sont représentées dans les figures C.2.1 et C.2.2 (Annexe C.2). L'ICMI est théoriquement distribué entre 0 et 1, 0 indiquant le niveau le plus faible de motivation intrinsèque, et 1 le niveau le plus élevé. Par construction, l'ICMI peut s'interpréter comme la probabilité de répondre positivement à un indicateur. La moitié des MGs de l'échantillon ont un score supérieur à 0,5 et le score estimé est compris entre un minimum de 0,31 (moins de 1 réponse positive sur 9) et 0,81 (plus 8 réponses positives sur 9), avec un écart type de 0,12. Ainsi, les

MGs figurant parmi les 10% les moins motivés intrinsèquement ont un score inférieur ou égal à 0,37 et les MGs figurant parmi les 10% les plus motivés intrinsèquement ont un score supérieur à 0,68. Enfin, pour 50% des MGs, l'ICMI est compris entre 0,44 et 0,57.

4.3.2. Les déterminants de la variabilité du score

Dans un premier temps, des analyses univariées sont menées en incluant une à une les caractéristiques des médecins dans le modèle multiniveaux (modèle 2). Puis, les variables significatives au seuil de 5% sont incluses dans un modèle multivarié final. Les résultats des différents modèles sont présentés dans le tableau 38. Dans les modèles univariés, trois variables ont un impact positif et significatif sur la variabilité inter-médecin : le fait d'être en groupe et de partager les dossiers ($p < 0,0001$), d'avoir une activité salariée en dehors de l'activité libérale ($p < 0,0001$) et de reporter un très bon état de santé ($p = 0,0248$). Ces 3 caractéristiques médecin sont associés à des niveaux plus élevés de motivations intrinsèques. En outre, ces trois variables restent significatives dans le modèle multivarié. Plus précisément, elles permettent d'expliquer près du quart de la variabilité inter-médecin (24,5%), celle-ci passant de 0,5122 dans le modèle 1 à 0,4051 dans le modèle 2. Enfin, il est intéressant de noter que l'âge, le sexe ou milieu d'exercice (urbain/rural) ne semblent pas avoir d'impact significatif sur le niveau de motivation intrinsèque.

Tableau 38. Résultats des modèles multiniveaux : impact des caractéristiques médecins sur la variabilité de l'ICMI⁴⁹

Effets	Variable	Analyses univariées		Analyses multivariées	
		Coeff.	P-value	Coeff.	P-value
Fixes	Constante			-1.2501	<.0001
	Age <=45 ans	-0.08828	0.5044		
	Age >=60 ans (ref, 45<Age<60)	-0.1018	0.4571		
	Sexe (ref, femme)	0.055	0.6216		
	Pratique en groupe (ref, solo)	0.4433	<.0001	0.3389	<.0001
	Urbain (ref, rural)	0.1265	0.2515		
	Très bon état de santé (ref, moyen-mauvais)	0.4197	0.0248	0.3701	0.0388
	Salarié (ref, libéral uniquement)	0.6145	<.0001	0.5807	<.0001
	Difficultés avec le profil de la patientèle (ref, non)	0.0384	0.7439		
Aléatoires				Variance	
MG				0.4051	

⁴⁹ Certaines variables comportaient des non réponses. De fait, nous avons créé des catégories spécifiques de non réponses qui ont été incluses dans les modèles, mais leurs résultats ne sont pas reportés dans le tableau 38.

4.3.3. Interaction avec les variables de ME

Comme le score de MI obtenu n'est pas parfaitement normal (cf figure C.2.2), nous utilisons les coefficients de corrélation de Spearman (tests du rang) afin d'étudier la relation entre l'ICMI et les variables quantitatives de ME. Comme 19% des MGs n'ont pas reporté le revenu, l'analyse est limitée aux 342 MGs répondants⁵⁰. Les résultats sont présentés dans le tableau 9. Nous n'obtenons pas de relation statistique significative entre le l'ICMI et le revenu ($p=0,8622$) et le revenu moyen par acte ($p=0,6586$).

Afin d'examiner la relation entre l'ICMI et les variables qualitatives de ME, nous utilisons la statistique non paramétrique de Wilcoxon, qui consiste à comparer les rangs de classement des médecins dans chaque catégorie des variables binaires utilisées (tableau 39). Les MGs déclarant être « rarement » ou « jamais » contraints par les exigences des patients concernant les horaires d'ouverture du cabinet (la durée de consultation, respectivement) ont en moyenne un rang de classement plus élevé sur l'échelle de MI ($p=0,0181$ et $p=0,0496$, respectivement). De plus, les MGs déclarant être satisfaits de leur revenu ont en moyenne un niveau de MI et donc un classement plus élevé ($p=0,0113$). Ce résultat pourrait aussi indiquer que les MGs ayant des motivations intrinsèques plus fortes aient moins d'attentes vis-à-vis de leur revenu, ou que leur revenu cible est déjà atteint.

Tableau 39. Coefficients de corrélation de Spearman entre l'ICMI et les variables de ME et tests des rangs de Wilcoxon

Variables quantitatives de ME (N=342)	Coefficient de corrélation de Spearman		p-value
Chiffre d'affaire		-0.00942	0.8622
Chiffre d'affaire/nb d'actes		0.02487	0.6586
Variables qualitatives de ME (N=423)	Rangs moyens *	Statistique de Wilcoxon	p-value
Satisfaction du revenu		31104	0.0113
Oui	222		
Non	194		
Contraints par les exigences des patients concernant :			
... les horaires d'ouverture du cabinet		41315	0.0181
rarement - jamais	220		
de temps en temps - souvent	196		
... la durée de la consultation		40770	0.0496
rarement - jamais	219		
de temps en temps - souvent	199		

* le rang moyen est calculé comme la différence entre l'ICMI du groupe considéré et l'ICMI moyen de l'échantillon.

⁵⁰ Ces non répondants ne sont pas significativement différents des autres MGs en termes d'âge, localisation, ou pratique de groupe.

4.4. Discussion

Du fait du renforcement des schémas incitatifs appliqués aux soins primaires, des questions se posent concernant leur effet potentiel indésirable sur les motivations des professionnels de santé. Cette problématique a fait l'objet de nombreux travaux théoriques ayant étudié les relations entre motivations intrinsèques et extrinsèques dans différents contextes (Frey 1994; Bénabou and Tirole 2003; James Jr. 2005; Sliwka 2006; Schnedler and Vanberg 2014). En revanche, peu d'études ont investigué empiriquement l'interrelation entre MI et ME et a fortiori dans le domaine des soins de premiers recours. Ainsi, l'objectif de notre étude était dans un premier temps de sélectionner un modèle simple permettant de construire un indicateur composite de la motivation intrinsèque (ICMI) des MGs et de comprendre les déterminants de sa variabilité. Dans un second temps, il s'agissait d'étudier les relations de complémentarité ou substituabilité entre ce score et différentes proxy de motivation extrinsèque (ME).

Les résultats du modèle multiniveaux utilisé pour construire l'ICMI ont montré qu'il existait une forte variabilité inter-médecin dans la distribution du score. De fait, nous avons introduit dans le modèle différentes variables caractérisant les médecins, afin d'analyser les déterminants de cette hétérogénéité. Nous avons tout d'abord montré que les médecins exerçant en groupe (partageant les dossiers) sont davantage caractérisés par des niveaux élevés de MI par rapport aux médecins exerçant en cabinet seul. Ce résultat est d'autant plus intéressant que de précédentes études avaient mis en évidence que le regroupement des médecins était associé à des incitations d'ordre extrinsèque, liées par exemple au fait de pouvoir bénéficier d'horaires de travail plus souples, d'une charge de travail réduite et / ou d'un partage des coûts liés à l'organisation du cabinet (Bourgueil et al 2007). Le fait que nous ayons considéré une définition plus précise de la pratique en groupe, c'est à dire avec partage des dossiers patients, peut avoir amené à sélectionner des médecins acceptant de s'exposer au jugement des pairs et de s'impliquer dans des prises de décision collégiales, et donc des médecins ayant des niveaux de MI plus élevés. Deux autres variables avaient un impact positif sur le niveau de motivation intrinsèque : le fait d'avoir une activité salariée en plus de l'activité libérale, pouvant témoigner d'une volonté de diversifier son activité, et le fait de déclarer un très bon état de santé. Toutefois, près des trois quarts de la variabilité inter-médecin associée à l'ICMI restait inexplicée dans le modèle. La variance résiduelle pourrait correspondre à des caractéristiques médecin non observées, ou d'autres caractéristiques non mesurables d'ordre psychologique par exemple.

Dans l'ensemble, cette étude a montré des résultats mitigés quant aux relations de substituabilité ou complémentarité entre MI et ME. En effet, les résultats diffèrent selon le type d'indicateur de ME utilisé. En considérant comme indicateur de ME, le chiffre d'affaire et le chiffre d'affaire moyen par acte, nous avons montré que l'ICMI n'était pas corrélé aux ME. Nous avons par ailleurs mené des analyses complémentaires considérant l'ICMI construit par le modèle de Birnbaum, mais celles-ci n'ont pas révélé de résultats différents. Toutefois, le fait que près d'un MG sur cinq (19%) n'ait pas reporté son revenu de 2008 pourrait en partie biaiser les résultats de cette analyse et invite à considérer d'autres proxy des ME. En considérant comme indicateur de ME des variables plus qualitatives, nous avons obtenu plusieurs relations négatives significatives avec l'ICMI : les médecins déclarant se sentir contraints par les exigences des patients concernant les horaires d'ouverture du cabinet ou les durées de consultations avaient un niveau plus faible de MI. De façon similaire, les médecins plus motivés intrinsèquement semblaient avoir des exigences plus faibles concernant leur niveau de rémunération car, à niveau de revenu égal, ils étaient plus souvent satisfaits de leur rémunération. Ces résultats semblent aller dans le sens d'une relation de substituabilité entre MI et ME. Par exemple, les médecins se sentant contraints concernant la durée de leur consultation pourraient raisonner en termes d'arbitrage économique en considérant le coût d'opportunité plus élevé d'une consultation plus longue en termes notamment de perte d'activité économique. En effet, comme nous l'avons vu dans le Chapitre 1, les durées de consultation plus longues sont associées à un nombre d'actes plus faible dans l'année et dans un contexte du paiement à l'acte, une activité plus faible se traduit généralement par une perte de chiffre d'affaire.

Conclusion

Cette étude montre l'existence d'une relation négative entre motivations intrinsèques et extrinsèques des MGs français. En ce sens, cette étude relance le débat sur la pertinence et les effets à long terme des incitatifs économiques appliqués aux soins primaires. En effet, le régulateur devrait prendre en compte les effets indésirables potentiels d'une augmentation des motivations extrinsèques par l'introduction de bonus financiers de type P4P. Afin de confirmer et de préciser les résultats de cette étude exploratoire, les études futures devraient se concentrer tout d'abord sur la mesure des MI dans le contexte des soins primaires par l'élaboration de questionnaires plus standardisés. Par ailleurs, il serait intéressant de recueillir des données longitudinales permettant d'analyser l'évolution des motivations intrinsèques des

médecins dans le temps, de façon à tester de manière plus précise la théorie de l'effet d'éviction des motivations. En effet, du fait du caractère transversal des données utilisées, nos résultats ne permettent pas de conclure quant à une relation de cause à effet entre MI et ME. En revanche, cette étude, combinée aux analyses des sections 2 et 3, peut être considérée comme préliminaire à un tel travail, dans le sens où elle a permis de mettre en évidence plusieurs indicateurs de MI ainsi que différentes méthodes statistiques permettant de mesurer ce concept dans le domaine des soins de premiers recours.

Annexes

Annexe C.1. Questionnaire

Identifiant n°

A - Organisation de votre pratique

1. Nature de l'exercice Libéral Libéral + salarié hospitalier Libéral + vacations ou autres (*élu professionnel, pompier, autre*)

2. Vous exercez en
a - Secteur 1
b - Secteur 1 avec droit de dépassements
c - Secteur 2

3. ... dans un
a - Cabinet individuel
b - Cabinet de groupe de médecins généralistes Nbre de médecins :.....
c - Cabinet de groupe de médecins de différentes spécialités Nbre de médecins :.....
d - Maison de santé pluridisciplinaire / centre de santé
e - Autre, quelle structure :
.....
.....

4. Si vous exercez en groupe partagez-vous ...

A. les dossiers de vos patients avec vos associés ? Oui Non
B. les dépenses professionnelles ? (salariales, locatives, etc.) Oui Non
b- Si oui, quelle est votre part moyenne des% dépenses :

5. Avez-vous informatisé vos dossiers patients ? Oui Non

6. « Télétransmettez-vous » à l'assurance maladie ? Oui Non

7. Si oui, quelle part des actes que vous réalisez :% des consultations% des visites

8. Pour la prise de RDV, avez-vous recours

à...

A - un standard téléphonique de prise de RDV

B - un(e) secrétaire salarié(e)

C - un(e) conjoint(e) « collaborateur »

D - vous prenez vous-même vos rendez-vous

E - - autre

9. Quelles ont été vos dépenses professionnelles au niveau individuel en 2008 ?

(Déclaration N° 2035 A -2008)

A- Charges salariales (secrétariat, etc.) [BB]+[BC] €

B- Loyer et charges locatives [BF] €

C- Achats [BA] €

D- Frais de transports [BJ] €

E- Autres charges [BR] – [charges mentionnées ci-dessus] €

F- Honoraires rétrocedés (remplaçants) [AC] €

10. Recevez-vous les patients sur RDV ?

Oui, uniquement Oui, la plupart du temps Non, pas souvent Non, jamais

11. Recevez-vous les visiteurs médicaux

Oui, environ 1 fois / semaine Oui, moins régulièrement Non, pas souvent Non, jamais

12. Participez-vous à la Permanence des soins ...

A. comme médecin régulateur ? (*plateforme téléphonique, Samu, etc.*) Oui Non

B. comme médecin de permanence ? (*astreintes, maison médicale de garde, association*) Oui Non

si oui, la nuit et /ou le WE et de façon ponctuelle, régulière ou sur réquisition

13. Echangez-vous en face à face ou au téléphone autour de cas concrets de patients avec ...

	<i>Très souvent</i>	<i>Souvent</i>	<i>De temps en temps</i>	<i>Rarement</i>
A - un / des médecins généralistes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B - un / des médecins spécialistes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C - un / des infirmiers	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D - un / des pharmaciens	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E - autre, qui ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14. Quel(s) professionnel(s) avez-vous le plus de difficultés à solliciter pour vos patients du fait de

(plusieurs réponses possibles)

	<i>Un spécialiste en ville</i>	<i>Un spécialiste à l'hôpital</i>	<i>Un infirmier / paramédical</i>	<i>Un pharmacien</i>
a – ... de l'éloignement	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

b – ... du délai de RDV	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
c – ... du dépassement d'honoraires	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d – ... de sa méconnaissance (identité, compétence, coordonnées)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

B - Contenu de votre activité et caractéristique de votre patientèle

Votre activité

1 - Avez-vous un Mode d'Exercice Particulier ? Oui Non

2 - Si oui lequel ?
.....

3 - Avez-vous une activité spécialisée ? (gynéco, pédiatrie, sport, gériatrie, ...) Oui Non

4 - Si oui laquelle ?
.....

5 - Quelle part de votre temps de travail consacrez-vous à cette ou ces activités ?%

6 - Exercez-vous des activités d'enseignement ? (université ou autre) Oui Non

7 - Êtes-vous maître de stage ? Oui Non

15. Quelles sont les indemnités ou rémunérations annuelles de vos différentes activités en 2008 ?
(Relevé SNIR 2008)

- A- Permanence des soins / gardes NBR [MAJF, MAJN, MAJU] (en nombre)
Indemnités permanence téléphonique MNT [FRG] €
- B- Forfaits de coordination MNT [FC 0-9] €
- C- Aide pérenne à la transmission MNT [FFS] €
- D- Indemnités de déplacement NBR [ID + IF + IK] (en nombre)
Dont frais de déplacement non remboursables MNT [FNR] €
- E- Formations diverses MNT [FF 0-9] + autres sources €
- F- Indemnités Maître de stage €
- G- Aides à l'installation (régionale, Cnamts) €
- H- Indemnités (élu professionnel, autre activité professionnelle) €
- G- Autres ? €

8 - Réalisez-vous ? (plusieurs réponses possibles)

	Très souvent	Souvent	De temps en temps	Rarement
a – Un suivi gynécologique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
b - De la petite chirurgie (plâtre, points de suture, etc.)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

c - De la petite traumatologie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
d - Un test à l'effort	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e - Un électrocardiogramme	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
e - De la prévention du tabac	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
f - De la prévention de l'alcool	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
g - De la prévention contre l'obésité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
h - Un suivi de grossesse	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
i - une prise en charge de toxicomane	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
j - Un suivi pédiatrique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

9 - Participez-vous à un réseau médical de prise en charge? Oui Non

10 - Si oui, le ou
lesquels ? :

11 - Combien de demi-journées travaillez-vous en moyenne par semaine ? demi-journées

12 - Combien d'heures travaillez-vous en moyenne par semaine ? heures

13 - Combien de semaines de vacances prenez-vous en moyenne par an ? semaines

14 - Combien de consultations avez-vous réalisé en 2008 ?
(RIAP 2008)

15 - Combien de visites avez-vous réalisé en 2008 ?
(RIAP 2008)

16 - Quel est le montant remboursable des actes que vous avez réalisés en 2008 ?
(RIAP 2008) €

☛ Votre patientèle

(RIAP 2008)

17 - Nombre de patients dont vous êtes le « médecin traitant »

18 - % de patients exonérés du TM %

19 - % de bénéficiaires de la CMU %

20 - % des actifs ayant eu des IJ %

21 - % de patients de moins de 16 ans %

22 - % de patients de 16 à 59 ans %

23 - % de patients de 60 à 69 ans %

24 - % de patients de 70 ans et plus %

C – Votre satisfaction et vos contraintes professionnelles

	<i>Souvent</i>	<i>De temps en temps</i>	<i>Rarement</i>	<i>Jamais</i>
1 - Êtes-vous satisfait de votre activité professionnelle ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2 - Vous sentez-vous libre de modifier votre agenda ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3 - Vous sentez-vous contraint par les exigences de vos patients sur...				
A- les horaires d'ouverture du cabinet	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B- la durée de vos consultations	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C - vos prescriptions médicamenteuses	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D - vos prescriptions d'analyses et tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
E - vos prescriptions d'arrêts de travail	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
F - les orientations vers les spécialistes dans le cadre du parcours de soins	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4 - Considérez-vous que votre profession est gratifiante sur le plan humain	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5 - Avez-vous un sentiment de <i>burn out</i> ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6 - Êtes-vous confronté à des patients de plus en plus informés ?			<input type="checkbox"/> Oui	<input type="checkbox"/> Non
7 - Si oui, considérez-vous que c'est une contrainte ?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8 - Avez-vous envie de changer ...				
A - de spécialité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
B - de lieu d'installation	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
C - de statut (<i>libéral / salarié</i>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
D - de mode d'installation (<i>individuel ≠ groupe</i>)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9 - Êtes-vous satisfait de la rémunération de votre activité médicale libérale ?			<input type="checkbox"/> Oui	<input type="checkbox"/> Non
10 - Que représentent vos salaires éventuels (bruts) par rapport à vos revenus totaux ?				... %
11 - Vous sentez-vous contraints par le contrôle de l'assurance maladie ?			<input type="checkbox"/> Oui	<input type="checkbox"/> Non
12 - Si oui, est-ce que cela modifie votre pratique ?			<input type="checkbox"/> Oui	<input type="checkbox"/> Non
13 - Rencontrez-vous des difficultés particulières dues au profil de votre patientèle ?			<input type="checkbox"/> Oui	<input type="checkbox"/> Non
14 - Si oui, lesquelles ?				
.....				

.....
.....
.....
.....

15 - Avez-vous déjà ressenti un sentiment d'insécurité dans votre exercice Oui Non

D – Quelques questions supplémentaires

- | | | | | |
|---|--|---|--|--|
| 1. Sexe : | <input type="checkbox"/> Homme | <input type="checkbox"/> Femme | | |
| 2. Age : | ans | | | |
| 3. Année d'installation : | | | | |
| 4. Avez-vous changé de mode d'installation dans les 5 dernières années (<i>individuel ⇄ groupe</i>) | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |
| 5. Département | | | | |
| 6. Dans une commune de ... habitants | <input type="checkbox"/> Moins de 10 000 | <input type="checkbox"/> De 10 à 20 000 | <input type="checkbox"/> De 20 à 100 000 | <input type="checkbox"/> Plus de 100 000 |
| 7. Exercez-vous en Zone Franche Urbaine (ZFU) | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |
| 8. Exercez-vous en Zone de Revitalisation Rurale (ZRR) | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |
| 9. Avez-vous un ou des enfants à charge de moins de 15 ans | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |
| 10. Si oui quel âge a le plus jeune | ... ans | | | |
| 11. Comment percevez-vous votre santé ? | Excellente <input type="checkbox"/> | Très bonne <input type="checkbox"/> | Bonne <input type="checkbox"/> | Médiocre <input type="checkbox"/> |
| 12. Êtes-vous abonné à une revue professionnelle ? | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |
| 13. Suivez-vous des cycles de formation médicale continue | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |
| 14. Avez-vous déjà participé à une action d'Evaluation des Pratiques Professionnelles | <input type="checkbox"/> Oui | <input type="checkbox"/> Non | | |

Annexe C.2. Graphiques additionnels

Figure C.2.1. Distribution de l'indicateur composite de motivations intrinsèques

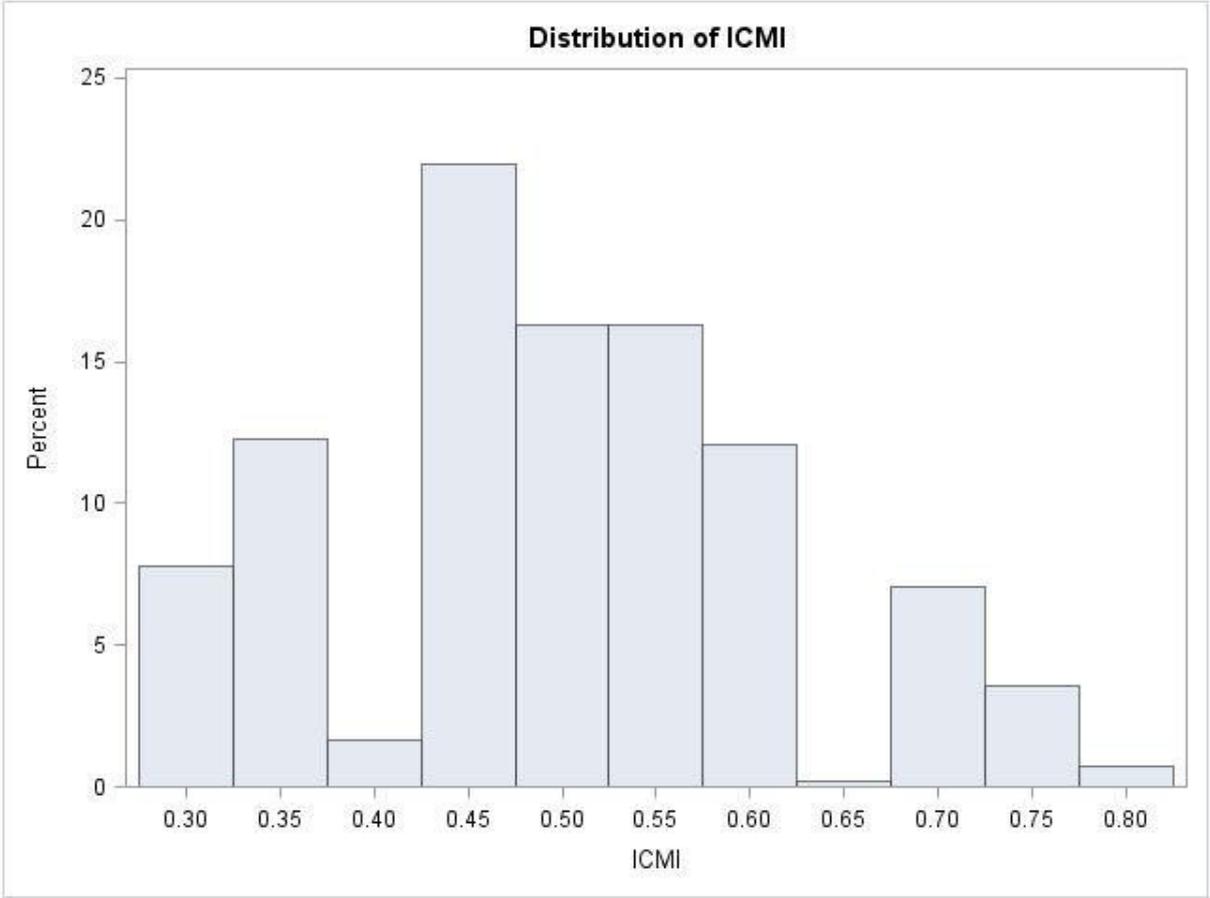
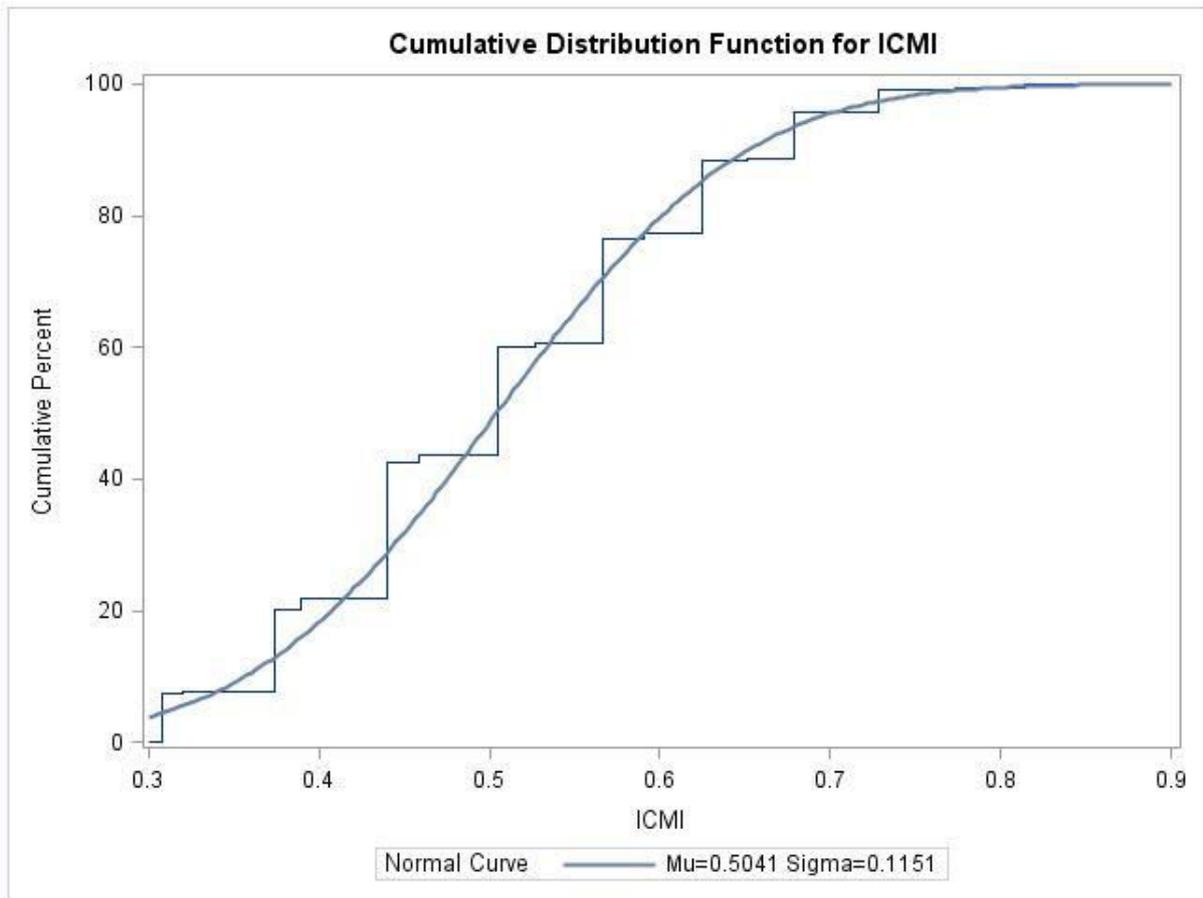


Figure C.2.2. Distribution cumulée de l'indicateur composite de motivation intrinsèque



Conclusion générale

Cette thèse s'intéresse à plusieurs questions posées par l'introduction, en France, d'incitatifs économiques de type paiements à la performance (P4P) appliqués à la médecine générale. Les P4P visent à intéresser financièrement les professionnels de santé aux résultats de leurs prises en charge médicale tels que mesurés par différents indicateurs, en allouant des bonus financiers en contrepartie de l'atteinte d'un objectif cible. De fait, ils ont donc un coût direct pour le système de santé (Cour des comptes 2014), mais aussi un coût indirect lié à leurs effets indésirables potentiels. Pourtant, peu d'études ont été menées en France pour comprendre, d'une part, l'impact de ces politiques incitatives (CAPI, ROSP) sur les pratiques des médecins généralistes (MGs) et, d'autre part, quelles alternatives pourraient, du point de vue des médecins eux-mêmes, compléter les dispositifs existants. Par ailleurs, il nous a semblé essentiel de comprendre les motivations intrinsèques des MGs concernant la fourniture de soins de qualité et leur interrelation avec les motivations financières. Dans un contexte où l'impact des P4P est débattu à la fois en termes d'efficacité et d'efficience, ces différentes questions et les approches mises en œuvre pour y répondre apparaissent complémentaires, et permettent de dresser un certain nombre d'hypothèses qui constituent autant de pistes de recherches futures.

Dans le **chapitre 1**, nous nous sommes intéressés aux effets des P4P et plus particulièrement, du CAPI, sur différents indicateurs de la qualité des soins. Construire un indicateur de qualité des soins est une démarche complexe, du fait de la difficulté à observer et à définir la qualité notamment dans le contexte des soins primaires caractérisé par une forte hétérogénéité des pratiques et des attentes des patients tout aussi variées. Aussi discutables que soient ces indicateurs, nous avons fait l'hypothèse qu'ils pouvaient du moins en partie approximer la qualité des soins en caractérisant au moins une des dimensions. Le premier indicateur que nous avons considéré est la durée de consultation. Cet indicateur est particulièrement pertinent dans le contexte du paiement à l'acte en France qui n'incite pas à une durée de consultation élevée. Nous avons mobilisé les données du projet Ecogen coordonné par le Professeur Laurent Letrilliart, qui compte plus de 20 000 observations de consultations en médecine générale (durée, motifs, résultats, caractéristiques des médecins et des patients). L'objectif de notre étude était d'analyser la relation entre signature du CAPI et durée de consultation, afin d'établir des hypothèses quant aux effets des incitatifs financiers

sur cet indicateur de qualité. Pour cela, nous avons mis en œuvre des modèles multi-niveaux, permettant de distinguer la part de la variabilité des durées de consultation liée aux caractéristiques des séances de celle liée aux caractéristiques des médecins. Nous avons montré que les médecins signataires du CAPI avaient en moyenne une durée de consultation significativement plus courte que les non signataires. En revanche, cet effet n'apparaissait plus significatif lorsque nous contrôlions du nombre de consultation effectuées en 2011, deux ans après la mise en place du CAPI. Ces résultats nous ont permis de proposer deux hypothèses quant à l'impact du CAPI sur la durée de la consultation. La première hypothèse est que le CAPI s'est traduit par une augmentation de l'activité médicale et de façon concomitante par une baisse de la durée de la consultation et donc de la qualité. Toutefois, cette hypothèse nous a semblé peu convaincante compte tenu du faible montant des bonus CAPI (environ 7% du revenu des MGs). Par ailleurs, un tel effet n'a jamais été montré dans des études internationales où, même si l'organisation des soins est différente, les montants de P4P sont généralement plus élevés. Dans la deuxième hypothèse, la durée de consultation plus faible des signataires du CAPI témoigne de phénomènes d'auto-sélection et d'un arbitrage entre nombre de consultations/durée de la consultation en faveur de la quantité plutôt que de la qualité. Quelque soit l'hypothèse retenue, nous montrons que le CAPI n'a pas eu d'impact significatif sur la qualité telle que mesurée par la durée de la consultation. Un prolongement de cette étude consisterait collecter et à analyser des données concernant l'évolution de l'activité médicale liée à la mise en place du CAPI. Une telle étude nécessiterait d'obtenir des données agrégées par médecin et en panel, et permettrait de conclure plus précisément quand à l'impact du CAPI sur cette dimension de la qualité des soins.

Le deuxième indicateur que nous avons considéré pour une évaluation plus ciblée du CAPI est le recours au dépistage des cancers. En effet, d'une part la prévention / dépistage est l'un des trois volets principaux du CAPI, d'autre part, le dépistage des cancers constitue un enjeu de santé publique majeur et pour lequel le rôle du médecin généraliste est reconnu. Dans un premier temps et pour bien comprendre le contexte précis de l'évaluation du CAPI, nous nous sommes intéressés, côté demande, aux déterminants du recours au dépistage des cancers (sein, col de l'utérus, côlon) et à leur évolution. L'objectif était d'identifier les freins et leviers au dépistage des cancers côté demande, alors que certains « dépistages organisés » existent et que des campagnes d'information et d'éducation parfois très médiatisées sont menées régulièrement. Dans ce but, nous avons analysé trois vagues de l'Enquête Santé Protection Sociale (ESPS) menées par l'IRDES, entre 2006 et 2010. Nous avons montré que

le recours au système de santé (mesuré au travers des visites chez le MG et le spécialiste), ainsi que son accès (mesuré par le fait de bénéficier d'une assurance maladie complémentaire), constituaient des déterminants importants du recours au dépistage de ces trois cancers. Par ailleurs, nous avons mis en évidence d'autres freins du recours au dépistage : l'adoption de comportements à risques (alcool, tabac) ou au contraire, de comportements parfaitement sains, pouvant témoigner d'arbitrages économiques différents (préférences temporelles, niveaux d'aversion au risque). Une fois ce contexte clarifié, dans un second temps, nous avons analysé l'impact du CAPI sur le recours au dépistage du cancer du sein. Nous avons testé l'hypothèse selon laquelle le fait d'avoir vu son MG traitant, et signataire du CAPI, augmentait la probabilité d'effectuer une mammographie pour une femme éligible dans le cadre des recommandations. Nous avons analysé les données de consommations de soins de plus de 60 000 femmes âgées entre 50 et 74 ans à partir de l'Echantillon Généralistes des Bénéficiaires (EGB). Nous avons montré un effet positif, mais statistiquement non significatif, d'une consultation auprès de son MG traitant signataire du CAPI, sur le recours à la mammographie. Ce résultat non significatif pourrait témoigner du fait que, dans le domaine particulier de la prévention / dépistage dans lequel l'exercice du médecin se heurte à de multiples freins (inégalités socioéconomiques, absence d'information quant à la réalisation du dépistage), les paiements à la performance ne constituent pas un levier suffisamment fort pour améliorer les pratiques de dépistage. Ce résultat nous a donc conforté dans la démarche d'étudier précisément les outils complémentaires aux P4P susceptibles d'être implémentés afin d'encourager les médecins généralistes à accroître leur offre de prévention et de dépistage des cancers.

Dans le **chapitre 2**, nous nous sommes intéressés aux préférences des MGs français concernant différents dispositifs susceptibles d'être mis en œuvre afin d'améliorer les pratiques de dépistage des cancers du sein, du col de l'utérus et du côlon. Nous souhaitons tester l'hypothèse selon laquelle, outre les incitatifs économiques, les MGs sont sensibles à d'autres types de dispositifs, non monétaires, permettant d'améliorer leur pratique. A partir d'une revue de la littérature et d'entretiens auprès d'experts, nous avons définis cinq dispositifs : quatre étaient de nature non monétaires, dont un concernait des transferts en nature (financement d'une formation), et le cinquième était un dispositif de rémunération additionnelle de type ROSP mais proposant des bonus plus élevés. Nous avons mobilisé la méthodologie des choix discrets afin de construire un questionnaire permettant de faire révéler aux médecins leurs préférences pour ces dispositifs. Plus précisément, le premier

objectif était de révéler les dispositifs ou combinaisons de dispositifs les plus plébiscités par les médecins, selon le type de cancer. Le deuxième objectif était d'analyser l'hétérogénéité des préférences afin d'identifier des profils de médecins plus ou moins sensibles à certains types de dispositifs. De fait, deux types de méthodes ont été mises en œuvre : la méthode des choix discrets (DCE) d'une part, et la méthode de classement des attributs du meilleur au pire (BWAS), d'autre part.

Tout d'abord, nous avons montré que les dispositifs plébiscités par les MGs dépendent de la nature des cancers ; ainsi, il ne semble pas y avoir un dispositif qui domine les autres pour tous les cancers (y compris la rémunération additionnelle). Nous avons également montré que, dans le cadre du dépistage des cancers féminins, les MGs semblent particulièrement sensibles à des transferts en nature (financement d'une formation sur le dépistage du cancer du sein ou du col de l'utérus) ainsi qu'au fait de disposer d'un listing pluriannuel des patientes dépistées, pour identifier les femmes éligibles non à jour de leur dépistage. Il est intéressant de noter que ces dispositifs seraient réalisables en pratique et ne seraient pas forcément aussi coûteux que des incitatifs financiers. En outre, concernant la méthode DCE, les résultats des taux marginaux de substitution calculés entre les attributs dits non monétaires et la rémunération additionnelle ont révélé que les MGs seraient prêts à renoncer à la totalité de leurs incitations financières et même prêts à payer « de leur poche » afin de bénéficier de certains de ces dispositifs. Ce résultat confirme l'hypothèse selon laquelle les motivations des médecins associées à des pratiques de qualité ne sont pas principalement financières.

Cette étude confirme les résultats de précédents travaux ayant montré l'importance des caractéristiques non monétaires dans la pratique des médecins (Scott 2001; Janus et al 2007; Ammi and Peyron 2010). Elle a permis de mettre en évidence que les médecins effectuent des arbitrages entre dispositifs monétaires d'une part et non monétaires d'autre part, et que les premiers ne sont pas forcément les plus plébiscités. Ce résultat apparaît d'autant plus important dans le contexte de renforcement des politiques de paiements à la performance, qui pourraient ne pas être les plus efficaces pour engendrer une amélioration des pratiques, du moins dans le domaine particulier de la prévention / dépistage des cancers. En outre, le modèle économique sur lequel repose le principe des P4P, issu principalement des résultats des modèles principal-agent (Laffont and Martimort 2002), est basé sur des hypothèses potentiellement discutables du comportement humain (Gneezy and Rustichini 2000; Fehr and

Falk 2002; Ellingsen and Johannesson 2006; Bowles 2008) ; parmi lesquelles, l'hypothèse d'indépendance entre motivations intrinsèques et extrinsèques (Frey 1997b).

Dans le **chapitre 3**, nous nous sommes intéressés à la problématique de l'interrelation entre motivations intrinsèques et extrinsèques dans le domaine des soins de premiers recours et plus particulièrement de la médecine générale. L'objectif était d'analyser empiriquement les relations de complémentarité ou substituabilité entre MI et ME, de façon à mieux éclairer les effets indésirables potentiels associés à l'augmentation des ME par l'instauration d'incitatifs financiers (les P4P). Nous nous sommes basés sur les résultats d'une enquête menée en 2010 auprès de 423 MGs de la région Midi-Pyrénées. Le premier objectif était de définir et de valider différentes mesures de la MI des MGs Français. Sur la base d'une revue de la littérature et de l'utilisation de techniques d'analyse factorielle, nous avons validé 9 indicateurs qui représentaient tous une seule dimension des MI. Ces indicateurs incluaient notamment le sentiment de satisfaction ou de gratification humaine de l'activité (Deci and Ryan 2000), ainsi que des activités témoignant d'une volonté d'améliorer la qualité des soins, notamment, la participation à un réseau de soins, les actions de prévention ou les échanges fréquents (avec d'autres professionnels de santé) autour de cas concrets de patients. Le deuxième objectif était d'identifier puis de comparer les résultats de différentes méthodes de construction d'un score composite de MI des MGs français. Nous avons montré une forte adéquation dans les résultats de chaque méthode témoignant de la robustesse des méthodes et de la validité de la mesure obtenue. Enfin, à partir du score de MI validé, nous avons analysé les relations entre MI et ME telles que mesurées par différents indicateurs présélectionnés (chiffre d'affaire, satisfaction vis-à-vis du revenu, sentiments d'être contraints par les patients). Nous avons montré une relation négative entre le score de MI et trois indicateurs de ME, reflétant d'une part un arbitrage économique en termes de calcul coûts/bénéfices plus marqué, et d'autre part des exigences en termes de rémunération plus élevées, pour les médecins moins motivés intrinsèquement. Ces relations négatives pourraient témoigner d'une substituabilité entre MI et ME. Même si la nature transversale des données ne nous permet pas de conclure directement à un effet d'éviction des MI par les ME, ces résultats invitent à une certaine prudence dans la conception et la généralisation des politiques d'incitations financières en médecine générale.

Pistes de recherches futures

Tout au long de cette thèse, nous avons montré des résultats très faibles et souvent non significatifs concernant l'impact du CAPI sur certains indicateurs de qualité des soins. Néanmoins, il ne s'agit pas de conclure qu'un mécanisme de P4P en général n'a pas de conséquence sur d'autres dimensions de la qualité approchées par différents indicateurs. En revanche, il est important de produire des analyses précises permettant d'évaluer l'efficacité réelle du dispositif dans différents domaines de la prise en charge, de façon à mieux réguler et améliorer la performance du système. Ce travail devra donc être complété par l'analyse de l'évolution d'autres indicateurs définis par le CAPI ou la ROSP. En premier lieu, il serait intéressant d'analyser l'évolution du recours au dépistage du cancer du col de l'utérus suite à la mise en place de la ROSP en 2012. Comme ce dispositif a été instauré à destination de l'ensemble des médecins généralistes, des techniques statistiques différentes devront être mobilisées (Serumaga et al 2011; McLintock et al 2014; Wang et al 2014). En second lieu, il serait pertinent d'analyser l'évolution d'autres types d'indicateurs relatifs au suivi des maladies chroniques, tels que par exemple le diabète ou l'hypertension artérielle, pour lesquels la littérature a, là encore, montré des effets contrastés (Lippi Bruni et al 2009; Serumaga et al 2011).

Les résultats de cette thèse invitent à une certaine prudence dans la conception des indicateurs de P4P et des modalités de versement des bonus. En premier lieu, comme pour toute politique économique, il est nécessaire de définir au préalable et précisément les objectifs du P4P, sachant que tous les indicateurs mobilisés sont susceptibles de ne pas servir ces objectifs comme attendu. En effet, les objectifs d'un P4P peuvent être variés : s'agit-il d'améliorer la qualité, de réduire les inégalités et/ ou de réduire les coûts ? Il est intéressant de rappeler que certains de ces objectifs sont incompatibles entre eux. Ainsi, il faut définir le type d'indicateur le plus pertinent pour l'objectif fixé: s'agit-il par exemple d'un indicateur de processus ou d'un indicateur de résultat ? (Bras et Duhamel 2008). Enfin, il est nécessaire de prendre en compte le contexte précis de prise en charge des patients, ainsi que les obstacles qui caractérisent une prise en charge de qualité. Comme nous l'avons vu dans le cadre du recours au dépistage des cancers, côté demande, de nombreux freins principalement d'ordre socioéconomiques constituent potentiellement des obstacles importants à l'engagement des médecins dans leur pratique de dépistage ciblée. Ainsi, les dispositifs (qu'ils soient monétaires ou non) à mettre en œuvre pour notamment améliorer la transmission de l'information entre le médecin et le patient devraient prendre en compte les déterminants de la

prise de décision autour du dépistage (Protière et al 2012). Ce type d'étude sur le partage d'information se situe dans le prolongement de certains travaux qui se sont intéressés aux attitudes des médecins et de leurs patients concernant la décision partagée pour l'adoption d'un traitement (Moumjid et al 2003; Nguyen et al 2014) ou encore aux préférences des patients concernant leur prise en charge pour un problème de santé chronique (Krucien et al 2014).

Les résultats de cette thèse révèlent la nécessité d'analyser de façon plus approfondie les comportements stratégiques des médecins associés à une pratique efficiente. Un premier prolongement de ce travail consisterait à développer des outils plus standardisés de la mesure des MI, lesquelles peuvent potentiellement remettre en question l'efficacité des incitatifs économiques. Il serait intéressant d'étudier de façon plus précise les conditions favorisant l'émergence des MI en médecine générale, ainsi que leur variabilité en fonction des conditions de la pratique du MG. En particulier, nous avons mis en évidence que les médecins travaillant en groupe avaient des MI plus élevées : ce résultat est notamment intéressant dans le contexte actuel d'incitations au regroupement des médecins (Mousquès 2011). Néanmoins, la pratique de groupe pose potentiellement d'autres types de problèmes : quelles modalités de distribution des bonus liés à la performance, quelle répartition des rôles et quelle gouvernance ? Ces problèmes sont susceptibles de modifier les résultats d'une évaluation des P4P en France, voire de modifier le design requis d'un P4P en lui-même, qui devrait alors être adapté à la pratique de groupe. A l'hôpital, des mécanismes d'incitations et des P4P existent, dont il faudrait peut être s'inspirer. Ces questions constituent des pistes de recherche pour approfondir nos connaissances sur la désirabilité et l'efficacité des P4P dans le contexte de l'offre de soins de premiers recours en France, et ainsi contribuer à la recherche d'outils permettant d'améliorer l'efficacité du système de santé.

Bibliographie

- Agurto I, Bishop A, Sánchez G, et al (2004) Perceived barriers and benefits to cervical cancer screening in Latin America. *Prev Med* 39:91–98. doi: 10.1016/j.jpmed.2004.03.040
- Amabile TM (1993) Motivational synergy: Toward new conceptualizations of intrinsic and extrinsic motivation in the workplace. *Hum Resour Manag Rev* 3:185–201. doi: 10.1016/1053-4822(93)90012-S
- Amabile TM, DeJong W, Lepper MR (1976) Effects of externally imposed deadlines on subsequent intrinsic motivation. *J Pers Soc Psychol* 34:92–98. doi: 10.1037/0022-3514.34.1.92
- Ammi M, Peyron C (2010) Incitations à l'offre de prévention et préférences en médecine générale : l'apport de la méthode DCE. *Économie Publique/Public Econ* 129–155.
- Anderson R, Manoogian ST, Steven J (1976) The undermining and enhancing of intrinsic motivation in preschool children. *J Pers Soc Psychol* 34:915–922. doi: 10.1037/0022-3514.34.5.915
- Argyris C (1964) *Integrating the individual and the organization*. Wiley, New York
- Arrossi S, Ramos S, Paolino M, Sankaranarayanan R (2008) Social inequality in Pap smear coverage: identifying under-users of cervical cancer screening in Argentina. *Reprod Health Matters* 16:50–58. doi: 10.1016/S0968-8080(08)32410-0
- Arrow KJ (1963) Uncertainty and the welfare economics of medical care. *Am Econ Rev* 941–973.
- Asselin L, Vu T (2008) Multidimensional poverty and multiple correspondence analysis. *Quant Approaches Multidimens Poverty Meas* 80–103.
- Assurance Maladie (2010) CAPI, 1ers résultats après une année d'existence. *Assur. Mal.*
- Azogui-Levy S, Rochereau T (2005) Comportements de recours aux soins et santé bucco-dentaire : Exploitation de l'enquête « santé et protection sociale » 2000. *Quest Déconomie Santé* 1–8.
- Bartholomew D, Knott M (1999) *Latent Variable Models and Factor Analysis*.
- Baumgartner H, Homburg C (1996) Applications of structural equation modeling in marketing and consumer research: A review. *Int J Res Mark* 13:139–161. doi: 10.1016/0167-8116(95)00038-0
- Becker GS (1976) *The Economic Approach to Human Behavior*. University of Chicago Press
- Becker GS, Mulligan CB (1997) The Endogenous Determination of Time Preference. *Q J Econ* 112:729–758. doi: 10.1162/003355397555334

- Beck F, Gautier A (2010) Les pratiques de dépistage des cancers en France. Baromètre cancer 2010. Inst. Natl. Prév. Educ. Pour Santé INPES
- Bedos C, Brodeur J-M, Benigeri M, Olivier M (2004) Inégalités sociales dans le recours aux soins dentaires. *Rev DÉpidémiologie Santé Publique* 52:261–270. doi: 10.1016/S0398-7620(04)99051-7
- Bellamy V (2010) Les revenus libéraux des médecins en 2007 et 2008. *Etudes Résultats* 1–8.
- Bell DNF, Hart RA (1999) Unpaid Work. *Economica* 66:271–290. doi: 10.1111/1468-0335.00169
- Bénabou R, Tirole J (2005) Incentives and Prosocial Behavior. National Bureau of Economic Research
- Bénabou R, Tirole J (2003a) Intrinsic and Extrinsic Motivation. *Rev Econ Stud* 70:489–520. doi: 10.1111/1467-937X.00253
- Bénabou R, Tirole J (2003b) Intrinsic and Extrinsic Motivation. *Rev Econ Stud* 70:489–520. doi: 10.1111/1467-937X.00253
- Benware CA, Deci EL (1984) Quality of Learning With an Active Versus Passive Motivational Set. *Am Educ Res J* 21:755–765. doi: 10.3102/00028312021004755
- Benzécri J (1979) Sur le calcul des taux d’inertie dans l’analyse d’un questionnaire. *Cah Anal Données* 4:377–8.
- Berchi C, Dupuis J-M, Launoy G (2006) The reasons of general practitioners for promoting colorectal cancer mass screening in France. *Eur J Health Econ HEPAC Health Econ Prev Care* 7:91–98. doi: 10.1007/s10198-006-0339-0
- Berchi C, Launoy G (2007) [Principle, strengths and weaknesses of discrete choice modelling for eliciting public preferences for health care]. *Rev Dépidémiologie Santé Publique* 55:133–139. doi: 10.1016/j.respe.2006.11.002
- Bingham A, Bishop A, Coffey P, et al (2003) Factors affecting utilization of cervical cancer prevention services in low-resource settings. *Salud Pública México* 45 Suppl 3:S408–416.
- Birnbaum A (1968) Some latent trait models and their use in inferring an examinee’s ability. *Stat. Theor. Ment. Test Scores*, Lord & M.R. Novick. Reading, MA: MIT Press, pp 397–479
- Bock RD, Aitkin M (1981) Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika* 46:443–459. doi: 10.1007/BF02293801
- Bolton P, Dewatripont M (1957) *Contract Theory*, MIT Press. Cambridge
- Borrell C, Fernandez E, Schiaffino A, et al (2001) Social class inequalities in the use of and access to health services in Catalonia, Spain: what is the influence of supplemental

- private health insurance? *Int J Qual Health Care* 13:117–125. doi: 10.1093/intqhc/13.2.117
- Bourgueil Y, Chambaretaud S, Marek A, Mousquès J (2007) Les nouvelles formes de coopération entre professionnels de santé: premiers enseignements des expériences internationales. *Enjeux Économiques Coop Entre Prof Santé* 46–59.
- Bowles S (2008) Policies Designed for Self-Interested Citizens May Undermine “The Moral Sentiments”: Evidence from Economic Experiments. *Science* 320:1605–1609. doi: 10.1126/science.1152110
- Bras et Duhamel (2008) Rémunérer les médecins selon leurs performances : les enseignements des expériences étrangères. <http://www.ladocumentationfrancaise.fr/rapports-publics/084000596/>. Accessed 10 Sep 2014
- Breuil-Genier P, Goffette C (2006) La durée des séances des médecins généralistes. *Etudes Résultats DREES*
- Brodin U, Fors U, Laksov KB (2010) The application of Item Response Theory on a teaching strategy profile questionnaire. *BMC Med Educ* 10:14. doi: 10.1186/1472-6920-10-14
- Brosig-Koch J, Hennig-Schmidt H, Kairies N, Wiesen D (2013) How effective are pay-for-performance incentives for physicians? A laboratory experiment. *Ruhr Econ. Pap.*
- Browne M., Cudeck R (1993) Alternative ways of assessing model fit. *Test. Struct. Equ. Models*, A.Bollen & J.S Long. Sage, Newsbury Park, pp 136–162
- Bungener M, Eisinger F, Aubin-Auger I (2010) Médecins généralistes et dépistage des cancers. Synthèse des résultats de l'enquête barométrique INCa/BVA Septembre 2010.
- Bush K, Kivlahan DR, McDonell MB, et al (1998) The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. *Arch Intern Med* 158:1789–1795.
- Campbell SM, Reeves D, Kontopantelis E, et al (2009) Effects of Pay for Performance on the Quality of Primary Care in England. *N Engl J Med* 361:368–378. doi: 10.1056/NEJMs0807651
- Carcaise-Edinboro P, Bradley CJ (2008) Influence of Patient-Provider Communication on Colorectal Cancer Screening. *Med Care* 46:738–745. doi: 10.1097/MLR.0b013e318178935a
- Carrère M-O, Moumjid-Ferdjaoui N, Charavel M, Brémond A (2000) Eliciting patients' preferences for adjuvant chemotherapy in breast cancer: development and validation of a bedside decision-making instrument in a French Regional Cancer Centre. *Health Expect Int J Public Particip Health Care Health Policy* 3:97–113.
- Carrieri V, Bilger M (2013) Preventive care: underused even when free. Is there something else at work? *Appl Econ* 45:239–253. doi: 10.1080/00036846.2011.597729

- Cathébras P, Begon A, Laporte S, et al (2008) Épuisement professionnel chez les médecins généralistes. /data/revues/07554982/00330022/1569/
- Chanel O, Paraponaris A, Protière C, Ventelou B (2010) Get paid more, work more? Lessons from French physicians' labour supply responses to hypothetical fee increases.
- Chattopadhyay SK, Ebrahim SH, Tao G, McKenna MT (2005) Use of cervical cancer screening among insured women: the extent of missed opportunities. *Health Policy* 73:194–201. doi: 10.1016/j.healthpol.2004.11.012
- Cheng J, Pullenayegum E, Marshall DA, et al (2012a) An empirical comparison of methods for analyzing correlated data from a discrete choice survey to elicit patient preference for colorectal cancer screening. *BMC Med Res Methodol* 12:15. doi: 10.1186/1471-2288-12-15
- Cheng S-H, Lee T-T, Chen C-C (2012b) A Longitudinal Examination of a Pay-for-Performance Program for Diabetes Care. *Med Care* 50:109–116. doi: 10.1097/MLR.0b013e31822d5d36
- Cheng S, Long JS (2007) Testing for IIA in the Multinomial Logit Model. *Sociol Methods Res* 35:583–600. doi: 10.1177/0049124106292361
- Choné P, Ma CA (2011) Optimal health care contract under physician agency. *Ann Econ Stat D'Économie Stat* 229–256.
- Christoffersson A (1975) Factor analysis of dichotomized variables. *Psychometrika* 40:5–32. doi: 10.1007/BF02291477
- Cohen E (2009) Applying best-worst scaling to wine marketing. *Int J Wine Bus Res* 21:8–23. doi: 10.1108/17511060910948008
- Cohen J (1968) Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 70:213–220.
- Commission des comptes (2011) Le contrat d'amélioration des pratiques individuelles (CAPI) [http://www.securite-sociale.fr/IMG/pdf/fiche_eclairage_maladie_capi_sept_2011.pdf]
- Com-Ruelle L, Dourgnon P, Jusot F, et al (2005) Identification et mesure des problèmes d'alcool en France: une comparaison de deux enquêtes en population générale. *IRDES* No 97.
- Conrad DA, Perry L (2009) Quality-Based Financial Incentives in Health Care: Can We Improve Quality by Paying for It? *Annu Rev Public Health* 30:357–371. doi: 10.1146/annurev.publhealth.031308.100243
- Cour des comptes (2014) Les relations conventionnelles entre l'assurance maladie et les professions libérales de santé [<https://www.ccomptes.fr/Publications/Publications/Les-relations-conventionnelles-entre-l-assurance-maladie-et-les-professions-liberales-de-sante>]

- Couture M-C, Nguyen CT, Alvarado BE, et al (2008) Inequalities in breast and cervical cancer screening among urban Mexican women. *Prev Med* 47:471–476. doi: 10.1016/j.ypmed.2008.07.005
- Crosson B, Propper C, Perkins A (2001) Do doctors respond to financial incentives? UK family doctors and the GP fundholder scheme. *J Public Econ* 79:375–398. doi: 10.1016/S0047-2727(00)00074-8
- De Bekker-Grob EW, Ryan M, Gerard K (2012) Discrete choice experiments in health economics: a review of the literature. *Health Econ* 21:145–172. doi: 10.1002/hec.1697
- DeCharms R (1968) Personal causation : the internal affective determinants of behavior.
- Deci EL (1971) Effects of externally mediated rewards on intrinsic motivation. *J Pers Soc Psychol* 18:105–115. doi: 10.1037/h0030644
- Deci EL (1972) The effects of contingent and noncontingent rewards and controls on intrinsic motivation. *Organ Behav Hum Perform* 8:217–229. doi: 10.1016/0030-5073(72)90047-5
- Deci EL, Koestner R, Ryan RM (1999a) The undermining effect is a reality after all— Extrinsic rewards, task interest, and self-determination: Reply to Eisenberger, Pierce, and Cameron (1999) and Lepper, Henderlong, and Gingras (1999). *Psychol Bull* 125:692–700. doi: 10.1037/0033-2909.125.6.692
- Deci EL, Koestner R, Ryan RM (1999b) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychol Bull* 125:627–668; discussion 692–700.
- Deci EL, Ryan RM (2000) The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychol Inq* 11:227–268. doi: 10.1207/S15327965PLI1104_01
- De Jaegher K (2010) Physician incentives: Cure versus prevention. *J Health Econ* 29:124–136. doi: 10.1016/j.jhealeco.2009.11.006
- Delattre E, Dormont B (2003) Fixed fees and physician-induced demand: A panel data study on French physicians. *Health Econ* 12:741–754. doi: 10.1002/hec.823
- Dennis J J, Moré J (1977) Quasi-Newton Methods, Motivation and Theory. *SIAM Rev* 19:46–89. doi: 10.1137/1019005
- De Roquefeuil L, Studer A, Neumann A, Merlière Y (2009) L'échantillon généraliste des bénéficiaires: représentativité, portée et limites. *Prat. Organ. Soins* 40:
- Devlin RA, Sarma S (2008) Do physician remuneration schemes matter? The case of Canadian family physicians. *J Health Econ* 27:1168–1181. doi: 10.1016/j.jhealeco.2008.05.006
- Dewatripont M, Jewitt I, Tirole J (2000) Multitask agency problems: Focus and task clustering. *Eur Econ Rev* 44:869–877. doi: 10.1016/S0014-2921(00)00059-3

- Doran T, Fullwood C, Gravelle H, et al (2006) Pay-for-Performance Programs in Family Practices in the United Kingdom. *N Engl J Med* 355:375–384. doi: 10.1056/NEJMsa055505
- Doran T, Fullwood C, Reeves D, et al (2008) Exclusion of patients from pay-for-performance targets by English physicians. *N Engl J Med* 359:274–84. doi: 10.1056/NEJMsa0800310
- Dormont B (2013) Le paiement à la performance : contraire à l'éthique ou au service de la santé publique ? *Trib Santé* 40:53–61. doi: 10.3917/seve.040.0053
- Dumont E, Fortin B, Jacquemet N, Shearer B (2008) Physicians' multitasking and incentives: empirical evidence from a natural experiment. *J Health Econ* 27:1436–1450. doi: 10.1016/j.jhealeco.2008.07.010
- Duport N, Ancelle-Park R (2006) Do socio-demographic factors influence mammography use of French women? Analysis of a French cross-sectional survey. *Eur J Cancer Prev Off J Eur Cancer Prev Organ ECP* 15:219–224. doi: 10.1097/01.cej.0000198902.78420.de
- Eaker S, Adami HO, Sparén P (2001) Attitudes to screening for cervical cancer: a population-based study in Sweden. *Cancer Causes Control CCC* 12:519–528.
- Eggleston K (2005) Multitasking and mixed systems for provider payment. *J Health Econ* 24:211–223. doi: 10.1016/j.jhealeco.2004.09.001
- Eijkenaar F, Emmert M, Scheppach M, Schöffski O (2013) Effects of pay for performance in health care: A systematic review of systematic reviews. *Health Policy Amst Neth* 110:115–130. doi: 10.1016/j.healthpol.2013.01.008
- Ellingsen T, Johannesson M (2006) *Pride and Prejudice: The Human Side of Incentive Theory*. Social Science Research Network, Rochester, NY
- Ellis P, Robinson P, Ciliska D, et al (2005) A Systematic Review of Studies Evaluating Diffusion and Dissemination of Selected Cancer Control Interventions. *Health Psychol* 24:488–500. doi: 10.1037/0278-6133.24.5.488
- Ellis RP, McGuire TG (1986) Provider behavior under prospective reimbursement: Cost sharing and supply. *J Health Econ* 5:129–151. doi: 10.1016/0167-6296(86)90002-0
- Ellis RP, McGuire TG (1993) Supply-side and demand-side cost sharing in health care. *J Econ Perspect J Am Econ Assoc* 7:135–151.
- Elmore JG, Fletcher SW (2012) Overdiagnosis in breast cancer screening: time to tackle an underappreciated harm. *Ann Intern Med* 156:536–537. doi: 10.7326/0003-4819-156-7-201204030-00012
- Falk A, Gächter S, Kovács J (1999) Intrinsic motivation and extrinsic incentives in a repeated game with incomplete contracts. *J Econ Psychol* 20:251–284. doi: 10.1016/S0167-4870(99)00009-4

- Fauvet L, Romain O, Buisine S, Laurent P (2012) Les affectations des étudiants en médecine à l'issue des épreuves classantes nationales en 2011. *Études Résultats Fr Minist Health* 802:1–8.
- Federici A, Giorgi Rossi P, Bartolozzi F, et al (2006) The role of GPs in increasing compliance to colorectal cancer screening: a randomised controlled trial (Italy). *Cancer Causes Control CCC* 17:45–52. doi: 10.1007/s10552-005-0380-9
- Fehr E, Falk A (2002) Psychological foundations of incentives. *Eur Econ Rev* 46:687–724. doi: 10.1016/S0014-2921(01)00208-2
- Feng Y, Farrar S, Sutton M, Ma A (2011) The effect of an increase in the rate of payment on general practitioners' intrinsic and extrinsic motivation. [<http://www.ohe.org/publications/article/the-effect-of-payment-on-gps-motivation-96.cfm>]
- Ferlay J, Shin H-R, Bray F, et al (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer J Int Cancer* 127:2893–2917. doi: 10.1002/ijc.25516
- Fiebig DG, Keane MP, Louviere J, Wasi N (2010) The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity. *Mark Sci* 29:393–421.
- Finn A, Louviere JJ (1992) Determining the appropriate response to evidence of public concern: the case of food safety. *J Public Policy Mark* 12–25.
- Flynn TN, Louviere JJ, Peters TJ, Coast J (2010) Using discrete choice experiments to understand preferences for quality of life. Variance-scale heterogeneity matters. *Soc Sci Med* 70:1957–1965. doi: 10.1016/j.socscimed.2010.03.008
- Flynn TN, Louviere JJ, Peters TJ, Coast J (2007) Best-worst scaling: What it can do for health care research and how to do it. *J Health Econ* 26:171–189. doi: 10.1016/j.jhealeco.2006.04.002
- Flynn TN, Louviere JJ, Peters TJ, Coast J (2008) Estimating preferences for a dermatology consultation using Best-Worst Scaling: comparison of various methods of analysis. *BMC Med Res Methodol* 8:76. doi: 10.1186/1471-2288-8-76
- Fon Sing M, Leuraud K, Duport N (2013) Characteristics of French people using organised colorectal cancer screening. Analysis of the 2010 French Health, Healthcare and Insurance Survey. *Prev Med* 57:65–68. doi: 10.1016/j.ypmed.2013.03.008
- Forsberg E, Axelsson R, Arnetz B (2001) Effects of performance-based reimbursement on the professional autonomy and power of physicians and the quality of care. *Int J Health Plann Manage* 16:297–310. doi: 10.1002/hpm.640
- Forsberg E, Axelsson R, Arnetz B (2002) Performance-based reimbursement in health care: Consequences for physicians' cost awareness and work environment. *Eur J Public Health* 12:44–50. doi: 10.1093/eurpub/12.1.44
- Franc C, Lesur R (2004) Systèmes de rémunération des médecins et incitations à la prévention. *Rev Économique Vol.* 55:901–922. doi: 10.3917/reco.555.0901

- Francois P (2000) "Public service motivation" as an argument for government provision. *J Public Econ* 78:275–299. doi: 10.1016/S0047-2727(00)00075-X
- Freeman GK (2002) Evolving general practice consultation in Britain: issues of length and context. *BMJ* 324:880–882. doi: 10.1136/bmj.324.7342.880
- Frew E, Wolstenholme JL, Whynes DK (2001) Willingness-to-pay for colorectal cancer screening. *Eur J Cancer Oxf Engl* 1990 37:1746–1751.
- Frey BS (1994) How Intrinsic Motivation is Crowded out and in. *Ration Soc* 6:334–352. doi: 10.1177/1043463194006003004
- Frey B. (1997) Not just for the money: an economic theory of personal motivation. Cheltenham, UK and Brookfield, US: Edward Elgar
- Frey BS (1999) *Economics As a Science of Human Behaviour: Towards a New Social Science Paradigm*. Springer
- Frey BS, Jegen R (2001) Motivation Crowding Theory. *J Econ Surv* 15:589–611. doi: 10.1111/1467-6419.00150
- Frey BS, Oberholzer-Gee F (1997) The Cost of Price Incentives: An Empirical Analysis of Motivation Crowding-Out. *Am Econ Rev* 87:746–55.
- Gautier A (2011) Baromètre santé médecins généralistes. INPES, Saint-Denis, p 266
- Gené-Badia J, Escaramis-Babiano G, Sans-Corrales M, et al (2007) Impact of economic incentives on quality of professional life and on end-user satisfaction in primary care. *Health Policy* 80:2–10. doi: 10.1016/j.healthpol.2006.02.008
- Gibbons R (1998) Incentives in Organizations. National Bureau of Economic Research No 6695. [<http://www.nber.org/papers/w6695>]
- Giuffrida A, Gosden T, Forland F, et al (2000) Target payments in primary care: effects on professional practice and health care outcomes. *Cochrane Database Syst Rev* CD000531. doi: 10.1002/14651858.CD000531
- Gneezy U, Rustichini A (2000) Pay Enough or Don't Pay at All. *Q J Econ* 115:791–810. doi: 10.1162/003355300554917
- Goetz K, Campbell SM, Broge B, et al (2012) The impact of intrinsic and extrinsic factors on the job satisfaction of dentists. *Community Dent Oral Epidemiol* 40:474–480. doi: 10.1111/j.1600-0528.2012.00693.x
- Gosden T, Forland F, Kristiansen IS, et al (2001) Impact of payment method on behaviour of primary care physicians: a systematic review. *J Health Serv Res Policy* 6:44–55. doi: 10.1258/1355819011927198
- Grady KE, Lemkau JP, Lee NR, Caddell C (1997a) Enhancing Mammography Referral in Primary Care. *Prev Med* 26:791–800. doi: 10.1006/pmed.1997.0219

- Grady KE, Lemkau JP, Lee NR, Caddell C (1997b) Enhancing Mammography Referral in Primary Care. *Prev Med* 26:791–800. doi: 10.1006/pmed.1997.0219
- Greenacre M (1984) Theory and applications of correspondence analysis. Academic Press, 364p.
- Greene D, Lepper MR (1978) The Hidden Costs of Reward: New Perspectives of the Psychology of Human Motivation. Hillsdale, NJ: L. Erlbaum Associates; New York
- Gregg P, Grout PA, Ratcliffe A, et al (2011) How important is pro-social behaviour in the delivery of public services? *J Public Econ* 95:758–766. doi: 10.1016/j.jpubeco.2011.03.002
- Grimshaw JM, Shirran L, Thomas R, et al (2001) Changing provider behavior: an overview of systematic reviews of interventions. *Med Care* 39:II2–45.
- Grossman, M (1972) On the concept of health capital and the demand for health. *J Polit Econ* 80:223–5.
- Gude T, Vaglum P, Anvik T, et al (2013) A few more minutes make a difference? The relationship between content and length of GP consultations. *Scand J Prim Health Care* 31:31–35. doi: 10.3109/02813432.2012.751698
- Hadley J, Mitchell JM, Sulmasy DP, Bloche MG (1999) Perceived financial incentives, HMO market penetration, and physicians' practice styles and satisfaction. *Health Serv Res* 34:307–321.
- Hall NJ, Rubin GP, Dobson C, et al (2013) Attitudes and beliefs of non-participants in a population-based screening programme for colorectal cancer. *Health Expect Int J Public Particip Health Care Health Policy*. doi: 10.1111/hex.12157
- Hays RD, Morales LS, Reise SP (2000) Item response theory and health outcomes measurement in the 21st century. *Med Care* 38:II28–42.
- Hennig-Schmidt H, Selten R, Wiesen D (2011) How payment systems affect physicians' provision behaviour--an experimental investigation. *J Health Econ* 30:637–646. doi: 10.1016/j.jhealeco.2011.05.001
- Heyes A (2005) The economics of vocation or “why is a badly paid nurse a good nurse”? *J Health Econ* 24:561–569. doi: 10.1016/j.jhealeco.2004.09.002
- Hillman AL, Ripley K, Goldfarb N, et al (1998) Physician financial incentives and feedback: failure to increase cancer screening in Medicaid managed care. *Am J Public Health* 88:1699–1701.
- Hole AR (2006) Small-sample properties of tests for heteroscedasticity in the conditional logit model. *Econ Bull* 3:1–14.
- Houle SKD, McAlister FA, Jackevicius CA, et al (2012) Does performance-based remuneration for individual health care practitioners affect patient care?: a systematic review. *Ann Intern Med* 157:889–899. doi: 10.7326/0003-4819-157-12-201212180-00009

- Howie JG, Heaney DJ, Maxwell M (1997) Measuring quality in general practice. Pilot study of a needs, process and outcome measure. *Occas Pap R Coll Gen Pract* i–32.
- Howie JG, Heaney DJ, Maxwell M, et al (2000) Developing a “consultation quality index” (CQI) for use in general practice. *Fam Pract* 17:455–461.
- Howie JG, Porter AM, Forbes JF (1989) Quality and the use of time in general practice: widening the discussion. *BMJ* 298:1008–1010.
- Howie JG, Porter AM, Heaney DJ, Hopton JL (1991) Long to short consultation ratio: a proxy measure of quality of care for general practice. *Br J Gen Pract* 41:48–54.
- Hsia J, Kemper E, Kiefe C, et al (2000) The Importance of Health Insurance as a Determinant of Cancer Screening: Evidence from the Women’s Health Initiative. *Prev Med* 31:261–270. doi: 10.1006/pmed.2000.0697
- Hu FB, Goldberg J, Hedeker D, et al (1998) Comparison of population-averaged and subject-specific approaches for analyzing repeated binary outcomes. *Am J Epidemiol* 147:694–703.
- Hughes D (1993) General practitioners and the new contract: promoting better health through financial incentives. *Health Policy* 25:39–50. doi: 10.1016/0168-8510(93)90101-T
- Iacobucci D (2010) Structural equations modeling: Fit Indices, sample size, and advanced topics. *J Consum Psychol* 20:90–98. doi: 10.1016/j.jcps.2009.09.003
- James Jr. HS (2005) Why did you do that? An economic examination of the effect of extrinsic compensation on intrinsic motivation and performance. *J Econ Psychol* 26:549–566. doi: 10.1016/j.joep.2004.11.002
- Janus K (2013) The effect of professional culture on intrinsic motivation among physicians in an academic medical center. *J Healthc Manag Am Coll Healthc Exec* 59:287–303; discussion 303–4.
- Janus K (2010) Managing motivation among health care professionals. *Adv Health Care Manag* 9:47–77. doi: 10.1108/S1474-8231(2010)0000009007
- Janus K, Amelung VE, Baker LC, et al (2008) Job Satisfaction and Motivation among Physicians in Academic Medical Centers: Insights from a Cross-National Study. *J Health Polit Policy Law* 33:1133–1167. doi: 10.1215/03616878-2008-035
- Janus K, Amelung VE, Gaitanides M, Schwartz FW (2007) German physicians “on strike”—Shedding light on the roots of physician dissatisfaction. *Health Policy* 82:357–365. doi: 10.1016/j.healthpol.2006.11.003
- Janz NK, Lakhani I, Vijan S, et al (2007) Determinants of colorectal cancer screening use, attempts, and non-use. *Prev Med* 44:452–458. doi: 10.1016/j.ypmed.2006.04.004
- Jensen LF, Mukai TO, Andersen B, Vedsted P (2012) The association between general practitioners’ attitudes towards breast cancer screening and women’s screening participation. *BMC Cancer* 12:254. doi: 10.1186/1471-2407-12-254

- Jha AK (2013) Time to get serious about pay for performance. *JAMA J Am Med Assoc* 309:347–348. doi: 10.1001/jama.2012.196646
- Jöreskog KG (1969) A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika* 34:183–202. doi: 10.1007/BF02289343
- Kaarboe O, Siciliani L (2011) Multi-tasking, quality and pay for performance. *Health Econ* 20:225–238. doi: 10.1002/hec.1582
- Kash KM, Holland JC, Halper MS, Miller DG (1992) Psychological distress and surveillance behaviors of women with a family history of breast cancer. *J Natl Cancer Inst* 84:24–30.
- Katapodi MC, Lee KA, Facione NC, Dodd MJ (2004) Predictors of perceived breast cancer risk and the relation between perceived risk and breast cancer screening: a meta-analytic review. *Prev Med* 38:388–402. doi: 10.1016/j.ypmed.2003.11.012
- Kenkel DS (1994) The demand for preventive medical care. *Appl Econ* 26:313–325. doi: 10.1080/00036849400000078
- Kerlikowske K, Grady D, Rubin SM, et al (1995) Efficacy of screening mammography. A meta-analysis. *JAMA J Am Med Assoc* 273:149–154.
- Kiran T, Wilton AS, Moineddin R, et al (2014) Effect of payment incentives on cancer screening in ontario primary care. *Ann Fam Med* 12:317–323. doi: 10.1370/afm.1664
- Klug SJ, Hetzer M, Blettner M (2005) Screening for breast and cervical cancer in a large German city: participation, motivation and knowledge of risk factors. *Eur J Public Health* 15:70–77. doi: 10.1093/eurpub/cki118
- Kobayashi LC, Wardle J, von Wagner C (2014) Limited health literacy is a barrier to colorectal cancer screening in England: evidence from the English Longitudinal Study of Ageing. *Prev Med* 61:100–105. doi: 10.1016/j.ypmed.2013.11.012
- Kreps D. (1997) Intrinsic motivation and extrinsic incentives. *Am Econ Rev Pap Proc* 87:356–364.
- Krucien N (2012) Analyse de la qualité de l'offre de soins de médecine générale du point de vue des patients. *Paris-Sud* 11, 237 p.
- Krucien N, Le Vaillant M, Pelletier-Fleury N (2014) What are the patients' preferences for the Chronic Care Model? An application to the obstructive sleep apnoea syndrome. *Health Expect Int J Public Particip Health Care Health Policy*. doi: 10.1111/hex.12222
- Laffont J, Martimort D (2002) The theory of incentives: The Principal-Agent model.
- Lairson DR, Chan W, Newmark GR (2005) Determinants of the demand for breast cancer screening among women veterans in the United States. *Soc Sci Med* 1982 61:1608–1617. doi: 10.1016/j.socscimed.2005.03.015
- Lancaster K (1966) A new approach to consumer theory. *J Polit Econ* 134–57.

- Lancsar E, Louviere J, Donaldson C, et al (2013) Best worst discrete choice experiments in health: Methods and an application. *Soc Sci Med* 76:74–82. doi: 10.1016/j.socscimed.2012.10.007
- Lanier DC, Roland M, Burstin H, Knottnerus JA (2003) Doctor performance and public accountability. *The Lancet* 362:1404–1408. doi: 10.1016/S0140-6736(03)14638-7
- Lantz PM, Weigers ME, House JS (1997) Education and income differentials in breast and cervical cancer screening. Policy implications for rural women. *Med Care* 35:219–236.
- Lazear EP, Shaw KL (2007) *Personnel Economics: The Economist’s View of Human Resources*. National Bureau of Economic Research
- Ledford G, Gerhart B, Fang M (2013) Negative effects of extrinsic rewards on intrinsic motivation: more smoke than fire. *World Work J*.
- Lepper MR, Greene D, Nisbett RE (1973) Undermining children’s intrinsic interest with extrinsic reward: A test of the “overjustification” hypothesis. *J Pers Soc Psychol* 28:129–137. doi: 10.1037/h0035519
- Liang K-Y, Zeger SL (1986) Longitudinal data analysis using generalized linear models. *Biometrika* 73:13–22. doi: 10.1093/biomet/73.1.13
- Liberalotto NA (2012) L’engagement des médecins généralistes à l’égard du dépistage des cancers féminins: un révélateur de leurs positionnements face aux transformations de leur contexte d’exercice. *EHESS*, 533 p.
- Li J, Hurley J, DeCicca P, Buckley G (2014) Physician response to pay-for-performance: evidence from a natural experiment. *Health Econ* 23:962–978. doi: 10.1002/hec.2971
- Lindenauer PK, Remus D, Roman S, et al (2007) Public Reporting and Pay for Performance in Hospital Quality Improvement. *N Engl J Med* 356:486–496. doi: 10.1056/NEJMsa064964
- Lindenberg S (2001) Intrinsic motivation in a new light. *Kylos*, p 317–343
- Lippi Bruni M, Nobilio L, Ugolini C (2009) Economic incentives in general practice: The impact of pay-for-participation and pay-for-compliance programs on diabetes care. *Health Policy* 90:140–148. doi: 10.1016/j.healthpol.2008.09.008
- Loewenstein G (1999) Because It Is There: The Challenge of Mountaineering... for Utility Theory. *Kyklos* 52:315–343. doi: 10.1111/j.1467-6435.1999.tb00221.x
- Lofters AK, Moineddin R, Hwang SW, Glazier RH (2010) Low Rates of Cervical Cancer Screening Among Urban Immigrants. *Med Care* 48:611–618. doi: 10.1097/MLR.0b013e3181d6886f
- Lostao L, Joiner TE, Pettit JW, et al (2001) Health beliefs and illness attitudes as predictors of breast cancer screening attendance. *Eur J Public Health* 11:274–279.

- Louviere JJ, Islam T (2008) A comparison of importance weights and willingness-to-pay measures derived from choice-based conjoint, constant sum scales and best–worst scaling. *J Bus Res* 61:903–911. doi: 10.1016/j.jbusres.2006.11.010
- Mangin D, Toop L (2007) The Quality and Outcomes Framework: what have you done to yourselves? *Br J Gen Pract* 57:435–437.
- Mannion R, Davies HTO (2008) Payment for performance in health care. *BMJ* 336:306–308. doi: 10.1136/bmj.39463.454815.94
- March JG (1999) *The Pursuit of Organizational Intelligence: Decisions and Learning in Organizations*. Blackwell Publishers, Inc., Cambridge, MA, USA
- Marley AAJ, Flynn TN, Louviere JJ (2008) Probabilistic models of set-dependent and attribute-level best–worst choice. *J Math Psychol* 52:281–296. doi: 10.1016/j.jmp.2008.02.002
- Marley AAJ, Islam T (2012) Conceptual Relations Between Expanded Rank Data and Models of the Unexpanded Rank Data. *J Choice Model* 5:38–80. doi: 10.1016/S1755-5345(13)70052-6
- Marley AAJ, Louviere JJ (2005) Some probabilistic models of best, worst, and best–worst choices. *J Math Psychol* 49:464–480. doi: 10.1016/j.jmp.2005.05.003
- Martinez-Huedo MA, Lopez de Andres A, Hernandez-Barrera V, et al (2012) Adherence to breast and cervical cancer screening in Spanish women with diabetes: associated factors and trend between 2006 and 2010. *Diabetes Metab* 38:142–148. doi: 10.1016/j.diabet.2011.09.007
- Martín-López R, Jiménez-García R, Lopez-de-Andres A, et al Inequalities in uptake of breast cancer screening in Spain: analysis of a cross-sectional national survey. *Public Health*. doi: 10.1016/j.puhe.2013.03.006
- Maynard A (2012) The powers and pitfalls of payment for performance. *Health Econ* 21:3–12. doi: 10.1002/hecl.1810
- McFadden D (1974) Conditional Logit Analysis of Qualitative Choice Behavior. *Front. Econom.*, P.Zarembka. Academic Press: New York, pp 105–142
- Mckalip D (2009) Pay for Performance and Public Reporting : Risks to Patients Outweigh Benefits. *J Am Physicians Surg* 113–117.
- McLintock K, Russell AM, Alderson SL, et al (2014) The effects of financial incentives for case finding for depression in patients with diabetes and coronary heart disease: interrupted time series analysis. *BMJ Open* 4:e005178. doi: 10.1136/bmjopen-2014-005178
- Mercer SW, Howie JGR (2006) CQI-2--a new measure of holistic interpersonal care in primary care consultations. *Br J Gen Pract J R Coll Gen Pract* 56:262–268.

- Mills B, Reyna VF, Estrada S (2008) Explaining contradictory relations between risk perception and risk taking. *Psychol Sci* 19:429–433. doi: 10.1111/j.1467-9280.2008.02104.x
- Morgan MAJ, Dunbar J (2011) Implementing pay-for-performance in Australian primary care: lessons from the United Kingdom and the United States. *Med J Aust* 194:213–214.
- Moumjid N, Carrère M-O, Charavel M, Brémond A (2003) Clinical issues in shared decision-making applied to breast cancer. *Health Expect Int J Public Particip Health Care Health Policy* 6:222–227.
- Mousquès J (2011) Le regroupement des professionnels de santé de premiers recours : quelles perspectives économiques en termes de performance ? *Rev Fr Aff Soc* 2-3:253–275.
- Mühlbacher AC, Bethge S (2014) Patients' preferences: a discrete-choice experiment for treatment of non-small-cell lung cancer. *Eur J Health Econ HEPAC Health Econ Prev Care*. doi: 10.1007/s10198-014-0622-4
- Mulaik SA (1988) Confirmatory Factor Analysis. In: Nesselroade JR, Cattell RB (eds) *Handb. Multivar. Exp. Psychol.* Springer US, pp 259–288
- Mullen KJ, Frank RG, Rosenthal MB (2010) Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. *RAND J Econ* 41:64–91. doi: 10.1111/j.1756-2171.2009.00090.x
- Mundlak Y (1978) On the Pooling of Time Series and Cross Section Data. *Econometrica* 46:69. doi: 10.2307/1913646
- Muthén B (1978) Contributions to factor analysis of dichotomous variables. *Psychometrika* 43:551–560. doi: 10.1007/BF02293813
- Muthén O, du Toit S, Spisic D (1997) Robust inference using weighted least squares and quadratic estimating equations in latent variable modelling with categorical and continuous outcomes. Tech. Rep.
- Narcy M (2011) Would nonprofit workers accept to earn less? Evidence from France. *Appl Econ* 43:313–326. doi: 10.1080/00036840802570447
- Nardo M, Saisana M, Saltelli A, et al (2005) Handbook on constructing composite indicators: methodology and user guide. OECD Stat. Work. Pap.
- Nguyen F, Moumjid N, Brémond A, Carrère M-O (2008) Validité théorique de la Méthode des Choix Discrets : le cas du Traitement Hormonal substitutif de la Ménopause. *J Gest Déconomie Médicales* 26:259–268.
- Nguyen F, Moumjid N, Charles C, et al (2014) Treatment decision-making in the medical encounter: comparing the attitudes of French surgeons and their patients in breast cancer care. *Patient Educ Couns* 94:230–237. doi: 10.1016/j.pec.2013.07.011
- OCDE (2013) Panorama de la santé 2013: les indicateurs de l'OCDE. [<http://www.oecd.org/fr/els/systemes-sante/Panorama-de-la-sante-2013.pdf>]

- ONDPS (2007) Synthèse générale de l'observatoire de la démographie des professionnels de santé. [http://www.sante.gouv.fr/IMG/pdf/Synthese_generale_ONDPS.pdf]
- Organization WH (2008) The Global Burden of Disease: 2004 Update. World Health Organization [http://www.who.int/healthinfo/global_burden_disease/GBD_report_2004update_full.pdf]
- Osterloh M, Frey BS (2000) Motivation, Knowledge Transfer, and Organizational Forms. *Organ Sci* 11:538–550. doi: 10.1287/orsc.11.5.538.15204
- Osterloh M, Frost J, Frey BS (2002) The Dynamics of Motivation in New Organizational Forms. *Int J Econ Bus* 9:61–77. doi: 10.1080/13571510110102976
- Osterloth M, Frey B. (2000) Motivation, knowledge transfer, and organizational form. *Organ Sci* 11:538–550.
- Petersen LA, Woodard LD, Urech T, et al (2006) Does Pay-for-Performance Improve the Quality of Health Care? *Ann Intern Med* 145:265–272. doi: 10.7326/0003-4819-145-4-200608150-00006
- Protière C, Moumjid N, Bouhnik A-D, et al (2012) Heterogeneity of cancer patient information-seeking behaviors. *Med Decis Mak Int J Soc Med Decis Mak* 32:362–375. doi: 10.1177/0272989X11415114
- Quinn M, Babb P, Jones J, Allen E (1999) Effect of screening on incidence of and mortality from cancer of cervix in England: evaluation based on routinely collected statistics. *BMJ* 318:904–908.
- RACGP (2011) RACGP Chronic conditions self management guidelines. Guidelines summary for GP's. [<http://curriculum.racgp.org.au/media/12314/chronicconditions.pdf>]
- Rice N, Jones A (1997) Multilevel models and health economics. *Health Econ* 6:561–575.
- Rigal L, Saurel-Cubizolles M-J, Falcoff H, et al (2011) Do social inequalities in cervical cancer screening persist among patients who use primary care? The Paris Prevention in General Practice survey. *Prev Med* 53:199–202. doi: 10.1016/j.ypmed.2011.06.016
- Rizzo JA, Blumenthal D (1994) Physician labor supply: do income effects matter? *J Health Econ* 13:433–453.
- Robinson JC (2001) Theory and Practice in the Design of Physician Payment Incentives. *Milbank Q* 79:149–177. doi: 10.1111/1468-0009.00202
- Rochaix L (1989) Information asymmetry and search in the market for physicians' services. *J Health Econ* 8:53–84. doi: 10.1016/0167-6296(89)90009-X
- Rochaix L (2004) Les modes de rémunération des médecins. *Rev Déconomie Financ* 76:223–239.

- Rosenthal MB, Dudley R (2007) Pay-for-performance: Will the latest payment trend improve care? *JAMA* 297:740–744. doi: 10.1001/jama.297.7.740
- Rosenthal MB, Frank RG (2006) What Is the Empirical Basis for Paying for Quality in Health Care? *Med Care Res Rev* 63:135–157. doi: 10.1177/1077558705285291
- Ryan AM (2010) Has Pay-for-Performance Decreased Access for Minority Patients? *Health Serv Res* 45:6–23. doi: 10.1111/j.1475-6773.2009.01050.x
- Ryan AM, Doran T (2012) The effect of improving processes of care on patient outcomes: evidence from the United Kingdom’s quality and outcomes framework. *Med Care* 50:191–199. doi: 10.1097/MLR.0b013e318244e6b5
- Ryan M, Gerard K (2003) Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Appl Health Econ Health Policy* 2:55–64.
- Sabatino SA, Habarta N, Baron RC, et al (2008) Interventions to increase recommendation and delivery of screening for breast, cervical, and colorectal cancers by healthcare providers systematic reviews of provider assessment and feedback and provider incentives. *Am J Prev Med* 35:S67–74. doi: 10.1016/j.amepre.2008.04.008
- Saint-Lary O, Bernard E, Sicsic J, et al (2013) Why Did Most French GPs Choose Not to Join the Voluntary National Pay-for-Performance Program? *PLoS ONE* 8:e72684. doi: 10.1371/journal.pone.0072684
- Saint-Lary O, Plu I, Naiditch M (2011) Adhérer ou pas au CAPI: de quel clivage des généralistes le paiement à la performance est-il le révélateur? *Rev Fr Aff Soc* n° 2-3:180–209.
- Samson A-L (2009) Faut-il remettre en cause le paiement à l’acte des médecins? *Regards Croisés Sur Léconomie* 5:144–158. doi: 10.3917/rce.005.0144
- Schnedler W, Vanberg C (2014) Playing “hard to get”: An economic rationale for crowding out of intrinsically motivated behavior. *Eur Econ Rev* 68:106–115. doi: 10.1016/j.eurocorev.2014.02.006
- Schwartz B (1990) The creation and destruction of value. *Am Psychol* 45:7–15. doi: 10.1037/0003-066X.45.1.7
- Scott A (2001) Eliciting GPs’ preferences for pecuniary and non-pecuniary job characteristics. *J Health Econ* 20:329–347.
- Scott A, Sivey P, Ait Ouakrim D, et al (1996) The effect of financial incentives on the quality of health care provided by primary care physicians. *Cochrane Database Syst. Rev.*
- Sébillé V, Hardouin J-B, Le Néel T, et al (2010) Methodological issues regarding power of classical test theory (CTT) and item response theory (IRT)-based approaches for the comparison of patient-reported outcomes in two groups of patients--a simulation study. *BMC Med Res Methodol* 10:24. doi: 10.1186/1471-2288-10-24

- Selvin E, Brett KM (2003) Breast and cervical cancer screening: sociodemographic predictors among White, Black, and Hispanic women. *Am J Public Health* 93:618–623.
- Serumaga B, Ross-Degnan D, Avery AJ, et al (2011) Effect of pay for performance on the management and outcomes of hypertension in the United Kingdom: interrupted time series study. *BMJ*. doi: 10.1136/bmj.d108
- Siciliani L (2009) Paying for performance and motivation crowding out. *Econ Lett* 103:68–71. doi: 10.1016/j.econlet.2009.01.022
- Simon HA (1991) Organizations and Markets. *J Econ Perspect* 5:25–44. doi: 10.1257/jep.5.2.25
- Simou E, Foundoulakis E, Kourlaba G, Maniadakis N (2011) Factors associated with the use of preventive services by women in Greece. *Eur J Public Health* 21:512–519. doi: 10.1093/eurpub/ckq103
- Sliwka D (2006) Trust as a Signal of a Social Norm and the Hidden Costs of Incentive Schemes. Social Science Research Network, Rochester, NY
- Smith R (2001) Why are doctors so unhappy? *BMJ* 322:1073–1074. doi: 10.1136/bmj.322.7294.1073
- Snijders T, Bosker R (1999) Multilevel analysis: An introduction to basic and advanced multilevel modeling. Sage Publ. Ltd, 368 p.
- Stiratelli R, Laird N, Ware JH (1984) Random-Effects Models for Serial Observations with Binary Response. *Biometrics* 40:961. doi: 10.2307/2531147
- Street DJ, Burgess L (2007) The Construction of Optimal Stated Choice Experiments: Theory and Methods. John Wiley & Sons
- Szeinbach SL, Barnes JH, McGhan WF, et al (1999) Using Conjoint Analysis to Evaluate Health State Preferences*. *Drug Inf J* 33:849–858. doi: 10.1177/009286159903300326
- Tacken MAJB, Braspenning JCC, Hermens RPMG, et al (2007) Uptake of cervical cancer screening in The Netherlands is mainly influenced by women's beliefs about the screening and by the inviting organization. *Eur J Public Health* 17:178–185. doi: 10.1093/eurpub/ckl082
- Tang S-H, Hall VC (1995) The overjustification effect: A meta-analysis. *Appl Cogn Psychol* 9:365–404. doi: 10.1002/acp.2350090502
- Thurstone LL (1927) A law of comparative judgment. *Psychol Rev* 273–286.
- Town R, Kane R, Johnson P, Butler M (2005a) Economic incentives and physicians' delivery of preventive care: A systematic review. *Am J Prev Med* 28:234–240. doi: 10.1016/j.amepre.2004.10.013

- Town R, Kane R, Johnson P, Butler M (2005b) Economic incentives and physicians' delivery of preventive care: A systematic review. *Am J Prev Med* 28:234–240. doi: 10.1016/j.amepre.2004.10.013
- Videau Y (2010) Essais sur l'offre de travail en médecine générale. Du rôle des incitations et des motivations. Aix Marseille 2, 280 p.
- Videau Y, Batifoulier P, Arrighi Y, et al (2010) Le cycle de vie des motivations professionnelles en médecine générale : une étude dans le champ de la prévention. *Rev DÉpidémiologie Santé Publique* 58:301–311. doi: 10.1016/j.respe.2010.04.009
- Walsh B, Silles M, O'Neill C (2011) The importance of socio-economic variables in cancer screening participation: A comparison between population-based and opportunistic screening in the EU-15. *Health Policy* 101:269–276. doi: 10.1016/j.healthpol.2011.02.001
- Walter LC, Lewis CL, Barton MB (2005) Screening for colorectal, breast, and cervical cancer in the elderly: a review of the evidence. *Am J Med* 118:1078–1086. doi: 10.1016/j.amjmed.2005.01.063
- Wang AT, Fan J, Van Houten HK, et al (2014) Impact of the 2009 US Preventive Services Task Force guidelines on screening mammography rates on women in their 40s. *PloS One* 9:e91399. doi: 10.1371/journal.pone.0091399
- Watt T, Groenvold M, Deng N, et al (2014) Confirmatory factor analysis of the thyroid-related quality of life questionnaire ThyPRO. *Health Qual Life Outcomes* 12:126. doi: 10.1186/s12955-014-0126-z
- Whynes DK, Baines DL (1998) Income-based incentives in UK general practice. *Health Policy Amst Neth* 43:15–31.
- Wilson A (1991) Consultation length in general practice: a review. *Br J Gen Pract* 41:119–122.
- Wilson A, Childs S (2002) The relationship between consultation length, process and outcomes in general practice: a systematic review. *Br J Gen Pract* 52:1012–1020.
- Wong GY, Mason WM (1985) The Hierarchical Logistic Regression Model for Multilevel Analysis. *J Am Stat Assoc* 80:513. doi: 10.2307/2288464
- Woolhandler S, Ariely D, Himmelstein DU (2012) Why pay for performance may be incompatible with quality improvement. *BMJ* 345:e5015.
- Wordsworth S, Ryan M, Skåtun D, Waugh N (2006) Women's preferences for cervical cancer screening: a study using a discrete choice experiment. *Int J Technol Assess Health Care* 22:344–350.