UNIVERSITÉ PARIS XIII – SORBONNE PARIS NORD

École doctorale Sciences, Technologies, Santé Galilée

Titre de la thèse en français

Mieux atteindre, cibler et caractériser les patients relevant du dépistage organisé du cancer colorectal (DOCCR) en France: Place des bases médico-administratives et des réseaux sociaux

Title of the thesis in english (optional)

Better reach, target and define the patients concerned by French population-based colorectal cancer screening program: Place of medico-administrative databases and social networks.

THÈSE DE DOCTORAT
présentée par

Akoï KOÏVOGUI

Laboratoire d'Informatique Médicale et d'ingénierie des connaissances en e-Santé (LIMICS)

pour l'obtention du grade de
DOCTEUR EN < INFORMATIQUE MEDICALE >

Soutenue le < 21/12/2023 > devant le jury d'examen constitué de :

BOUZILLE Guillaume, MCU-PH, HDR, Université de Rennes 1, Rapporteur
APARICIO Thomas, PU-PH, Université Paris Cité, Rapporteur
JAY Nicolas, PU-PH, Université de Lorraine, Examinateur
SABATE Jean Marc, PU-PH, Université Sorbonne Paris Nord, Examinateur
DUCLOS Catherine, PU-PH, Université Sorbonne Paris Nord, Co-Directrice de thèse
BENAMOUZIG Robert, PU-PH, Université Sorbonne Paris Nord, Co-Directrice de thèse

RESUME DE LA THESE EN FRANÇAIS

Toutes les conditions justifiant la mise en place d'une politique de dépistage sont remplies par le cancer colorectal (CCR) : il est fréquent (43 000 nouveaux cas en France chaque année), grave (on ne guérit actuellement qu'un cancer sur deux) et s'il est diagnostiqué au début, il est habituellement guéri. Ce cancer peut être prévenu par la détection et l'exérèse des adénomes qui le précèdent le plus souvent. Enfin, il existe une stratégie de dépistage reposant sur la recherche d'un saignement occulte dans les selles par un test réalisé tous les 2 ans. En France, le taux de participation aux campagnes de dépistage atteint rarement 50% dans la plupart des départements français. Nous cherchions à vérifier deux hypothèses : i)l'exploitation de plusieurs sources de données, notamment les bases de données médico-administratives et les registres de cancers, permettrait de mieux cibler et caractériser les personnes éligibles aux campagnes de dépistage en France, ce qui améliorerait considérablement le taux d'exclusion médicale ; ii) la promotion du dépistage à travers les réseaux sociaux permettrait d'atteindre les personnes non participant aux campagnes de dépistage, ce qui améliorerait considérable l'effectif de participants à une campagne de dépistage.

Dans cette thèse, nous nous sommes fixés pour objectif de développer des approches pour mieux cibler, caractériser et atteindre les patients relevant du dépistage en France, en s'intéressant en particulier à l'apport de la réutilisation de données médico-administratives et l'usage de réseaux sociaux.

Notre approche méthodologique a été de comparer d'abord les définitions internationales des populations cible du dépistage et la place des bases de données médico-administratives dans les méthodologies de ciblage des populations éligibles aux campagnes de dépistage à travers le monde. Une synthèse des différents profils de personnes atteintes d'une situations morbides pouvant justifier d'une exclusion médicale a été faite. La définition de la personne a été faite à l'aide de critères d'exclusion consensuels dans les programmes internationaux dans le de la transcrire en requête pour pouvoir exclure les personnes du dépistage à partir des données du Système national des données de santé (SNDS). A l'aide d'un module de géolocalisation et de sensibilisation sur le réseau social Facebook, nous avons analyser l'apport des réseaux sociaux comme sources supplémentaires de motivation des personnes à participer aux campagnes de DOCCR.

Cette thèse met en évidence des stratégies et moyens d'exploitation de plusieurs sources de données pour une optimisation des campagnes de DOCCR en France. La thèse montre aussi que ces sources de données n'étaient pas encore exploitées en France dans le cadre du dépistage du cancer colorectal. Pour consolider la définition de la population éligible au DOCCR, l'intérêt du développement d'une méthode d'identification à partir des données de consommation de soins du SNDS a été largement argumenté dans cette thèse. Nous avons argumenté l'accès aux données du SNDS et le projet suit son cours en guise de perspective à la thèse. Nous envisageons aussi, la mise en place et l'évaluation à grande échelle d'une stratégie complémentaire d'optimisation des campagnes de DOCCR à travers les réseaux sociaux notamment Facebook car l'expérimentation n'avait ciblé six communes d'un seul département.

Mots Clés : Dépistage du cancer colorectal ; Taux de participation, Taux d'exclusion médicale ; Base de données médico-administrative ; Registre de cancer ; Réseaux sociaux.

RESUME DE LA THESE EN ANGLAIS

All the conditions justifying the implementation of a screening policy are met by colorectal cancer (CRC): it is frequent (43,000 new cases in France each year), serious (only one in two CRC case is currently cured) and if diagnosed early, it is usually cured. This cancer can be prevented by the detection and excision of the adenomas which most often precede it. Finally, there is a screening strategy based on the search for occult bleeding in the stools by a test carried out every 2 years. In France, the participation rate in screening campaigns rarely reaches 50% in most French districts. We sought to verify two hypotheses in this research work: i) The use of several data sources, in particular medico-administrative databases and cancer registries, would make it possible to better target and define people eligible for screening campaigns in France, which would improve significantly the medical exclusion rates; ii) Promoting screening through social networks would make it possible to reach people who do not participate in screening campaigns, which would considerably improve the number of participants in a screening campaign.

In this thesis, we aimed to develop approaches to better target, define and reach patients subject to screening in France, focusing in particular on the contribution of the reuse of medico-administrative data and the use of social networks.

Our methodological approach was to first compare international definitions of screening target populations and the place of medico-administrative databases in methodologies for targeting populations eligible for screening campaigns around the world. A summary of the different profiles of people suffering from morbid situations that could justify medical exclusion was highlighted. The definition of the person was made using consensual exclusion criteria in international programs in order to translate it into a query to be able to exclude people from screening campaign using the data of the French medico-administrative database usually called “SNDS” (“Système national des données de santé”). Using a geolocation and awareness module on the social network Facebook, we analyzed the contribution of social networks as additional sources of motivation for people to participate in colorectal cancer screening campaigns.

This thesis highlights strategies and means for using several data sources for optimization of screening campaigns in France. The thesis also shows that these data sources were not yet used in France for colorectal cancer screening. To consolidate the definition of the campaign’s eligible population, the interest in developing an identification method based on healthcare consumption data from SNDS has been widely argued in this thesis. We have argued for access to SNDS data, and the project is continuing as a perspective for the thesis. We are also considering the implementation and large-scale evaluation of a complementary strategy for optimizing screening campaigns through social networks, particularly Facebook, because the experiment only targeted six municipalities in a single district.

Keywords: Colorectal cancer screening; Population-based screening program; Participation rate, medical exclusion rate; Medical-administrative database; Cancer Registry; Social networks.

REMERCIEMENTS

SOMMAIRE

RESUME DE LA THESE EN FRANÇAIS **2**

RESUME DE LA THESE EN ANGLAIS **3**

REMERCIEMENTS **4**

LISTE DES ABREVIATIONS **8**

LISTE DES TABLEAUX **10**

LISTE DES FIGURES **11**

VALORISATION SCIENTIFIQUE **12**

1. INTRODUCTION **16**

1.1.Le programme dépistage organisé du cancer colorectal en France (DOCCR) 16

1.1.1 – INTERET de la mise en place du programme dépistage organisé du cancer colorectal (DOCCR) 16

1.1.2 – ORGANISATION du DOCCR en France 17

1.1.3 – Problématique de la faible participation aux campagnes de DOCCR 19

1.1.4 – Stratégies et moyens actuels d'augmentation de la participation aux campagnes 20

1.1.5 – Fiabilité des données d'invitation de la population cible de DOCCR 20

Figure 1.1.5.1 : Causes de non-participation à la campagne de DOCCR 2015-2016 en Ile-de-France 22

1.1.6 – Fiabilité de la méthode d'estimation du taux de participation 23

Figure 1.1.6.1: Taux d'exclusion à la campagne de DOCCR 2015-2016 en Ile-de-France 25

1.2. Sources de données exploitées et exploitables pour une meilleure détermination du taux d'exclusion 26

1.2.1 – Revue sur les Bases de données medico-administratives et registres de cancers le depistage 26

1.2.2 – Les Bases de données exploitables par le programme de depistage en France 27

1.2.2.A. – Le Système National des Données de Santé (SNDS) 27

Figure 1.2.2.1 Sources d'alimentation du SNDS. Source : Health-Data-Hub [129] 28

Figure 1.2.2.2 : Schéma relationnel de la base de données DCIR. Source Figure Health-Data-Hub (HDH)[129] 29

1.2.2.B. – Les Registres de cancers en France 30

Figure 1.2.2.3 : Cartographie des registres de cancers. 31

1.3. Hypothèse 32

2. OBJECTIFS DE LA THESE ET AXE DE RECHERCHE **32**

2.1. – Objectif principal 32

2.2. – Objectifs spécifiques 32

3. AXE 1 : ANALYSE DE LA DEFINITION ACTUELLE DE LA POPULATION CIBLE DES CAMPAGNES ET PLACE DES BASES DE DONNEES ADMINISTRATIVES DANS LES METHODOLOGIES DE CIBLAGE **33**

3.1 Place des bases de données médico-administratives dans la sélection des personnes éligibles aux campagnes de dépistage : Étude 1. 33

3.1.1 – Contexte 33

3.1.2 – Méthodes 34

3.2 Performance des algorithmes de recherche dans les bases de données médico-administratives des facteurs morbides justifiant d'une exclusion médicale du dépistage organisé, Revue de la littérature: Étude 2. 71

3.2.1 – Contexte 71

3.2.2 – Méthodes 71

3.2.3 – Resultats 72

Tableau-3.2.3.1: Exemples d'études de validation des phénotypes de CCR, coloscopie et IBD dans des BDMA. 73

3.3 Synthèse des résultats sur la définition de la personne cible. 75

4. AXE 2 : DEFINIR UN PATIENT A EXCLURE DU DEPISTAGE A PARTIR DES DONNEES DU SNDS **76**

4.1 Pratiques actuelles des structures de gestion du dépistage en France: Étude 3. 76

4.1.1 – Contexte 76

4.1.2 – Méthodes 76

4.2 Filtrage de la population cible du dépistage du CCR en France, État de lieux en prélude de l'usage de la base médico-administrative de l'assurance maladie: Étude 4. 103
4.2.1 – Contexte 103
4.2.2 – Méthodes 103

Figure-4.2.1 : Principales stratégies utilisées pour l'exclusion médicale au cours d'une année (N) de campagne de DOCCR. 104

4.3 Algorithme d'exclusion médicale et faisabilité de l'exclusion des personnes non éligibles au dépistage à partir du SNDS : Étude 5. 108
4.3.1 – Contexte 108
4.3.2 – Méthodes 108

4.4 Synthèse de la définition opérationnelle de la personne à exclure du dépistage à partir du SNDS. 132

Tableau 4.4.1 : Analyse d'impact des scénarios d'accès au dépistage du CRC en France. 133

5. AXE 3 : MONTRER QUE LES RESEAUX SOCIAUX PEUVENT ETRE UTILISES COMME SOURCES SUPPLEMENTAIRES DE MOTIVATION DES PERSONNES A PARTICIPER AU DOCCR 135

5.1 – Faisabilité d'une incitation à la participation aux campagnes de dépistage du cancer colorectal, par la motivation des personnes à travers les réseaux sociaux: Étude 6. 135
5.1.1 – Contexte 135
5.1.2 – Méthodes 135
5.1.3 – Synthèse des résultats 136

5.2. – Synthèses de l'axe 3 et perspectives 149

6. DISCUSSION GENERALE 150

6.1. Principaux résultats 150

6.2. Apport des sources de données pour une meilleure définition de la population cible du dépistage 151

7. CONCLUSION 153

8. PERSPECTIVES DE LA THESE 154

8.1 – Validation de l'algorithme d'extraction des données du SNDS 154
8.1.1 – Test de l'éfficacité de l'alogrithme d'exclusion médicale en termes d'extraction de données des bases du SNDS 154
8.1.1.A. – Phénotyper la personne cible à exclure des campagnes à l'aide du SNDS 154
8.1.1.B. – Analyser l'apport des données du SNDS pour une estimation du taux de participation à une campagne de DOCCR en France 155
8.1.2. – Projet de chainage des données du programme de dépistage du cancer colorectal et des données du SNDS. 155

8.2 Projet de mise en œuvre d'une stratégie complémentaire basée sur la géolocalisation et la motivation des personnes par le biais de leur connexion sur les réseaux sociaux 157

Annexe 1 : Analyse du flux et projections faites par le module. 159

Annexe 2 : Projet de phénotypage de la personne cible à exclure ou pas des campagnes de dépistage organisé du cancer colorectal à l'aide des données du SNDS. 160

9. BIBLIOGRAPHIE 172

LISTE DES ABREVIATIONS

ALD	Affection longue durée
APHP	Assistance publique hôpitaux de paris
ARS	Agence régionale de santé
BDMA	Base de données médico-administratives
CCAM	Classification commune des actes médicaux
CCR	Cancer colorectal
CEPIDC	Centre d'épidémiologie sur les causes médicales de décès
CESREES	Comité éthique et scientifique pour les recherches, les études et les évaluations dans le domaine de la santé
CHU	Centre hospitalier universitaire
CIM-X	Classification internationale des maladies et des problèmes de santé connexes $X^{ème}$ révision
CNAM	Caisse nationale de l'assurance maladie
CNIL	Commission nationale de l'informatique et des libertés
CNRS	Centre national de la recherche scientifique
CPAM	Caisse primaire d'assurance maladie
CRCDC	Centre régional de coordination du dépistage des cancers
DCIR	Données de consommation inter régime
DOCCR	Programme de dépistage organisé du cancer colorectal
DRG	Diagnosis Related Group
FIT	Test fécal immunologique
GHM	Groupe homogène de malades
GHS	Groupe homogène de séjours
HCL	Hospices civils de Lyon
HDH	Health Data Hub
HMO	Health Maintenance Organization
HAD	Hospitalisation à domicile
IDF	Ile-de-France
IMC	Indice de masse corporelle
Inca	Institut national du cancer

INSERM	Institut national de la santé et de la recherche médicale
LIMICS	Laboratoire d'informatique médicale et d'ingénierie des connaissances en e-Santé
MCO	Médecine chirurgie obstétrique et odontologie
MICI	Maladies inflammatoires chroniques de l'intestin
MSA	Mutualité sociale agricole
NIR	Numéro d'inscription au répertoire des personnes physiques
OMS	Organisation mondiale de la santé
PMSI	Programme de médicalisation des systèmes d'information
RG	Régime général de l'assurance maladie
RSI	Régime social des indépendants
SPF	Santé publique France
SNDS	Système national des données de santé
SNIIRAM	Système national d'information inter-régimes de l'assurance maladie
UE	Union Européenne
USPN	Université sorbonne Paris nord

LISTE DES TABLEAUX

Tableau 5.4.1 : Analyse d'impact des scénarios d'accès au dépistage du CRC. …….. 77

LISTE DES FIGURES

Figure-4.2.1 : Principales stratégies utilisées pour l'exclusion médicale au cours d'une année (N) de campagne de DOCCR. ……47

VALORISATION SCIENTIFIQUE

I - Articles dans les revues scientifiques à comité de lecture

I-1 En rapport avec la thèse:

a. Articles publiés:

Koïvogui A, *Benamouzig R, Duclos C. Refinement of the Target Population for Colorectal Cancer Screening in France, Inventory as a Prelude to the Use of Medico-Administrative Database.* ***Stud Health Technol Inform.*** *2023 May 18;302: 394-395. doi: 10.3233/SHTI230157.PMID: 37203702*

Koïvogui A*, Levi S, Finkler M, Lewkowicz S, Gombeaud T, Sabate JM, Duclos C, Benamouzig R. Feasibility of encouraging participation in colorectal cancer screening campaigns by motivating people through the social network, Facebook.* ***Colorectal Dis.*** *2020 May 12. doi: 10.1111/codi.15121. PMID: 32397003*

b. Articles soumis:

Koïvogui A*, Benamouzig R, Balamou C, Binefa G, Hoeck S, Duclos C. Place of medico-administrative databases in the selection of the target population in population-based colorectal cancer screening programs.* ***Nature Medicine****, Manuscript ID: NMED-A129223, Submitted October 30, 2023.*

Koïvogui A*, Balamou C, Benamouzig R, Duclos C. Enquête transversale sur les pratiques d'exclusion médicale lors des campagnes de dépistage du cancer colorectal en France.* ***Rev Epidemiol Sante Publique.*** *Manuscript ID: RESPE-D-23-00180, Submitted October 28, 2023*

Koïvogui A*, Balamou C, Benamouzig R, Duclos C. Computational definition of medical exclusion and feasibility of excluding people not eligible for French population-based colorectal cancer screening from the French medico-administrative database.* ***BMC Med Inform Decis Mak,*** *Manuscript ID: d4d8b993-b1ec-4914-86f7-4021812b6bc4, Submitted July 11, 2023.*

I-2 Sans rapport avec la thèse:

a. Articles publiés:

Koïvogui A*, Abihsera G, Le Trong T, Ait- Hadad H, Bernoux A, Delattre H, Vincelet C, Caroll C, Nicolet J. Evaluation of the severity of screened colorectal cancer in the context of the health crisis linked to COVID19 in Ile-de-France region.* ***Rev Epidemiol Sante Publique****, 2023 Jun 26;71(5):102124. doi: 10.1016/j.respe.2023.102124.PMID: 37451076*

Uzan J, ***Koivogui A****, Benamouzig A. Dépistage du cancer colorectal.* ***EMC Gastro-entérologie*** *[9-008-E-50] - doi : 10.1016/S1155-1968(23)46338-9*

Koïvogui A*, Vincelet C, Abihsera G, Ait- Hadad H, Delattre H, Le Trong T, Bernoux A, Caroll C, Nicolet J. Supply and quality of colonoscopy according to the characteristics of gastroenterologists in the French population-based colorectal-cancer screening program.* ***World J Gastroenterol.*** *2023 Mar 7;29(9):1492-1508. doi: 10.3748/wjg.v29.i9.1492.PMID: 36998423.*

*Guilavogui T, Verdun S, **Koïvogui A**, Viscogliosi E, Certad G. Prevalence of Intestinal Parasitosis in Guinea: Systematic Review of the Literature and Meta-Analysis. **Pathogens**. 2023 Feb 16;12(2):336. doi: 10.3390/pathogens12020336.PMID: 36839608*

*Balamou C, **Koïvogui A,** Zysman K, Rodrigue CM, Rymzhanova R. Impact of tomosynthesis on the evolution of the cancer detection rate in the French National Breast Cancer Screening Program. **Public Health** 2022 Sep ;210:65-73. doi: 10.1016/j.puhe.2022.06.011. Epub 2022 Jul 20. PMID: 35870323*

*Balamou C, **Koïvogui A**, Rymzhanova R, Cornelis S, Rodrigue-Moulinie C, Sellier N. Breast Cancer Incidence by Age at Finding of Mammographic Abnormality in Women Participating in French Organized Screening Campaigns. **Public Health.** 2022 Jan;202:121-130. doi: 10.1016/j.puhe.2021.11.012. Epub 2021 Dec 21. PMID: 34952431*

*Vincelet C, Forzy ML, Bernoux A, **Koivogui A**. Le dépistage du cancer du sein dans deux régions françaises dans le contexte de la crise sanitaire liée à la COVID19. **Imagérie de la femme** 2021 Oct; 31(3): 130–135. Doi : 10.1016/j.femme.2021.05.004. PMCID: PMC8429352*

*Balamou C, **Koïvogui A**, Rodrigue-Moulinie C, Clerc A, Piccotti C, Deloraine A, Exbrayat C. Prediction of the severity of colorectal lesion by fecal hemoglobin concentration observed during previous test in the French screening program. **World J Gastroenterol.** 2021 Aug 21;27(31):5272-5287. doi: 10.3748/wjg.v27.i31.5272. PMID: 34497450*

***Koïvogui A**, Vincelet C, Ait-Hadad H, Pellissier H, Valibay S, Kaufmanis A, Benamouzig R. [Colorectal cancer screening program: Test completion rate and follow-up results after selective mailing of the test kit, based on likelihood of participation]. **Rev Epidemiol Sante Publique**. 2021 Oct;69(5):265-276. doi: 10.1016/j.respe.2021.06.002. PMID: 34344563*

*Deutsch D, **Koivogui A**, Benamouzig R. Dépistage du cancer colorectal: un taux de participation qui stagne ? **la revue du praticien medecine générale** - tome 34 - n°1037 - mars 2020*

b. **Articles en cours de publication**:

II – Communication aux congrès

II-1 En rapport avec la thèse :

Koïvogui A. Analyse comparative des stratégies et moyens d'exclusion médicale dans les programmes de dépistage organisé du cancer colorectal des pays membres de l'Union européenne. *Journées Francophones d'Hépato-gastroentérologie et d'Oncologie Digestive, 16-19 mars 2023,* 6ème forum du dépistage organisé du cancer colorectal 17 mars 2023, Communication orale. Programme accessible : https://www.jfhod.com/sites/www.jfhod.com/files/2023-01/JFHOD23_programme_sp.pdf

Koivogui A**, Levi S., Lewkowicz S, Gombeaud T, Le Mab G, Benamouzig R. People´s motivation through social networks could increase participation in colorectal cancer screening campaigns. Congès Barcelone, **UEG (United European Gastroenterology)** Abstract book Oct 2019, Poster P1843. **United European

Gastroenterology Journal *Volume 7 Issue 8 (Supplement), October 2019. doi: 10.1177/2050640619854671 PMID: 32213001*

A. Koivogui*, S. Levi, S. Lewkowicz, T. Gombeaud, J.M. Sabaté, R. Benamouzig. Faisabilité d'une incitation à la participation aux campagnes de dépistage du cancer colorectal, par la motivation des personnes à travers le réseau social Facebook. Journées Francophones d'Hépato-gastroentérologie et d'Oncologie Digestive, mars 2020, Poster P.144. Accessible: https://www.snfge.org/content/faisabilite-dune-incitation-la*

II-2 Sans rapport avec la thèse :

A. Koivogui*, H. Ait Hadad, G. Abihsera, C. Vincelet, T. Le Trung, H. Delattre, A. Bernoux, J. Nicolet, D. Lamarque. Dépistage organisé du cancer colorectal en France : délai de réalisation d'un second test après un premier test non analysable et gravité des lésions cancéreuses.* ***Journées Francophones d'Hépato-gastroentérologie et d'Oncologie Digestive, mars 2022****, Poster P.048. Accessible: https://www.snfge.org/content/depistage-organise-du-cancer-colorectal-en*

A. Koivogui*, C. Vincelet, T. Le Trung, G. Abihsera, H. Delattre, H. Ait Hadad, A. Bernoux, J. Nicolet, D. Lamarque. Accès à la coloscopie de dépistage en Ile-de-France (IDF), dans le contexte de la crise sanitaire liée au Covid-19.* ***Journées Francophones d'Hépato-gastroentérologie et d'Oncologie Digestive, mars 2022****, Poster P.047. Accessible: https://www.snfge.org/content/acces-la-coloscopie-de-depistage-en-ile*

C. Vincelet, M. Husson, H. Ait Hadad, A. Bernoux, H. Delattre, ***A. Koivogui****, T. Le Trung, C. Pelisson, A. Kaufmanis. Participation des Franciliens au dépistage du cancer colorectal suite à l'envoi d'un test à domicile en relance 2.* ***Journées Francophones d'Hépato-gastroentérologie et d'Oncologie Digestive****, mars 2022, Poster P.020. Accessible: https://www.snfge.org/content/participation-des-franciliens-au-depistage*

A. Koivogui*, C. Balamou, H. Ait Hadad, R. Rymzhanova, S. Cornelis, A. Kaufmanis, J.M. Sabaté, R. Benamouzig. Résultats et fréquence des complications liées à la coloscopie réalisée après l'âge de 74 ans dans le programme de dépistage organisé du cancer colorectal en France.* ***Journées Francophones d'Hépato-gastroentérologie et d'Oncologie Digestive, mars 2020****, Poster P.143. Accessible: https://www.snfge.org/content/resultats-et-frequence-des-complications*

A. Kaufmanis, C. Vincelet, ***A. Koivogui****, J. Deyra, H. Ait Hadad, H. Delattre-Massy, Z. Brixi, S. Bercier, T. Le Trung. Colorectal cancer screening and choice of the positive cut-off level. Are French decision making criteria optimal in relation to a person's equal access to colonoscopy in regards to advanced neoplasia risk? Congès Barcelone, UEG (****United European Gastroenterology****) Abstract book oct 2019, Poster P1147.* ***United European Gastroenterology Journal*** *Volume 7 Issue 8 (Supplement), October 2019. doi: 10.1177/2050640619854671 PMID: 32213001*

1. INTRODUCTION

1.1. Le programme dépistage organisé du cancer colorectal en France (DOCCR)

1.1.1 – INTERET DE LA MISE EN PLACE DU PROGRAMME DEPISTAGE ORGANISE DU CANCER COLORECTAL (DOCCR)

Toutes les conditions justifiant la mise en place d'une politique de dépistage sont remplies par le cancer colorectal (CCR) : il est fréquent (43 000 nouveaux cas en France chaque année), grave (on ne guérit actuellement qu'un cancer sur deux) et s'il est diagnostiqué au début, il est habituellement guéri par exérèse. C'est un cancer qui a la particularité d'être le plus souvent précédé par une tumeur bénigne, le polype adénomateux. Le cancer peut être prévenu par la détection et l'exérèse des adénomes. Enfin, il existe une stratégie de dépistage reposant sur la recherche d'un saignement occulte dans les selles par un test réalisé tous les 2 ans[1-3].

Cependant, il a fallu attendre plusieurs années après la publication des premières études apportant la preuve de l'efficacité du dépistage [3, 4], pour voir émerger de véritables programmes régionaux ou nationaux de dépistage du CCR en Europe. Plusieurs années se sont écoulées entre la publication des études princeps et la diffusion (fin 1999) des recommandations faites à Vienne par le groupe des experts cancérologues de la Commission Européenne et l'inscription du dépistage du CCR dans le Code européen contre le cancer [5, 6]. A l'exception de l'Allemagne qui avait mis en place un programme en 1977 sous décision du Parlement (en l'absence de preuve établie d'efficacité du dépistage)[6], la plupart des pays membres de l'Union Européenne (UE) ont mis en place un programme seulement après 2005. En France, le programme national de dépistage organisé du CCR (DOCCR) a débuté au début des années 2000 dans 13 puis 23 départements pilotes (Allier–03–, Alpes-Maritimes–06–, Ardennes–08–, Bouches-du-Rhône–13–, Calvados–14–, Charente–16–, Côte-d'Or–21– , Finistère–29–, Hérault–34–, Ille-et-Vilaine–35–, Indre-et-Loire–37–, Isère-38–, Marne–51–, Mayenne–53–, Moselle–57–, Nord–59–, Orne–61–, Puy-de-Dôme–63–, Pyrénées-Orientales–66–, Haut-Rhin–68–, Saône-et-Loire–71–, Essonne–91–, Seine-Saint-Denis–93–)[7]. Au vu, des premiers résultats, la généralisation du dépistage à tout le territoire a pu être annoncée en 2005 et devenue une réalité en 2009[6-8].

Une évaluation récente du DOCCR montrait que la réalisation d'un test fécal immunologique (FIT) tous les deux ans dans une population à risque moyen était la stratégie la plus rentable parmi les stratégies évaluées[9]. Le test

utilisé actuellement a fait la preuve de son efficience en termes de diminution de la mortalité par cancer colorectal [2, 3, 10-13].

1.1.2 – ORGANISATION DU DOCCR EN FRANCE

En France, les campagnes de DOCCR sont organisées selon les modalités d'invitation, de réalisation du test, et de suivi des résultats positifs conformément à un cahier des charges du DOCCR[8]. De son lancement officiel en 2009 jusqu'en 2019, l'organisation des campagnes de DOCCR était gérée par des structures départementales de gestion du DOCCR (SGDO). Ces SGDO étaient majoritairement (>95%) des associations Loi 1901 ou des groupements d'intérêt public (GIP). Pour faire suite à l'arrêté du 23 mars 2018[14], les centres régionaux de coordination des dépistages des cancers (CRCDC) ont été créés (100% associations Loi 1901) dans chaque région par une démarche de fusion/absorption des SGDO départementales.

Bien que l'organisation des campagnes de DOCCR ait été transférée à un niveau régional, les démarches d'invitation des personnes cibles et de suivi des résultats de dépistage, sont restées inchangées (***Figure 1.1.2.1)***. A un rythme trimestriel, une mise à jour des fichiers de personnes éligibles était faite après la transmission des données individuelles (Nom et Prénom, date de naissance, sexe, Numéro de sécurité sociale, adresse de résidence) par les différents régimes d'assurance maladie (RAM). Les personnes étaient ainsi invitées à leur 50ème anniversaire (primo-invitation) et tous les deux ans après un test de dépistage (Invitation-subséquente) jusqu'à l'âge de 74ans révolus.

Pour chaque campagne, toute personne invitée et n'habitant pas à l'adresse indiquée (NPAI) était réinvitée lors d'une campagne ultérieure si son éligibilité était confirmée par son RAM. Par contre, toute personne ayant refusé de participer n'était pas réinvitée lors d'une campagne ultérieure.

Il existait, de 2003 à 2014 puis à partir de 2018, deux phases (Médicale et Postale) de distribution des kits de tests de dépistage. Une phase médicale, au cours de laquelle, les personnes ciblées par la campagne recevaient, à la date anniversaire, une lettre leur invitant à retirer le test de dépistage chez le médecin généraliste. Les personnes n'ayant pas réalisé le test après cette invitation initiale étaient relancées une première fois, trois mois après R0. Trois mois après la première relance, une phase postale était organisée à l'attention de toutes les personnes n'ayant pas retiré le test chez leur médecin généraliste. Appelé relance postale avec test (RPT), elle consistait en l'envoi postal par la SGDO, d'un kit de test de dépistage, accompagné d'un deuxième courrier de relance.

Au terme de la séquence d'invitation/relance, toute personne ayant réalisé un test de dépistage, était réinvitée deux ans après un résultat négatif. En cas de résultat positif, la personne était réinvitée cinq ans après une coloscopie négative et exclue du DOCCR après une coloscopie positive (adénome à risque, CCR). Pour celles qui n'avaient pas réalisé un test mais qui avaient (ou leurs représentant) adressé une réponse postale à la SGDO, une exclusion définitive du DOCCR était décidée si celle-ci rapportait un diagnostic d'une pathologie chronique du colon-rectum (Exemples : CCR, Maladie de Crohn, adénome à risque) ou des antécédents familiaux de CCR (CCR chez un parent 1er degré avant 65ans, CCR chez les deux parents 1er degré, Polypose Adénomateuse Familiale) justifiant d'une exclusion médicale. De même, les personnes décédées, celles qui ont déménagé étaient exclues définitivement du DOCCR, du moins dans le département de résidence. Une exclusion temporaire était décidée dans tous les cas où la réponse postale évoquait une coloscopie de moins de cinq ans ou un coloscanner de moins de deux ans ou une recherche de sang dans les selles de moins de deux ans. A partir du 1er mai 2019, les SGDO de chaque région ont fusionné pour constituer une SGDO régionale. Cette régionalisation n'a apporté aucun changement dans l'organisation du DOCCR sus-décrite.

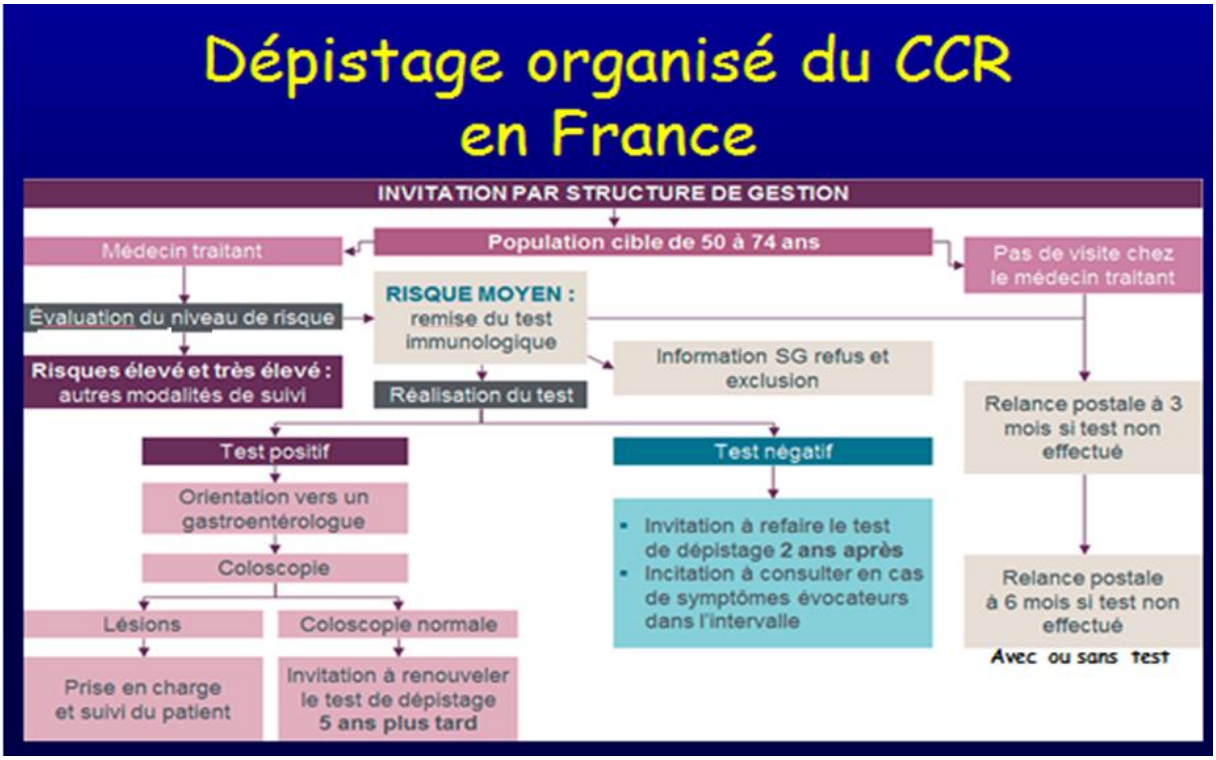


Figure 1.1.2.1 : Schéma d'organisation du DOCCR

1.1.3 – PROBLEMATIQUE DE LA FAIBLE PARTICIPATION AUX CAMPAGNES DE DOCCR

Malgré l'intérêt démontré du programme de dépistage, la problématique de la participation aux campagnes reste encore non résolue dans la plupart des programmes européens. Les premières évaluations faites avaient révélé que peu de programmes organisés de dépistage du CCR et d'études pilotes atteignaient le taux de participation minimal de 45% que recommande l'UE encore moins le taux de participation souhaitable entre 65 et 70%[15, 16]. La couverture du dépistage (proportion de ceux effectivement dépistés de l'ensemble de la population cible éligible) variait de 7,0% à Bruxelles et en Wallonie (Belgique) à 67,7% en Finlande[17]. La participation suite à une première invitation variait de 7,0% (Bruxelles & Wallonie, Belgique) à 64,3% (Région basque, Espagne)[17].

En France, le taux de participation aux campagnes biannuelles dépasse rarement 50% dans la plupart des départements français [7, 18, 19]**.** Dans le département de la Seine-Saint-Denis comme dans les autres départements franciliens de la petite couronne, le taux de participation n'a jamais dépassé 35 % depuis la première évaluation du programme [20].

Toutefois, avec ce niveau actuel du taux de participation, le programme français réduirait l'incidence du CCR de 5% et la mortalité par CCR de 14% par rapport à une population indemne de tout programme de dépistage [21]. Le programme permettrait de réduire l'incidence de 3% et 8% supplémentaires pour des taux de participation de 45% et 65% respectivement[21].

Plusieurs raisons ont déjà été évoquées pour expliquer la faible participation à une campagne de dépistage [18, 22-25]. Le lien entre la participation à une campagne et le sexe a été documenté [18]. De même, l'augmentation du taux de participation avec l'âge est établie[7]. La forte mobilisation des personnes ayant un régime spécial a été aussi établie [22]. Pour Pollack, le niveau socioéconomique et le type d'assurance du patient sont des facteurs influençant la décision du médecin à recommander ou non le dépistage [23]. La littérature rapporte aussi que de par ses caractéristiques, le type de test utilisé peut être un obstacle à la participation. En exemple, l'Hemoccult©(utilisé dans plusieurs programme, notamment en France) n'identifiait que 50 % des lésions cancéreuses et 1/3 des adénomes[24]. En voyant certains de leurs patients passer entre les mailles du filet, une partie des Médecins généralistes (MG) avait fini par s'en méfier. La nécessité de prélever plusieurs échantillons de selles pendant trois jours avait créé des réticences côté patient[24, 25].

1.1.4 – STRATEGIES ET MOYENS ACTUELS DEPLOYES POUR AUGMENTER LA PARTICIPATION AUX CAMPAGNES DE DOCCR EN FRANCE ET AILLEURS

Pour optimiser le taux de participation, plusieurs stratégies ont été proposées [26-34], avec des résultats très variables. Parmi elles, l'envoi du test aux personnes ciblées, utilisé avec des taux de participation de 50 à 70% au Royaume-Uni, en Hollande ou en Finlande ne permettait d'obtenir qu'une participation de l'ordre de 15% en France[35] avec un coût individuel estimé à 33,59 € [36]. Ceci avait conduit à l'idée qu'en France, il fallait associer le médecin généraliste et l'envoi postal pour obtenir un taux de participation d'au moins 45%[37]**.**

Ainsi, après la généralisation du dépistage en 2009 en France, l'arrêté du 29/09/2006 (JORF 21/12/2006) prévoyait une phase de distribution médicale et une phase de distribution postale[8]. Ce qui permettrait d'atteindre les individus qui ne consultent pas un médecin ou ceux dont le médecin n'est pas motivé pour participer au programme de dépistage. L'arrêté du 22/02/2013 a rendu non obligatoire la Relance Postale avec Test (RPT), le laissant à la discrétion des Agences Régionales de Santé (ARS)[37]. La conséquence a été un refus de financer le coût de la RPT dans la plupart des départements[37]. La RPT a été réintroduite dans le programme français par l'Arrêté du 19 mars 2018 recommandant qu'un kit de dépistage soit joint au second courrier de relance lorsque la personne a participé au moins, à l'une des trois précédentes campagnes[38]. Le 1er mars 2022, un site de commande des kits de test via internet a été mis en place par la Caisse nationale de l'assurance maladie (CNAM). Un rôle actif des pharmaciens pour la distribution des tests a ensuite été permis par les évolutions réglementaires. Quelques mois plus tard (en juin 2022), les SGDO ont été instruites pour mettre en place une formation des pharmaciens d'officine à la remise du kit de test en pharmacie.

1.1.5 – FIABILITE DES DONNEES D'INVITATION DE LA POPULATION CIBLE DE DOCCR

Dans le contexte français, il semble pertinent d'investiguer la fiabilité des données d'invitation fournies par la CNAM et la méthode d'estimation du taux de participation. Dans la région Ile-de-France (IDF), il ressort du tableau de bord régional du DOCCR, que seulement 4,5% de causes de non-participation (hors causes médicales) sont renseignées au niveau régional, avec des variations départementales importantes, allant de 0,1% pour Paris à 11,5% en Seine-Saint-Denis (***Figure 1.1.5.1***). Représentant environ 90% de ces causes de non-participation renseignées, les plis non distribuables (PND) parce que la personne invitée à la campagne n'habite pas à l'adresse indiquée (NPAI), seraient la première cause de non-participation au DOCCR[39]. En plus d'une variation interdépartementale cette proportion de NPAI connait une variation intercommunale considérable au

sein d'un même département. Au cours d'une récente évaluation des campagnes départementales de DOCCR dans le département de la Seine-Saint-Denis, six villes s'étaient distinguées avec une proportion élevée de NPAI supérieure à 10% des personnes invitées à la campagne (*Source : Rapport bureau CDC93, 2017*). Ces NPAI décomptés seulement après l'envoi des courriers d'invitation/relance pourraient avoir leur origine dans les données sources des différents régimes d'assurance maladie. Une forte proportion de NPAI tel que connue en IDF, pourrait avoir un impact considérable sur le taux de participation.

Tableau 12 : Nombre de non participants au DOCCR au cours de la campagne 2015-2016 et part des causes connues (Source : SGDO)

	75	77	78	91	92	93	94	95	Ile de France
Nombre de femmes et d'hommes non participants après une invitation en 2015-2016	558 584	258 929	304 840	235 068	347 704	305 156	267 688	237 024	2 514 993
Nombre de femmes et d'hommes pour lesquelles la (les) cause(s) est (sont) connue(s), hors exclusions médicales, après invitation en 2015-2016	562	11 864	13 667	7 218	25 853	35 096	18 774	1 208	114 242
% de causes connues	*0,1%*	*4,6%*	*4,5%*	*3,1%*	*7,4%*	*11,5%*	*7,0%*	*0,5%*	*4,5%*
Répartition des causes de non-participation après invitation en 2015-2016, hors causes médicales									
Décès		*1,5%*	*1,6%*	*2,8%*	*0,0%*	*0,3%*	*0,3%*		*0,8%*
PND		*86,9%*	*86,2%*	*81,8%*	*98,9%*	*93,4%*	*87,5%*		*90,0%*
Déménagement		*2,0%*	*1,5%*	*4,8%*	*0,1%*	*0,6%*	*1,0%*		*1,5%*
Refus		*9,6%*	*10,7%*	*10,7%*	*1,0%*	*5,7%*	*11,3%*		*7,7%*
Ensemble		*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*	*100,0%*		*100,0%*

Figure 1.1.5.1 : Causes (hors causes médicales) de non-participation à la campagne de DOCCR 2015-2016 en Ile-de-France, synthèse du Tableau de bord régional [39]

1.1.6 – FIABILITE DE LA METHODE D'ESTIMATION DU TAUX DE PARTICIPATION

Dans son protocole d'évaluation du programme de DOCCR, Santé Publique France (SPF) définit le taux de participation par le rapport entre le nombre de personnes dépistées et la population éligible, soit la population cible du programme (hommes et femmes de 50 à 74 ans) dont sont soustraites les personnes exclues du programme pour raisons médicales[40]. La définition de cette population éligible est basée sur la notion de risque moyen caractérisé par l'âge (50-74 ans) et l'absence de risque personnel ou familial connu [8]. Ainsi des personnes à haut risque (antécédents personnels ou familiaux d'adénomes colorectaux, de cancers ou de maladies inflammatoires : colite ulcéreuse et maladie de Crohn) ou à très haut risque (polypose adénomateuse familiale ou syndrome de Lynch) de développer un cancer colorectal et les personnes atteintes de pathologies extra-intestinales sévères sont censées être exclues de cette population éligible. De même, les personnes ayant réalisé une coloscopie dans les 5 ans sont temporairement exclues si le résultat de cette coloscopie est normal [7]. Lors de la première évaluation du programme, ce taux d'exclusion était estimé à 10.5% dans une population de 9.710.960 personnes invitées à la campagne de 2008-2009 dans 46 départements de France [7]. En IDF, le taux d'exclusion standardisé était de 9,9% (***Annexe-2***) dans la cohorte de personnes invitées à la campagne biannuelle de 2015-2016 comparativement au taux national qui était de 12,3% [39].

Une enquête réalisée en Seine-Saint-Denis avait montré que cette approche d'estimation du taux d'exclusion était sous-optimale et que de nombreux sujets étaient considérés comme à risque moyen alors qu'ils étaient à risque élevé et que d'autres étaient considérés comme non participants alors qu'ils étaient en fait « couverts » par une coloscopie récente non déclarée. Dans ce département de Seine-Saint-Denis, le taux de participation avait été estimé à 22,37% dans la cohorte des 388.652 personnes âgées de 50-74 ans, invitées à la campagne biannuelle de 2018-2019. Dans cette cohorte, la proportion des exclusions pour cause médicale n'était que de 2,39%. Ces chiffres pourraient être sous-estimés car en 2018 dans le département, le taux standardisé des personnes prises en charge pour maladies inflammatoires chroniques intestinales était de 4‰ ; le taux standardisé des personnes prises en charge pour un cancer colorectal actif était de 2,2‰ et pour un cancer colorectal sous surveillance de 3,2‰ [41]. Le nombre de coloscopies réalisées en France, en 2016, était estimé à 1.402.047 dont 76,9% chez les personnes âgées de plus de 50 ans [42]. En tenant compte de la part de Seine-Saint-Denis (2,5% en 2017) dans la population française, on estimerait à 26.954 le nombre annuel de coloscopies réalisées chez les Séquano-Dionysiens âgés de plus de 50 ans, ce qui représente environ 7% de la population ciblée.

Le fait que le taux d'exclusion soit estimé seulement après un recueil de déclarations volontaires des sujets (questionnaire joint au courrier d'invitation) et des médecins traitant (questionnaire structuré) pourrait compromettre un recueil exhaustif de ces cas potentiels d'exclusion médicale.

A propos de la méthode de recueil, SPF confirmait que le nombre de personnes exclues était fourni par les structures de gestion à partir de diverses sources de données exploitables (réponses des personnes invitées, fichiers des actes de coloscopies remboursés par les caisses primaires d'assurance maladie, fichiers des actes diagnostiques et thérapeutiques, comptes rendus standardisés des résultats de coloscopie et des examens anatomo-cytopathologiques…), mais dont l'accès variait selon les structures de gestion [40].

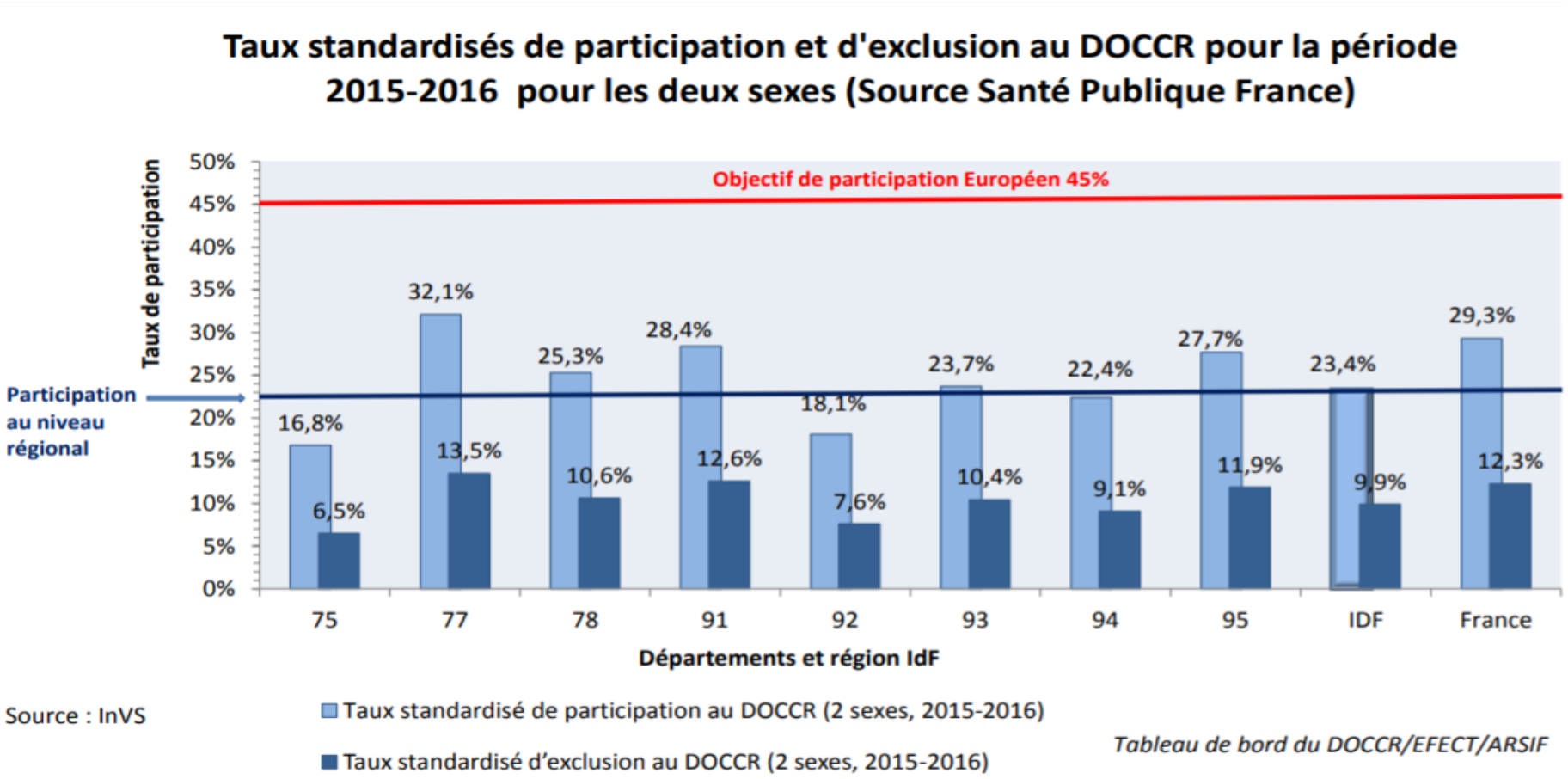


Figure 1.1.6.1: Taux d'exclusion à la campagne de DOCCR 2015-2016 en Ile-de-France, synthèse du Tableau de bord régional [39]

1.2. Les sources de données exploitées et exploitables pour une meilleure détermination du taux d'exclusion aux campagnes

1.2.1 – REVUE SUR LES BASES DE DONNEES MEDICO-ADMINISTRATIVES ET REGISTRES DE CANCERS ACTUELLEMENT EXPLOITES DANS LES PROGRAMMES DE DEPISTAGE

Il existe dans la plupart des pays pratiquant le dépistage du CCR au moins une base de données médico-administrative (BDMA)[43-63] ou un registre de cancers[43, 46, 49, 62, 64-77], couvrant totalement ou partiellement, la population. L'usage des BDMA pour la sélection ou l'affinage des populations éligibles au DOCCR est peu documenté dans la littérature. Aux Etats-Unis d'Amérique (USA), bien que l'approche du dépistage soit majoritairement opportuniste, il existait des programmes organisés (1-*National Demonstration Program, Centers for Disease Contrôle and Prevention -CDC-, 2-Delaware Statewide Cancer Control Program, 3-Cancer Prevention, Education, Screening and Treatment Program, Maryland, 4-New York City Program, 5-Kaiser Permanente of Northern California -KPNC-, 6-Group Health Cooperative of Puget Sound, 7-Veterans Health Administration (VHA), 8-Independence Blue Cross -IBC-*) ciblant des populations ou groupes spécifiques[78] et utilisant des BDMA dédiées. En illustration, la Medicare claims database (BDMA qui couvre 95 à 98 % des personnes de ≥ 65 ans et celles de moins de 65ans ayant un handicap ou avec une insuffisance rénale terminale) est utilisée par les projets CDC[79]. En Europe, bien que rare, l'usage des BDMA dans les programmes de dépistage est aussi documenté dans la littérature. Au Portugal [80], la population cible des campagnes est directement extraite de la BDMA.

Bien que l'usage des BDMA et registres de cancer soit assez fréquent dans les programmes, il n'existe pas à ce jour, un phénotype standard de la personne à exclure des campagnes de DOCCR à partir de ces BDMA et registre. Cependant, pour la plupart des facteurs morbides justifiant une exclusion médicale du DOCCR, il existe des algorithmes de recherche de phénotypes validés et reproductibles dans les BDMA répertoriées[43, 62, 81-128].

1.2.2 – LES BASES DE DONNEES EXPLOITABLES PAR LE PROGRAMME DE DEPISTAGE EN FRANCE

1.2.2.A. – LE SYSTEME NATIONAL DES DONNEES DE SANTE (SNDS)

Architecture et données disponibles

Le Système National des Données de Santé (SNDS) a été instauré par la loi du 26 janvier 2016 de modernisation du système de santé et dont les modalités de mise en œuvre ont été approfondies par la loi du 24 juillet 2019 puis le décret du 29 juin 2021. Son objectif est de développer l'usage des données de santé, d'analyser et d'améliorer la santé de la population en croisant plusieurs sources de données[129]. Le SNDS est constitué de 3 sources principales chaînées : i)Les données du Datamart de Consommation Inter-Régimes (DCIR) issues du Système National d'Information Inter-régimes de l'Assurance Maladie (SNIIRAM), qui sont exhaustives sur les dépenses de l'Assurance Maladie et gérées par la Cnam ; ii) Le programme de médicalisation des systèmes d'information (PMSI) qui est géré par l'ATIH et qui collecte les données concernant les établissements de soins ; iii) Les causes médicales de décès qui sont des données issues des certificats de décès, gérées par le CépiDc de l'Inserm [129-131]. Deux nouvelles sources qui avaient été créées lors de l'épidémie de COVID-19 pour suivre différents indicateurs, ont été implémentées (ou sont en phase de l'être) conformément à un arrêté ministériel datant du 12 mai 2022 : i)La base vaccin covid, VAC-SI (labélisé : Système d'Information Vaccin COVID) contient les informations concernant les vaccinations contre la COVID-19 ; ii) La base SI-DEP qui est une base du Système d'Information national de Dépistage Populationnel de la COVID-19, contient les informations sur les résultats des tests de dépistage de la COVID-19[129].

Globalement, le SNDS est une base de données (BD) collectant les données individualisées et anonymes de l'intégralité des remboursements de dépenses de santé des secteurs publics ou privés pour les sujets affiliés à l'un des régimes obligatoires de l'assurance maladie (***Figure 1.2.2.1***), soit près de 99% (67 millions de personnes) des personnes résidant en France.

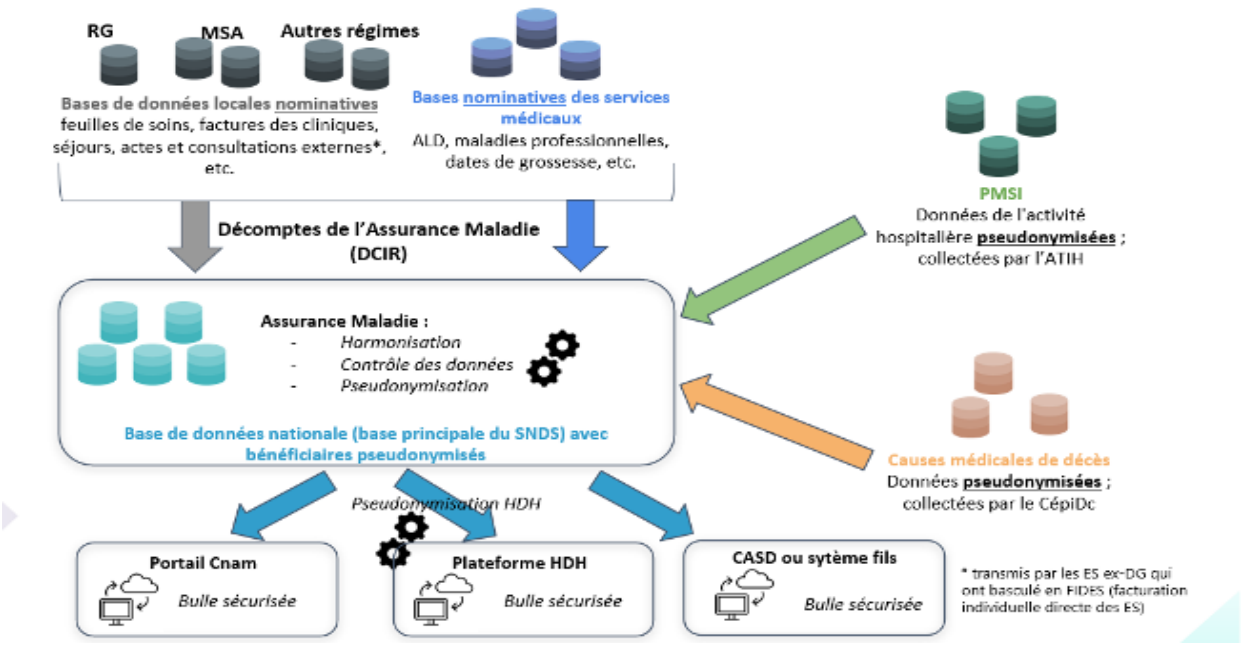


Figure 1.2.2.1 Sources d'alimentation du SNDS. Source : Health-Data-Hub [129]

Les données du SNDS sont restituées, sur le portail SNDS géré par la CNAM, sous trois systèmes: i) Les 15 bases de données agrégées thématiques, notamment les datamarts sur différents domaines spécifiques (soins libéraux, biologie médicale, pharmacie, dispositifs médicaux, etc.) ; ii) Les données individuelles exhaustives, comprenant les données de soins de ville, des séjours hospitaliers et des causes médicales de décès de la totalité des bénéficiaires en France, permettant de réaliser des études sur près de 99% de la population française ; iii) L'Echantillon de Bénéficiaires du SNDS (ESND), constitué de données individuelles échantillonnées (échantillon issu d'un tirage aléatoire simple au 2/100ème des bénéficiaires de la base principale du SNIIRAM). L' ESND est représentatif par âge, sexe, grands régimes d'affiliation, région et département de résidence des bénéficiaires. Cet échantillon possédant la même structure que les données individuelles exhaustives, permet la réalisation d'études longitudinales à l'aide de données de villes (consommations de soins) et hospitalières[129].

En accédant au SNDS à travers ce portail géré par la CNAM, un identifiant anonymisé unique permet de suivre une personne dans tout le système de soins mais surtout dans chacune des bases du SNDS. Les nombreuses sources de données qui constituent le SNDS et le complètent (en exemples, les informations complémentaires sur les bénéficiaires, sur les professionnels ou les établissements de santé, les nomenclatures, etc.) sont actuellement des bases de données organisées et structurées. La consultation et les opérations sur ces bases sont facilitées par leur modèle relationnel[129]. De plus, le principal atout du SNDS est d'avoir apparié l'ensemble de ces données entre elles, tel que schématisé (***Figure 1.2.2.2***) par le Health-Data-Hub (HDH).

Les maladies sont identifiées selon la dixième révision de la Classification Internationale des Maladies de l'OMS (CIM-10)[130]. Les actes médicaux sont codés selon la classification commune des actes médicaux (CCAM). Les médicaments sont enregistrés avec leur code CIP pour la dispensation en ville et associé à un code ATC (classification anatomo thérapeutique et chimique de l'OMS), les actes de biologie sont codés à l'aide de la NABM. Les données hospitalières du PMSI comportent les informations médicales associées à un séjour hospitalier (diagnostics pris en charge, actes médicaux, médicaments en sus du groupe homogène de séjours -GHS- administrés en milieu hospitalier, …). Les diagnostics gérés en hospitalisation sont codés en diagnostic principal (DP) qui peut être associé à un diagnostic relié (DR) et des diagnostics associés (DAS). Les patients ayant une maladie chronique bénéficient d'une prise en charge à 100% de leurs frais de santé au titre d'une affection de longue durée (ALD). L'ALD est enregistrée avec les 3 premiers caractères du code de la CIM-10 et une date de début de validité. En outre, le SNDS contient des données sociodémographiques ainsi que, le cas échant, la date et la cause de décès du patient [131].

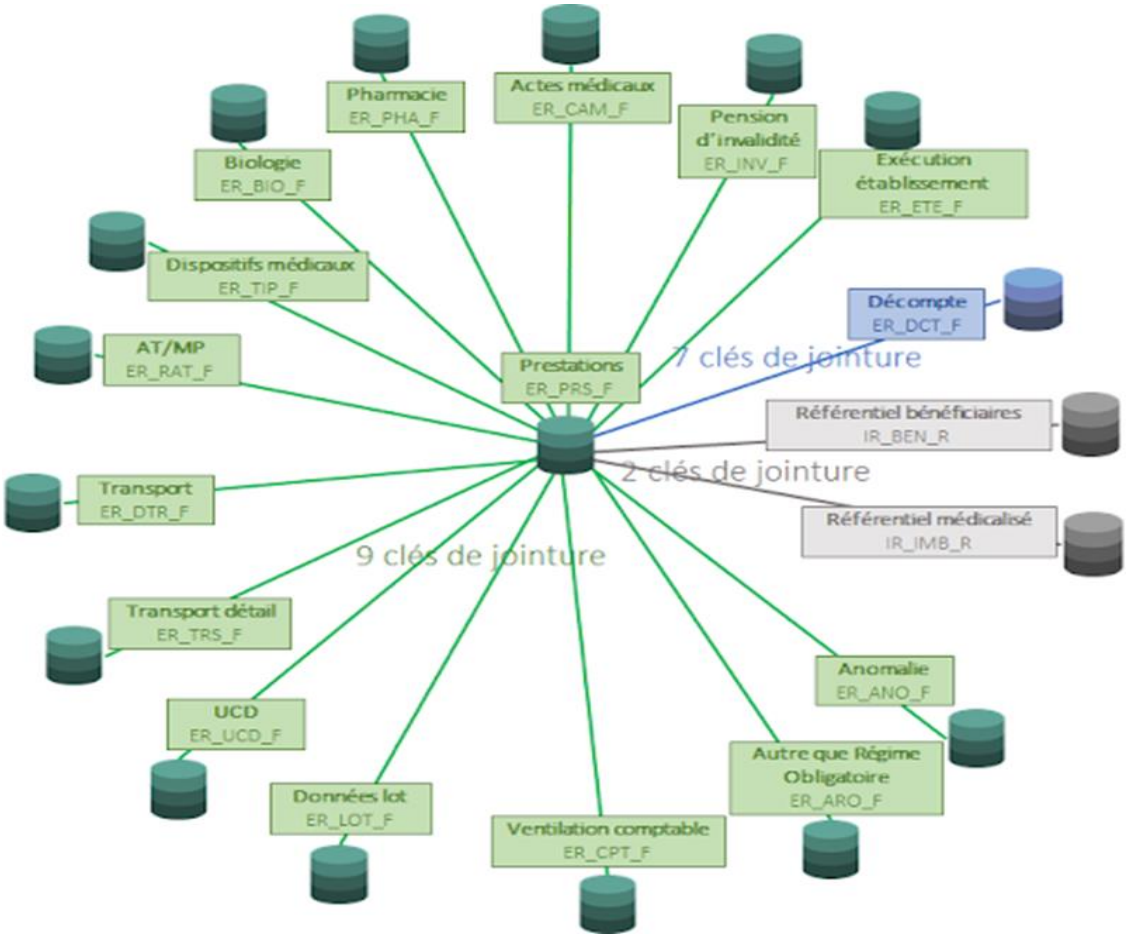


Figure 1.2.2.2 : Schéma relationnel de la base de données DCIR. Source Figure Health-Data-Hub (HDH)[129]

Exemples d'exploitation du SNDS

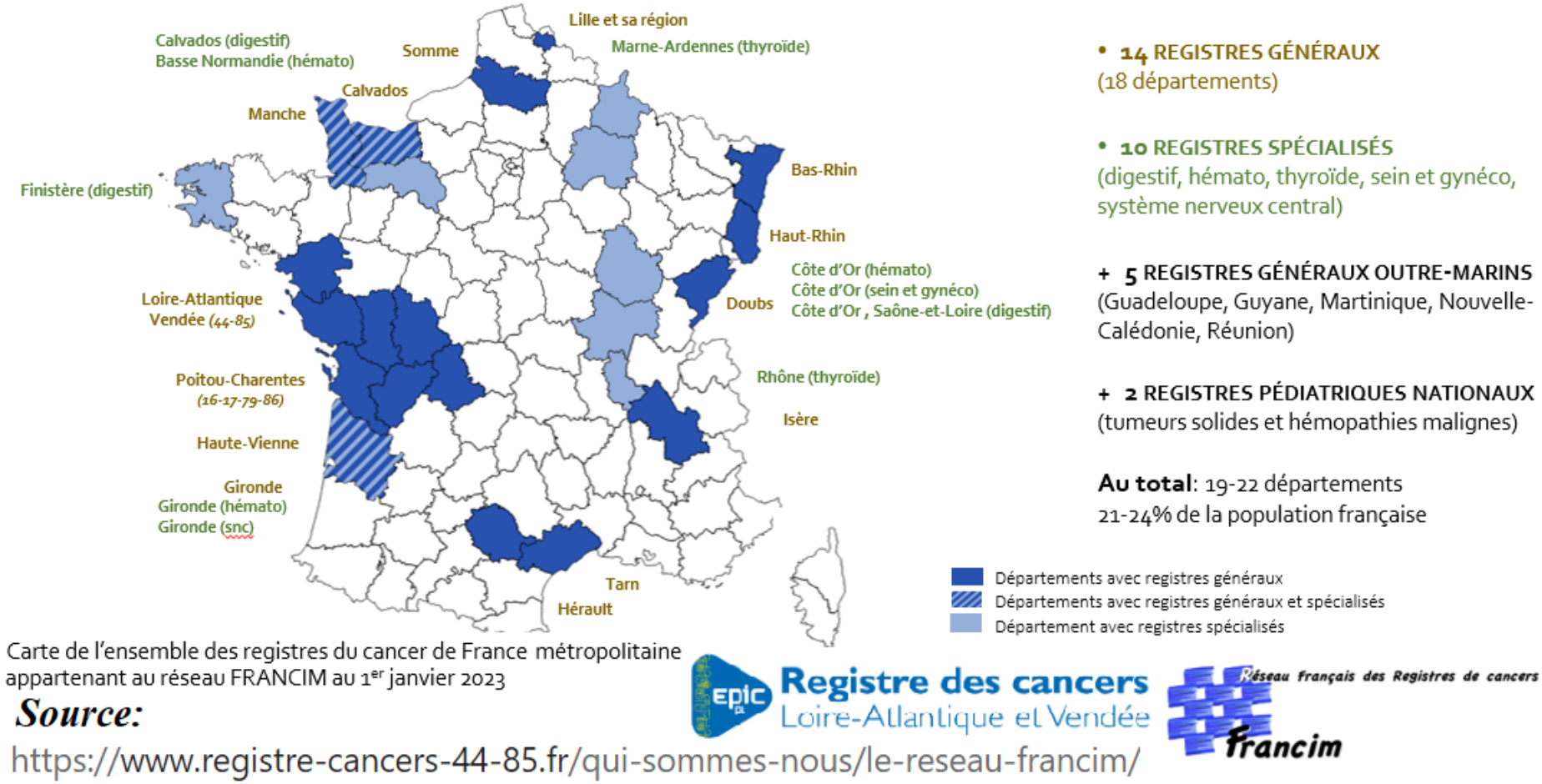
La Caisse nationale de l'assurance maladie (CNAM) tient à jour dans un répertoire dédié (https://assurance-maladie.ameli.fr/etudes-et-donnees/etudes-publications/references-publications-articles-assurance-maladie), une liste exhaustive de travaux réalisées sur les bases du SNDS. Il s'agit de travaux réalisés par l'Assurance Maladie ou d'autres équipes institutionnelles en partenariat avec la CNAM, publiés dans des revues à comité de lecture. A la date du 25 mars 2022, nous avons répertorié 339 travaux réalisés sur au moins une des bases du SNDS (PMSI, ALD, SNIIRAM, CEPIDC) et publiés entre janvier 2010 et décembre 2021.

1.2.2.B. – LES REGISTRES DE CANCERS EN FRANCE

Architecture et exploitation des données disponibles des registres de cancers

A ce jour, la France est l'un des rares pays de l'UE qui n'ont pas encore un registre national de cancer. Il ressort d'une analyse récente qu'environ 20% de la population française est couverte par un registre de cancer[132]. Toutefois, au 30 novembre 2019, 32 registres des cancers (RC), couvrant chacun un ou plusieurs départements, ont été évalués par le Comité d'évaluation des registres (CER). 27 registres des cancers envoient leurs données annuellement pour alimenter la base commune du réseau français des registres des cancers (FRANCIM). Ces 27 registres intègrent le partenariat financier et scientifique « FRANCIM, Hospices civils de Lyon (HCL), Institut national du cancer (Inca) et SPF» (Source : Inca, https://www.e-cancer.fr). Parmi ces 27 registres, 23 (RC-CCR) pourraient avoir le CCR parmi les cancers faisant l'objet d'une collecte de données ((***Annexe-3***).

Chaque RC recense les cancers diagnostiqués chez des personnes résidant dans le(s) département(s) couvert(s) par le RC. En illustration, le RC de la Loire-Atlantique et Vendée (https://www.registre-cancers-44-85.fr) recense les cancers/tumeurs de la Loire-Atlantique (depuis 1991 pour les cancers du sein et du colon-rectum et depuis 1998 pour tous les cancers/tumeurs) et de la Vendée (tout cancer depuis 1998). Pour une année N donnée, le processus de recueil et de croisement d'informations de différentes sources (récupérées à N+1), d'enquêtes et de saisie des données (N+1/N+2), de réalisation des contrôles de qualité (N+2/N+3) et de validation des données (N+3) nécessite un travail d'environ 3 ans (https://www.registre-cancers-44-85.fr). Les données issues des RC servent à la fois à l'estimation de l'incidence des cancers en France et au suivi de l'évolution de cette incidence[133, 134].



Carte de l'ensemble des registres du cancer de France métropolitaine appartenant au réseau FRANCIM au 1er janvier 2023

Source:

https://www.registre-cancers-44-85.fr/qui-sommes-nous/le-reseau-francim/

Figure 1.2.2.3 : Cartographie des registres de cancers.

1.3. Hypothèse

Il ressort de la synthèse de la littérature (ci-dessus) un faible taux de participation aux campagnes de dépistage en France mais aussi un mauvais ciblage de la population éligible à une campagne. Il existe pourtant des sources de données qui pourraient aider à mieux sélectionner les personnes éligibles aux campagnes de dépistage mais ces sources ne sont pas systématiquement utilisées.

Au regard de ce qui précède, nous avons émis l'hypothèse que l'exploitation de plusieurs sources de données, notamment les bases de données médico-administratives et les registres de cancers, permettrait de mieux cibler et caractériser les personnes éligibles aux campagnes de dépistage en France, ce qui améliorerait considérablement le taux d'exclusion médicale. L'exploitation des réseaux sociaux quant à elle, permettrait d'atteindre les personnes non participant aux campagnes de dépistage, ce qui améliorerait considérable l'effectif de participants à une campagne de dépistage. Si au moins l'une des deux hypothèses est affirmative, cette thèse aura documenté une stratégie d'amélioration du taux de participation aux campagnes de dépistage en France.

2. OBJECTIFS DE LA THESE ET AXE DE RECHERCHE

2.1. – Objectif principal

Développer des approches pour mieux cibler, caractériser et atteindre les patients relevant du DOCCR en s'intéressant en particulier à l'apport de la réutilisation de données médico-administratives et l'usage de réseaux sociaux.

2.2. – Objectifs spécifiques

Axe 1 : Comparer les définitions internationales des populations cible du dépistage et la place des bases de données médico-administratives dans les méthodologies de ciblage des populations éligibles aux campagnes de dépistage.

Axe 2 : Définir un patient à exclure du dépistage à partir des données du SNDS.

Axe 3 : Analyser l'apport des réseaux sociaux comme sources supplémentaires de motivation des personnes à participer aux campagnes de DOCCR.

3. AXE 1 : ANALYSE DE LA DEFINITION ACTUELLE DE LA POPULATION CIBLE DES CAMPAGNES DE DEPISTAGE ET LA PLACE DES BASES DE DONNEES ADMINISTRATIVES DANS LES METHODOLOGIES DE CIBLAGE

3.1 Place des bases de données médico-administratives dans la sélection des personnes éligibles aux campagnes de dépistage : Étude 1.

Tel que recommandé par les instances scientifiques en Europe [135] et en Amérique [136], la plupart des cahiers de charge de DOCCR recommandent le ciblage des personnes à risque moyen de CRC. Ceci sous-entend que les personnes à haut risque (antécédents personnels ou familiaux d'adénomes colorectaux, de cancers ou de MICI : colite ulcéreuse et maladie de Crohn) ou à très haut risque (polypose adénomateuse familiale –PAF– et syndrome de Lynch) de développer un CCR ne sont pas éligibles aux campagnes de DOCCR. Mais à ce jour il n'existe pas une définition standard de la personne à exclure des campagnes de dépistages, ce qui compromet une meilleure segmentation de la population cible sur le niveau de risque de CCR.

3.1.1 – CONTEXTE

Il existe une corrélation positive entre le taux de participation à une campagne de dépistage du cancer colorectal et le taux d'exclusion médicale. Des avis antérieurs sont concordants sur le fait que l'exclusion médicale soit sous-évaluée en France. Par ailleurs, pour la collecte des données d'exclusion, il n'existe pas en France une méthodologie standard et reproductible. Cependant, il existe des bases de données notamment le SNDS, contenant des informations médicales qui pourraient être utiles pour le repérage des cas d'exclusion médicale. Pour extraire les données d'exclusion médicale du SNDS l'usage d'un algorithme performant est indispensable. Il n'existe pas à ce jour en France, un algorithme d'extraction de données du SNDS qui soit exhaustif sur tous les facteurs morbides pouvant justifier d'une exclusion médicale. Pour construire un tel algorithme, il nous a semblé nécessaire d'analyser comment divers organismes de gestion du DOCCR définissaient et identifiaient ces situations morbides afin de retransposer ces définitions pour interroger les bases du SNDS.

3.1.2 – METHODES

L'étude a couplé une enquête transversale sur les stratégies d'exclusion médicale et une revue non systématique de la littérature pour compiler les phénotypes et algorithmes validés dans les bases de données médico-administratives (BDMA) ou registres de cancer des pays ayant un programme de DOCCR et pratiquant une exclusion médicale.

L'enquête transversale a inclus prioritairement, tous les programmes des pays de la zone Europe de l'OMS qui avaient été inclus dans des enquêtes précédentes [137, 138]. L'enquête a été étendue aux non européens ayant un programme organisé à déploiement complet[137]. Le questionnaire adressé dans un courrier électronique a recueilli les informations en rapport avec le type (Opportuniste ou Organisé) du programme de dépistage en cours en 2021, en référence à la définition d'un DOCCR[137]; les critères d'exclusion en vigueur en 2021, conformément aux critères énumérés dans la littérature[8, 16, 17] ; l'apport possible des BDMA et registres de cancers dans l'exclusion médicale. L'enquête a collecté les informations sur l'usage des BDMA pour l'optimisation des exclusions médicale. Pour toutes les bases dont l'existence a été initialement révélée, qu'elles soient connectées ou pas à la base du dépistage, l'enquête a collecté : i) le Type de BDMA (Base de données de remboursement des soins et prestation médicale/paramédicales des compagnies et régime d'assurance maladie ou autres BDMA de données de morbidité hospitalière) et les registres de cancer; ii)la date de création(ou date de début de collecte des données), iii)la zone géographique couverte, iv)la disponibilité des données (Permanente ou durée de conservation limitée).

Pour chaque facteur morbide justifiant une exclusion médicale, une revue non systématique a été faite dans le but de compiler les différentes définitions opérationnelles en vigueur dans la BDMA dont était extraite la population éligible ou les BDMA potentiellement utilisées/utilisables pour affiner la population éligible.

Manuscript #	NMED-A129223
Current Revision #	0
Submission Date	30th Oct 23
Current Stage	Manuscript under consideration
Title	Place of medico-administrative databases in the selection of the target population in population-based colorectal cancer screening programs
Manuscript Type	Article
Collection	N/A
Manuscript Comment	na
Corresponding Author	Dr Akoï Koïvogui (aakoivogui@live.fr) (Université Sorbonne Paris Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, LIMICS, Sorbonne Université)
Contributing Authors	Professor ROBERT BENAMOUZIG , Dr CHRISTIAN BALAMOU , Dr GEMMA BINEFA , Dr Sarah Hoeck , Dr DOMINIKA NOVAK-MLAKAR , Professor CATHERINE DUCLOS
Authorship	Yes
Abstract	Colorectal cancer(CRC) screening in average-risk populations involves filtering a target population based on medical information. The development of a consensual definition of each morbid situation that can justify exclusion from the target population and its translation by a filtering method applicable to medico-administrative databases (MADB) are challenges for programs. Twenty population-based programs out of the thirty-one programs (Australia, England, Manitoba, Ontario, Washington and 26 European countries) that participated in a survey in 2022 practiced medical exclusions. Five consensus exclusion criteria (personal history of: CRC, inflammatory bowel disease-IBD-, adenoma; recent colonoscopy; genetic risk) were highlighted. These consensus criteria did not have the same definition in the MADBs (i.e. in Catalonia, the CRC phenotype includes ICD-10 codes C18-C21, while the C21 code is excluded elsewhere; in France, IBD phenotypes include codes M07.4, M07.5 whereas in Denmark, phenotypes found only include ICD-10 codes K50/K51). The study highlights a disparity in the choice of the type/number of medical exclusion criteria. There is a limited number of consensual criteria, which seem the least subtle because they are easy to collect thanks to the current means available to program management structures. These consensus criteria can be queried in most of the referenced MADBs. However, the use of MADB was not in effect in all countries, compromising comparability between programs.
Subject Terms	Health sciences/Health care/Disease prevention/Preventive medicine Health sciences/Health care/Public health
Show Author Information	Allow Reviewers to see Author information.
Research Square author dashboard	I understand that my manuscript and associated personal data will be shared with Research Square for the delivery of the author dashboard.
Preprint Deposition	No, my co-authors and I would not like to benefit from *In Review*
Competing interests policy	There is **NO** Competing Interest.
Applicable Funding Source	No Applicable Funding
Previous Interactions	Content related to my submission was recently published by this journal Related content/promotion on researcher networks (ResearchGate etc.) Related content/promotion in an e-mail/e-alert

Place of medico-administrative databases in the selection of the target population in population-based colorectal cancer screening programs

Akoï Koïvogui[1,2], Robert Benamouzig[3], Christian Balamou[4], Gemma Binefa[5], Sarah Hoeck[6], Dominika Novak-Mlakar[7], Catherine Duclos[1,8]

[1]Université Sorbonne Paris Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, LIMICS, Sorbonne Université, INSERM, 93000, Bobigny, France.

[2]CRCDC-IDF, Site de Seine-Saint-Denis, 93146, Bondy, France.

[3]AP-HP, Hôpital Avicenne, Service d'Hépato Gastro Entérologie, 93000, Bobigny, France.

[4]CRCDC-AURA, Site de l'Ain Bourg-en-Bresse, 01000, Bourg-en-Bresse, France.

[5]Cancer Prevention and Control Program, IDIBELL, Catalan Institute of Oncology, Barcelona, Spain.

[6]Centre for Cancer Detection, 8000 Bruges, Belgium.

[7]National Institute for Public Health, Ljubljana, Slovenia.

[8]AP-HP, Hôpital Avicenne, Service d'Information médicale, 93000, Bobigny, France.

Correspondence to: Dr Akoï KOÏVOGUI, Université Sorbonne Paris Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, LIMICS, Sorbonne Université, INSERM, 93000, Bobigny, France.
mail: akoi.koivogui@inserm.fr

Abstract

Colorectal cancer(CRC) screening in average-risk populations involves filtering a target population based on medical information. The development of a consensual definition of each morbid situation that can justify exclusion from the target population and its translation by a filtering method applicable to medico-administrative databases (MADB) are challenges for programs. Twenty population-based programs out of the thirty-one programs (Australia, England, Manitoba, Ontario, Washington and 26 European countries) that participated in a survey in 2022 practiced medical exclusions. Five consensus exclusion criteria (personal history of: CRC, inflammatory bowel disease-IBD-, adenoma; recent colonoscopy; genetic risk) were highlighted. These consensus criteria did not have the same definition in the MADBs (i.e. in Catalonia, the CRC phenotype includes ICD-10 codes C18-C21, while the C21 code is excluded elsewhere; in France, IBD phenotypes include codes M07.4, M07.5 whereas in Denmark, phenotypes found only include ICD-10 codes K50/K51). The study highlights a disparity in the choice of the type/number of medical exclusion criteria. There is a limited number of consensual criteria, which seem the least subtle because they are easy to collect thanks to the current means available to program management structures. These consensus criteria can be queried in most of the referenced MADBs. However, the use of MADB was not in effect in all countries, compromising comparability between programs.

Introduction

Colorectal cancer (CRC) is the third most common cancer and the second leading cause of cancer-related deaths worldwide [1]. CRC screening in average risk populations, by looking for occult bleeding in the stool, carried out after every two years in average risk population, is associated with a reduction in mortality from CRC. The decrease in mortality becomes significant when CRC screening exceeds 50% of the target population [2, 3]. The CRC screening approach common in several European countries is the population-based CRC screening programs (CRCSP) with the systematic invitation of a target population and follow-up of people whose primary screening test was positive. In contrast, other countries such as the United States of America (USA) have developed an opportunistic approach with screening by colonoscopy or sometimes by fecal test [4].

CRCSP targets a population at average risk of CRC, defined on criteria of age [4, 5], absence of personal or family risk of CRC [6, 7], or other digestive pathology such as inflammatory bowel diseases (IBD). High risk people (personal or family history of colorectal adenomas, CRC or IBD: ulcerative colitis and Crohn's disease) or very high risk people (familial adenomatous polyposis –FAP– and Lynch syndrome) of developing CRC, people with severe extra-intestinal pathologies and those with symptoms of an intestinal disease, are not eligible for average-risk screening campaigns in most CRCSPs [6, 7]. Similarly, people who have undergone a colonoscopy within 5 years or a CT-colonography within 2 years, are temporarily excluded if the result of this colonoscopy/CT-colonography was normal [6-8]. It follows that screening for CRC in average risk populations involves filtering a target population based on medical information.

While the exclusion criteria appear to be clearly listed[6, 7, 9], but the way in which they are used to select those eligible for a campaign still is poorly documented to date. Some CRCSPs offered the fecal test without exclusion criteria (Brussels and Wallonia in Belgium, Finland, Stockholm in Sweden) [6]. In addition to the variable ineligibility criteria, there was a disparity (1-15%) in terms of the proportion of ineligible people among the people invited to CRCSP's campaign [7]. Wang et al. [9] revealed heterogeneity in targeting populations and inaccessibility of selection data in some national programs. Although these previous studies[6, 7, 9] have described some means of medical exclusion data [7] and the timing of exclusion (before and after invitations) [7], the reasons for choosing a morbid situation as exclusion criteria in such programs and not in another program, were not explicit in their studies.

Similarly, the data used for the campaign invitation can be extracted (frequently) from the medico-administrative databases (MADB), in particular the health insurance companies reimbursement databases (Claims-DB) [10, 11].

The wide range of the exclusion rate reported by Buillard et al. [7], could be attributable to a disparity in the strategies and means used for exclusion, in particular the use of Claims-DB. In France, where the health insurance databases (SNDS: "*Système national des données de santé*") are currently inaccessible to CRCSP's management structures, the exclusion rate (12.9% out of the 2016-2017 campaign) was largely underestimated because 20% of the target population completed a colonoscopy in the last 5 years[12]. Buillard [7] reports that a participation rate estimated at 45% in a target population without excluding potential cases of ineligibility from the denominator, would rise to 50% (gain of 5 points) if 10% of the population were deemed ineligible. These previous studies do not help to understand whether a filter of the target population is systematically carried out on medical information, or whether there was a consensual definition of these morbid situations translatable by a filtering method applicable on databases. The development of a consensual definition of each morbid situation that can justify exclusion of the target population and its translation by a filtering method applicable to MADB are therefore challenges for programs. This study aims of (1) to describe the strategies and means for most correctly targeting people to be invited in the CRCSP, (2) to describe the place of the MADB and cancer registries in the strategies for collecting the exclusion data.

Results

Practice of medical exclusion according to the cross-sectional survey

The form was sent to contacts in forty-one countries and a response was obtained from thirty-one programs in 30 countries (in Canada: Ontario and Manitoba). Among the thirty countries, two did not have a screening program (Bulgaria, Romania), 4 had an opportunistic program (Austria, Greece, Iceland, Slovakia) and 24 had a CRCSP at regional or national level (**Fig. 1**). In six countries, the CRCSP was either in the pilot phase (Lithuania since 2020) or in experimentation in a randomized trial (Norway, Poland), or specific to a target population chosen on other socio-demographic criteria (USA) or in deployment from 2022 (Finland, Germany, Norway). In Norway, a national CRCSP program was started in 2022, following an experiment (FIT every 2 years between 50-74 vs. Flexible sigmoidoscopy once between 50-74) which had been ongoing since 2012[13]. In the USA, although the approach is mostly opportunistic, the existence of 8 population-based programs has been described [14]. Following the success of its demonstration program at 5 sites in pilot states (Maryland, New York, Missouri, Washington), the CDC is funding the implementation of other population-based projects targeting populations covered by Federally qualified health centers (FDHCs). Other FQHCs have participated in large research projects and programs (i.e. Sea Mar Community Health Centers and the PRECISE project [15], conducted at Kaiser Permanente Center for Health Research (KPCHR) and funded by the National Cancer Institute). Population registers were the main sources for identifying the target population. In thirteen out of 25 CRCSPs, the campaigns targeted people aged 50-74 and the FIT was the primary test (**Table 1**).

Of the 25 CRCSPs listed, seventeen carried out exclusions and 2 programs (Finland, Norway) planned to do so in 2022 (***Table-1***). For different reasons, six organized programs (Denmark, England, Germany, Lithuania, Luxembourg, Stockholm (Sweden)) did not refine their target populations by quantifiable exclusions. In England, the NHS Bowel Cancer Screening Program does not assume that a medical problem would exclude people from the program, other than if they have had a full removal of their bowel. This is then confirmed with the clinician to ensure that all the bowel has been removed (as people with any bowel remaining can still partake in the program). In Germany and Lithuania (pilot), the CRCSP protocols implemented in 2020 did not support any exclusion. In Luxembourg, the list of people to be invited was provided each month by the social security center. To exclude patients suffering from a pathology that could potentially influence the test positivity, the patient's referring doctor received the test result 2 days before the patient, which allows for the doctor to explain to the patient the probable reason for the positive test, to advise the patient on undertaking a colonoscopy and

whether to leave the CRCSP. In Stockholm, the endoscopy units had access to the computer system used by the program management structure (MS-CRCSP), they regularly recorded the cases of CRC or polyp diagnosed in a patient included in the program. These patients were thus automatically excluded from subsequent invitations. In Denmark, although the program targeted average risk people, no exclusion was made from the campaign invitation module. However, in the invitation letter, it was brought to the attention of people at high risk or at very high risk that it was not necessary to complete a screening test if they were already in a personalized follow-up program. In addition, using the unique identifier of the people, the number and the participation of these people who were not eligible for the CRCSP campaigns were regularly quantified [16]. In seven programs, almost all exclusions were made only after invitations, using data collected from patients and their treating physicians (***Table-1***).

Practice of medical exclusion in line with the aim of the study

In total, twenty programs were qualified as practicing (or potential practicing) medical exclusion. In fact, in addition to the seventeen cases of actual practice, the two programs (Finland, Norway) planning to practice medical exclusion have been added. Likewise, Denmark was added because their practice is like a medical exclusion in the sense of this study. Among the eleven exclusion criteria listed, only 5 (personal history of CRC, IBD, adenoma, recent colonoscopy/CT-colonography, genetic risk) were consensual in the twenty programs refining their target populations by exclusions, with some particularities (**Table 2**).

In Flanders (Belgium), Tuscany (Italy) and the Netherlands, a personal history of CRC did not lead to permanent exclusion. In Flanders, people who had been diagnosed with CRC in the last 10 years, those who had undergone opportunistic screening or a recent complete colonoscopy (<10y)/CT-colonography (<4 years) and those with a full colectomy were excluded from invitations using MADBs. In addition, during the invitation, a flyer was sent to people to inform them that participation was not recommended in case of symptoms or family or genetic risk of CRC. The criteria falling within the definition of average risk were applied in the programs only if there was proof of follow-up by colonoscopy (i.e., Italy) or only if the information was available in the main collection sources (i.e., Ontario). Whatever the program, anyone who received an invitation could be excluded if they or their attending physician made the request for any reason.

Use of medico-administrative databases and cancer registries in medical exclusion strategies

Pilot programs (Poland) and those with only a plan to achieve medical exclusion (Finland, Norway) were not included in the analysis of database use. The survey reveals that there was at least one MADB [16-40] or cancer registry [17, 20, 23, 38, 41-54], fully or partially covering each of the 17 states/countries/regions having a complete deployment of the CRCSP and practicing an exclusion. Claims-DBs only existed in ten countries/regions out of the seventeen and a regular connection with the CRCSP database was established in seven programs (***Table-3***). Two countries (Israel, Switzerland) were distinguished by the absence of national MADBs (**Extended data: Tab1**); the existing MADBs were those of insurance companies (i.e., CLALIT Database). In 2014, Washington State set up a Claims-DB (All-Payer Claims Database, WA-APCD) which was supplied by several other Claims-DB such as the bases of the Centers for Medicare & Medicaid service. Similarly, there was a cancer registry in all the countries, but it was regularly connected with the CRCSP database only in 8 provinces/regions/countries (***Table-3***).

In three countries (Italy, Denmark, Canada) the interconnection between screening bases and other bases was facilitated by the existence of a permanent personal health identification number (PHIN). In Croatia, the cancer screening register was regularly updated from taxpayer databases, with the name of the general practitioner and the health insurance number provided by the Croatian health Insurance Institute. There was not (in Croatia) a direct link between the screening database and the Claims-DB, but a daily transmission of cases of exclusion was made by the health insurance institute. In France, no link was established between the SNDS and the CRC screening databases. Each health insurance scheme (>10 in France) made available to each regional MS-CRCSP, each quarter, a target population list without any guarantee that an exclusion had been made upstream. Exclusions were made either after returning invitation letters by post, or using cancer registries which only covered a few departments, or as part of a partnership between certain MS-CRCSPs and the primary health insurance agency.

Most of the terminologies were MADB specific except for 3 that were used by ≥2 MADB: 1) The All-patient refined diagnosis related group (APR-DRG); 2) the WHO ICD in its various versions and modifications (i.e., 100% of cancer registries); 3) the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). The morbid situations attributable to the five consensual exclusion criteria did not have the same definition in the extracted phenotypes. In Catalonia, the CCR phenotype extracted from a cancer registry includes the ICD-10 codes C18-C21, while the C21 code was excluded elsewhere. In France, the IBD phenotypes include the codes M07.4, M07.5 which is not the case in Denmark or Portugal where the listed phenotypes only include the ICD-

10 codes K50 and K51 (***Table-3***). Literature review found only colonoscopy (or colonoscopy/sigmoidoscopy) phenotypes in Claims-DB. The phenotypes of the other four consensual criteria were accessible only in hospital morbidity databases with a rarity of genetic risk phenotypes.

The survey reveals the existence of two types of strategies for selecting the eligible population in the 17 CRCSPs completely deployed and performing exclusion (**Fig. 2**). In type-A (i.e., Australia, Belgium, Czech Republic, Israel, Ontario, Portugal, Slovenia, Washington-state) as summarized in a Portuguese study[55], the target population was either directly extracted from MADBs or from electronic health records (i.e. Washington) or linked with the MADB in order to extract all the medico-clinical characteristics allowing to qualify eligibility and quantify upstream of the invitations, the number of people to be excluded, the duration (definitive or temporary) of the exclusion from the program. In type B (i.e., Catalonia, Croatia, France, Italy, Manitoba, Netherlands, Switzerland), a list of target population was extracted from a source (MADB or not) and made available to the MS-CRCSP which should go through several means to to qualify eligibility and quantify the number of people to be excluded, the duration (permanent or temporary) of the exclusion, usually after invitations through the people invited or their attending physicians.

In the programs refining target population upstream of invitations using Claims-DB, in addition to the high exclusion rate, the target population coverage rate was estimated with precision. As summarized in the Flanders report (2019-2020), target population was 2,006,959 in 2019, the exclusion rate was 40.7% and the coverage rate was 64.6% (**Table 4**). In the cohorts of people invited to the 2019-2020/2021 campaign in 3 CRCSPs refining the eligible population usually after campaign invitations, the proportion of exclusions (all causes) was 4.4% in Slovenia and 4.8% in Catalonia. In France, where the CRCSP is officially evaluated at a national level, the data on exclusion are unavailable because the public health agency in charge of the CRCSP's evaluation only disseminates a global number of exclusions (medical and non-medical). To date, to have an exclusion count by criterion in France as reported in Slovenia and Catalonia, an extraction must be requested in the databases of the ninety-nine departments subject to regular evaluations. As an illustration, in the department of Isère, 344,973 people were invited to the 2019-2020 campaign. Among them, 45,820(13.3%) had been excluded (all causes of exclusion combined). The participation rate was estimated at 42% in Isère, while it was 33.8% in Haute-Savoie. These two French departments did not have the same means to refine their eligible populations (i.e., existence of a cancer registry in Isère), hence the significant difference in proportions.

196 **Discussion**

197 This study reveals a diversity of strategies and means of exclusion in CRCSPs. Although some programs target
198 higher or lower extremes, the average age range in the programs was as recommended by the EU commission [4,
199 56]. A few programs did not refine their target populations by exclusions. Most programs, especially in Europe,
200 used population registers as the primary sources for identification of the target population while carrying out
201 exclusions before or after campaign invitations. Despite the disparity in terms of number and types of exclusion
202 criteria, the study highlights the existence of five consensual criteria (personal history of CRC, personal history
203 of IBD, personal history of adenoma, recent colonoscopy, high risk genetic syndrome), which are applied by
204 more than 50% of programs. To carry out these exclusions, MADB, specially Claims-DB, were systematically
205 used in certain programs while others only used information provided by patients or their attending physicians to
206 refine their campaign target populations.

207 As recommended by scientific bodies in Europe [57] and USA [58], most CRCSP specifications recommend
208 targeting people at average risk of CRC. This notion of average risk was not considered to date because no
209 program applies all the exclusion criteria that fall within the definition of high risk. The number of consensual
210 criteria was certainly limited but consistent with the definition of CCR risk. The disparity between programs
211 revealed by this study, in terms of the number and type of exclusion criteria applied, argues the need for a
212 redefinition of the CRCSP's eligible population.

213 Although the previous surveys were carried out at a time when several programs were in the pilot phase [6, 7],
214 this study shows that the conclusions made are still valid, which implies that the recommendations for
215 improvement made by these authors have not been applied. Klabunde et al. [6] concluded that more meaningful
216 comparisons of CRC screening participation indicators across programs are possible if participation indicators
217 are calculated using consistent definitions and differences in program organization and population characteristics
218 are considered. For this standardization of definitions, our study suggests keeping only the consensual exclusion
219 criteria which are probably the least subtle among the eleven criteria listed.

220 Transient benign pathologies, other serious disease in particular those in the terminal phase, are certainly
221 morbid situations that can prevent the performance of a screening test but, it must be admitted that they are not
222 necessarily CRC risk factors, their inapplicability in many programs is therefore justified. Regarding the family
223 history of CRC, their subtlety is reinforced by the inaccessibility of patient's selection process data, already
224 mentioned by Wang et al. [9]. The inaccessibility of data was highlighted in this study by the Ontario program in

which application of the recommendations was obsolete due to a lack of information in the MADB used for exclusions: "*These populations were not eligible for CRC average-risk screening according to program recommendations; however, these information were not accessible via the administrative databases therefore the program does not exclude/alter the correspondence campaign for these populations. Participants who meet the definition of increased risk in Ontario are not eligible for average risk CRC screening; data regarding family history of CRC is not available through the administrative databases in Ontario therefore, exclusions to the correspondence campaign based on family history directly cannot be applied, however, people with a recent colonoscopy are excluded, indirectly excluding people who are being screened through the increased risk arm of the program*". However, people who should be excluded from campaigns because of a family history of CRC can be excluded indirectly because of the third consensual criteria (recent colonoscopy), provided that they are in a process of regular follow-up (high-risk people). This was the case in France where the associations of the ICD-10 code Z80.X and a code for total colonoscopy are fairly frequent in homogeneous groups of diseases in SNDS database.

This study contributes to the understanding of the processes for construction of the CRCSP's target populations. It provides an understanding of the exclusion practices in force in the CRCSPs while highlighting consensual criteria which, if applied by all, would improve the comparability of participation rates between programs. Indeed, by observing a variation (1-15%) in the proportion of ineligible people, Bulliard et al. [7] admitted that this led to differences in estimates of campaign participation rate. The comparison of exclusion data (2019-2020) shows that the number of exclusion criteria has no impact on the exclusion rate. We can deduct from these exclusion data that connecting the screening database with other databases (or their regular use) would perfect the exclusion rate.

In addition to its impact in perfecting the exclusion rate, the use of MADB allows reproducible filtering, which could guarantee a consensual method for defining the eligible population. Cancer registries are recognized as essential for an adequate evaluation of cancer screening programs, but they are not involved in the evaluation of screening in several European countries [59]. For these authors, the lack of involvement of cancer registries was a major obstacle to improving the effectiveness of European programs. About 8 years after this finding, there are still countries (i.e., France) not fully covered by a cancer registry and several programs without interconnections with existing cancer registries.

Prior to any use of MADBs as a standard exclusion means, standardization of the querying algorithms is necessary, particularly in the European Union area where there is a prospect of setting up a European health data space (EHDS). In its current form, the European Commission proposal does not stipulate specific standards that must be universally adopted to ensure semantic and syntactic interoperability[60]. The MADBs revealed in this study use various and non-interoperable terminologies for some, the definition of standards is a requirement for the migration of screening data in this EHDS.

Before using the MADBs, it will also be necessary to think about implementing new strategies to minimize the number of people wrongly included or not included wrongly in the campaigns because none of the MADB's the querying algorithms has an accuracy of 100% [17, 21, 26, 34]. In addition, there is a diversity of computational definitions of morbid situations justifying exclusion, this compromises any standardization of the definition of the person to be excluded from CRCSP campaigns.

Conclusion

The notion of average risk is partially considered in the CRCSPs because no program collects data on all the morbid situations justifying medical exclusion. In addition, the practice of exclusion was disparate and varied according to the means of exclusion available. Despite the disparity in terms of the number of exclusion criteria and the diversity in the choice of criteria by the programs practicing exclusion, the study reveals the existence of a limited number of consensual criteria, which seem to be the less subtle because easily collectable thanks to the current means available to program management structures. These consensual criteria can be queried in most of the MADB referenced in countries with a CRCSP, although the use of these types of health databases is not in force in several programs. However, with a view to standardizing evaluation indicators for better comparability of programs, the use of this type of database would allow reproducible filtering and a consensual method for defining the eligible population.

Declarations

Regulatory issues

All methods were carried out in accordance with relevant guidelines and regulations. According to the current French legislation, a study that does not change the care of patients does not require the opinion of the Clinical Research Ethics Committee. This article does not contain any studies with human participants performed by any

of the authors. This study does not involve human participants and informed consent was therefore not required. This article does not contain any studies with animals performed by any of the authors.

Consent for publication

Not Applicable

Availability of data and materials

Survey data is available upon request. Screening activity data provided by the coordinators should be requested from the program coordinators in the respective countries/regions.

Competing Interests

All authors declare no conflict of interest.

Funding Sources

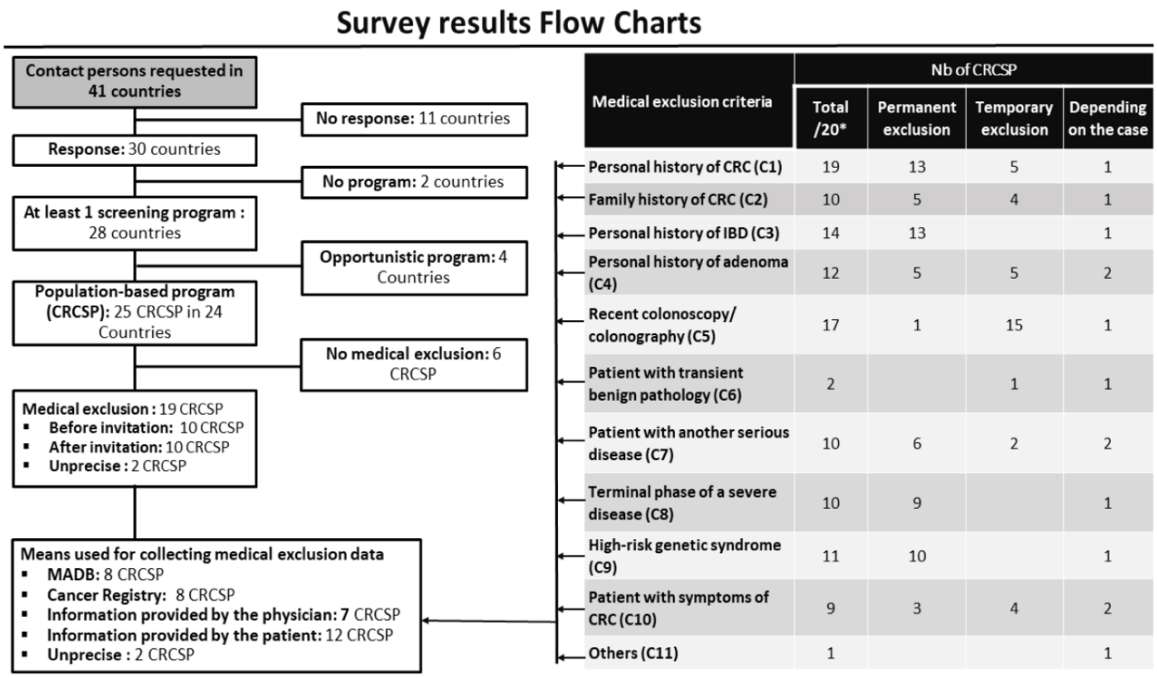
The study did not receive any special funding.

Authors' contributions

AK: *Conceptualization, Methodology, Data curation, Original draft preparation* ***CB, GB, SH and DNM****: Data curation, Writing Original draft preparation* ***RB****: Visualization, Supervision, Validation* ***CD****: Conceptualization, Methodology, Visualization, Supervision, Validation.* All the authors revised the manuscript, read, and approved the final version of this manuscript.

Acknowledgments

The authors are grateful to all collaborators for the quality of the data used in this study. The authors are grateful to Kabelo Leonard Mauco (Institute of Health Sciences, Botswana) and Taiwo Adedeji (University of Portsmouth, UK) for the thorough editing of the manuscript and to all the collaborators who took part in the final revisions.



Medical exclusion criteria	Nb of CRCSP			
	Total /20*	Permanent exclusion	Temporary exclusion	Depending on the case
Personal history of CRC (C1)	19	13	5	1
Family history of CRC (C2)	10	5	4	1
Personal history of IBD (C3)	14	13		1
Personal history of adenoma (C4)	12	5	5	2
Recent colonoscopy/ colonography (C5)	17	1	15	1
Patient with transient benign pathology (C6)	2		1	1
Patient with another serious disease (C7)	10	6	2	2
Terminal phase of a severe disease (C8)	10	9		1
High-risk genetic syndrome (C9)	11	10		1
Patient with symptoms of CRC (C10)	9	3	4	2
Others (C11)	1			1

Fig. 1: Practice of medical exclusion flow charts; * According to the results of the survey there is no practice of medical exclusion in Denmark (left: 19 programs). But in the sense of the study the practice is comparable to a medical exclusion (right: 20 programs).

307 **Table 1: Type of program by country (Situation in 2021 revealed in 2022).**

Country (Region or District)	Program type at the start of the program (Year of implementation of a screening program)	Year of implementation of the most recent screening protocol (program type according to this protocol)	Test used (Target age); time (in month) between two screening tests	Source of identification of the screening target population (in organised program)	Practice (YES/NO) of a medical exclusion in organised program (Means used if YES)	Practice of a medical exclusion in organised program: Timeline
Australia	Pop-based (2006)	2017 (Pop-based)	FIT (50-74); 24	B	YES (M2)	T1
Austria	Opportunistic (1980)	2005 (Opportunistic)	gFOBT/FIT (≥50); 12 Colo (≥50); 120	-	-	-
Belgium (Flanders)	Pop-based (2013)	2021 (Pop-based)	FIT (50-74); 24	A	YES (M1, M2, M3)	T1
Bulgaria	No program	No program	-	-	-	-
Canada (Ontario)	Pop-based (2008)	2015 (Pop-based)	FIT (50-74); 24	A	YES (M1, M2, M3)	T1, T2
Canada (Manitoba)	Pop-based (2008)	2015 (Pop-based)	FIT (50-74); 24	A	YES (M1, M3)	T1
Croatia	Pop-based (2008)	2013 (Pop-based)	FIT (50-74); 24	A, B	YES (M2, M3)	T2
Czech Republic	Pop-based (2000)	2020 (Pop-based)	FIT (50-54); 12 FIT (>54); 24 Colo (>54); 120	B	YES (M4)	T2
Denmark	Pop-based (2014)	2018 (Pop-based)	FIT (50-74); 24	A	NO	-
England	Pop-based (2006)	2021 (Pop-based)	FIT (56-74); 24	A, D	NO	-
Finland	Pop-based (2004) [a]	2022 (Pop-based) [a]	FIT (50-74); 24	D	YES (?)	?
France	Pop-based (2009)	2018 (Pop-based)	FIT (50-74); 24	B	YES (M1, M2, M3, M4) [b]	T1 [b], T2
Germany	Opportunistic (1974)	2020 (Pop-based) [a]	FIT (50-54); 12 FIT (>54); 24 Colo (Men 50-74); 120 Colo (Women 55-74); 120	B	NO	-
Greece	Opportunistic (?)	? (Opportunistic)	-	-	-	-
Iceland	Opportunistic (?)	?(Opportunistic)	Colo (>50);	-	-	-
Israel (CLALIT HMO) *	Pop-based (1995)	2005 (Pop-based)	FIT (50-74); 12	B	YES (M1, M2, M3, M4)	T1

Italy (Tuscany) †	Pop-based (2005)	2015 (Pop-based)	FIT (50-70); 24	A	YES (M3, M4)	T2
Lithuania	Pop-based (2009) [a]	2020 (Pop-based) [a]	FIT (50-74); 24	A	NO	-
Luxembourg	Pop-based (2016)	2021 (Pop-based)	FIT (50-74); 24	A, B	NO	-
Malta	Pop-based (2013)	2011 (Pop-based)	FIT (55-72); 24	A, C	YES (M3)	T2
Norway	Pop-based (2012) [a]	2022(Pop-based) [a]	FIT (55-65); 24	A	YES (M1)	T1
Poland	Opportunistic (2000)	2012 (Pop-based) [a]	Colo (55-64); once in the lifetime	A	YES (?)	?
Portugal (Northern) †	Pop-based (2008)	2018 (Pop-based)	FIT (50-74); 24	B	YES (M1, M2)	T1
Romania	No program	No program	-	-	-	-
Slovakia	Opportunistic	Opportunistic	-	-	-	-
Slovenia	Pop-based (2009)	2015 (Pop-based)	FIT (50-74); 24	A, B	YES (M2, M3)	T2
Spain (Catalonia)	Pop-based (2000)	2021 (Pop-based)	FIT (50-69); 24	A	YES (M1, M2, M3, M4)	T1, T2
Sweden (Stockholm)	Pop-based (2008)	2014 (Pop-based)	FIT (60-70); 24	A	NO	-
Switzerland (Vaud) †	Pop-based (2015)	2015 (Pop-based)	FIT (50-69); 24 Colo (50-69); 120	A	YES (M1, M3, M4)	T2
The Netherlands	Pop-based (2014)	2021 (Pop-based)	FIT (55-75); 48	A	YES (M3)	T2
USA (KPCHR Project) ‡	Pop-based (2014)	2018 (Pop-based)	FIT (50-75); 12 Colo (50-75); 120	A	YES (M4)	T1 [c]
No response	Cyprus, Estonia, Hungary, Ireland, Japan, Latvia, North Ireland, Scotland, Singapore, South Korea, Wales					

*Four official not-for-profit Health Maintenance Organizations (HMO) provide CRC screening in Israel, Clalit Health Services (CHS)is the largest HMO that insures about 50% of the population ; †In the country, none or another program than the one involved in this survey had been targeted by previous surveys; ‡Kaiser Permanente Center for Health Research(KPCHR) manages the PRECISE research project funded by the National Cancer Institute (2019-2023), that is taking place in a community health network in Washington of 30 clinics (Sea Mar Community Health Centers). All community clinics have an electronic health record (EHR), which is their medical record system and which is used to identify patients due for screening or follow-up; [a] Pilot program or randomized trial or recent program in the reorganization; [b] No structural framework but there are partnerships in some departments between the screening structure and the cancer registries or with the primary health insurance agency; [c] All community clinics have an electronic medical record system which is used in PRECISE project to identify patients due for screening or follow-up; **Colo:** Colonoscopy; **FIT:** Fecal immunochemical test; **FS:** Flexible sigmoidoscopy; **gFOBT**: Guaiac fecal occult blood test; **Source of identification of the screening target population**: (**A**): Population register, (**B**): Health insurance companies files, (**C**): Electoral lists, (**D**): Patient lists of general practitioners or primary health care centres, **MADB**: Medical administrative database; **Means used for medical exclusion**: **M1:** Exclusion made using a Cancer Registry, **M2**: Exclusion made using a MADB; **M3**: Exclusion made using information provided by the patient, **M4**: Exclusion made using information provided by the attending physician or primary health care facility; **Pop-based**: Population-based; **Timeline (Chronology) of medical exclusions** : **T1**: Before the campaign invitation using data from a cancer registry or using data from a MADB; **T2** : After the campaign invitation using information provided by the patient or by their treating physician or by a primary health care structure.

Table 2: Duration (in years) of the exclusion according to the exclusion criteria in force in the organized programs.

Country (Region or Program)	Number of medical exclusion criteria in the programs	Duration (in years) of the exclusion by medical exclusion criteria										
		Personal history of CRC (C1)	Family history of CRC (C2)	Personal history of IBD (C3)	Personal history of adenoma (C4)	Recent colonoscopy/ sigmoidoscopy/ CT colonography (C5)	Patient with transient benign pathology (C6)	Patient with another serious disease (C7)	Terminal phase of a severe disease (C8)	High-risk genetic syndrome (C9)	Patient with symptoms of CRC (C10)	Others (C11)
Australia	1					2						
Belgium (Flanders)	6	10	P [a]			4				P [a]	P [a]	P/T [b]
Canada (Ontario)	9	P	P[c]	P[c]	P[c]	10			P[c]	P[c]	P[c]	
Canada (Manitoba)	8	P	P	P	P	5				P	T	
Croatia	4	P			5	5		P [d]				
Czech Republic*	10	T/P	T/P	T/P	T/P	T/P	T/P	T/P	T/P	T/P	T/P	
Denmark	6	P	P	P	T	5				P		
Finland	2	10		P								
France	7	P	P	P	P	5/2 [e]				P	T	
Israel (CLALIT HMO)	3	P	P			T				P		
Italy (Tuscany) ‡	6	T		P	T	T			P	P		
Malta	6	5		P	2	2		T[f]	P			
Norway (Southeast)	1	P										
Poland	4	P						P	P		P	
Portugal (Northern)	10	P	P	P	P	10/5[g]	T	P	P	P	T/P	
Slovenia	4	P		P	P	3		T/P				
Spain (Catalonia)	9	P	P	P	T/P[h]	T		P	P	P	T	
Switzerland (Vaud)	7	P	P	P		9.5		P	P	P		
The Netherlands	7	T		P	T	T		T	P		T	
USA (KPCHR Project)	5	P		P		P		P	P			
Number of programs using the criterion / Total		**19/20**	**10/20**	**14/20**	**12/20**	**17/20**	**1/20**	**10/20**	**10/20**	**11 /20**	**9/20**	**1/20**

Responsible for the evaluation (Permanent, temporary, duration) of the criteria was the GP (alternatively primary care gynecologist) who evaluates the criteria during periodic check-up;* *‡Personal history of CRC****: Permanent exclusion if proof of specific endoscopic and/or oncological follow-up, Temporary exclusion of 5 years if the course of care over time is not known.* ***Personal history of IBD, Lynch disease and PAF:*** *Permanent exclusion if proof of endoscopic follow-up or appropriate medical follow-up re-enter screening program if no proof of follow-up.* ***Complete colonoscopy performed spontaneously outside the screening program****: 10 years in low-risk patients if negative examination or 1-2 tubular adenomas <10 mm with low-grade dysplasia, 5 years if no information is available on the risk of patient.* ***Virtual colonoscopy performed spontaneously outside the screening program****: 10 years if negative, 2 years if detection of polyps < 6 mm, without immediate colonoscopy;* [a] *The information is given at the time of the invitation, that any person concerned by the morbid situation is not concerned by the screening, they are free to participate or not;* [b] *Opportunistic screening(T);* [c] *These populations were not eligible for colorectal cancer screening according to program recommendations; however, these information were not accessible via the administrative databases therefore the program does not exclude/alter the correspondence campaign for these populations. ;* [d] *Exclusion if data available;* [e] *5 years if colonoscopy and 2 years if CT Colonography;* [f] *Patients undergoing chemotherapy or radiotherapy for other cancers are suspended for about 6 months after finishing their treatments;* [g] *10 years if colonoscopy and 5 years if sigmoidoscopy;* [h] *The duration of the temporary exclusion depends on the type, number and size of the adenomas.* **In C5:** *Any unique value corresponds to the duration (in year) of temporary exclusion after a colonoscopy****; P:*** *Permanent Exclusion****; T:*** *Temporary exclusion with duration in years not communicated;* ***IBD:*** *Inflammatory bowel Disease (Chron's disease or Ulcerative colitis).*

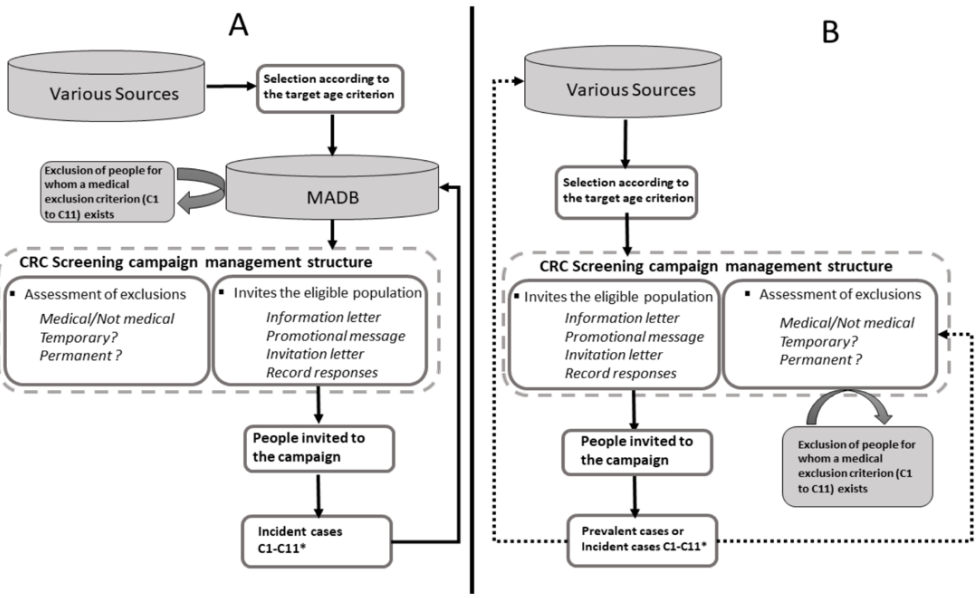
Table 3: Examples of phenotypes and algorithms defining the morbid situations which define the consensual criteria in indexed databases, connection between the screening database and these databases.

Country (Region or Program)	Claims-DB [A], example of Other-MADB [B] and example of cancer registries referenced and collecting data on consensual medical exclusion criteria [Database type]: Database name (Acronym)	Link between the Screening Database and other Databases	Terminologies of interest currently used for coding in the Database	Consensual medical exclusion criteria identified (Phenotype examples) [reference studies];
Australia	**[A]**: Medical benefits schedule (MBS)	Linked	MBS item codes	**Colo** (D_1) [17, 18];
	[B]: National hospital morbidity database (NHMD)	Not linked	ACHI, ARP-DRG, ICD-0-3, ICD-10, ICD-10-AM, SNOMED CT,	**CRC** (A_2 or A_8 or A_9 or A_{10}) [17, 19]; **IBD** (B_1 or B_3 or B_4) [19]; **Colo** (D_2 or D_3 or D_4) [19]; **Adenoma** (C_4 or C_5 or C_6) [19];
	[C]: Australian Cancer Database (ACD)	Not linked	ICD-10-AM	**CCR** (A_1) [17, 41];
Belgium (Flanders)	**[A]:** The Intermutualistic agency (IMA)	Linked†	Nomenclature IMA	**Colo** (D_5) [20];
	[B]: Minimum hospital data set (MHD-MZG-RHM)	Not Linked*	3BT, APR-DRG, LOINC, ICD-10-CM, ICD-10-PCS, SNOMED CT	?
	[C]: Belgian Cancer Registry	Linked	ICD-10-CM	**CCR** (A_1) [20];
Canada (Manitoba[a] & Ontario)	**[A]:** Ontario health insurance plan (OHIP) claims database	Linked	OHIP codes (billing & diagnostic),	**Colo** (D_6 or D_7) [21, 22];
	[B]: Hospital morbidity database (HMDB)	Not Linked	CCI, ICD-10-CA,	**CRC** (A_2 or A_8) [22, 23]; **IBD** (B_1 or B_3) [23, 24]; **Colo** (D_8) [22]; **Adenoma** (C_2) [25]
	[C]: Manitoba Cancer Registry and treatment	Linked	ICD-O-3	**CRC** (A_2) [42];
	[C]: Ontario Cancer Registry	Linked	ICD-9, ICD-10-CM	**CRC** (A_2) [23]; **CRC** (A_8) [22];
Croatia	**[A]:** Croatian primary health care database	Not Linked*	?	?
	[B]: Croatian hospital discharge Database	Not Linked*	APR-DRG,	?
	[C]: Croatian National Cancer registry	Not linked*	ICD-9, ICD-10-CM	**CCR** (A_{16}) OR (A_{17}) [43];
Czech Republic	**[A]:** Health insurance database	Linked	?	?
	[B]: National register of hospitalized patients (NRHOSP)	Linked	ICD-10	**CRC** (A_1) [31]; **IBD** (B_1) [30];
	[C]: Czech National Cancer Registry	Linked	ICD-10	**CCR** (A_1) [45]
Denmark [a]	**[B]:** Danish national patient registry (DNPR)	Linked	ICD-10, NOMESCO, SKS, SNOMED CT	**CRC** (A_1 or A_5 or A_{13}) [16, 26, 27]; **IBD** (B_1 or B_5) [16, 26, 27]; **Adenoma** (C_7) [16]; **Colo** (D_{12} or D_{13}) [29]; **High-risk** (E_1) [16];
	[C]: The Danish Cancer Registry	Linked	ICD-10	**CCR** (A_1) [44];
France	**[A]/[B]:** Système National des données de santé (SNDS)	Not linked**	ICD-10; CCAM	**CCR** (A_1) [31, 32]; **IBD** (B_1 or B_2) [31]; **Adenoma** (C_1) [32]; **Colo** (D_9) [33];
	Isere Cancer Registry	Not linked**	ICD-10	**CCR** (A_1) [47]
Israel [a]	**[B]:** CLALIT Health Services (CHS) database	Linked	ICD-9	**IBD** (B_3) [34][35]
	[C]: The Israel National Cancer Registry (INCR)	Linked	ICD-10	**CRC** (A_1+A_{11}) [47]
Italy [a]	**[B]:** National hospital discharge database	Not Linked	ICD-9	**CCR** (A_{14}) [35]; **IBD** (B_3) [36];
	[C]: Tuscany Region Tumor Registry	Not linked*	ICD-10	**CCR** (A_1) [48];
Malta [a]	**[B]:** National hospitals information system (NHIS)	Not Linked	ICD-10; ICD-9 CM (procedures);	?
	[C]: Malta National Cancer Registry	Not linked	ICD-O-2	**CCR** (A_1) [49];
Portugal (Northern)	**[A]:** National patient database	Linked	?	?
	[B]: National hospital morbidity database	Not Linked	APR-DRG, ICD-9-CM, ICD-10-CM, ICD-10-PCS	**IBD** (B_3) [37];
	[C]: Registo Oncológico Nacional (RON)‡	Linked	ICD-10	**CCR** (A_1) [50];
Slovenia	**[A]:** Health insurance institute data (ZZZS)	Linked	ICD-10-AM	
	[B]: Hospital discharge database	Not Linked	?	
	[C]: Cancer Registry of Republic of Slovenia	Linked	ICD-10	**CCR** (A_1) [51];
Spain (Catalonia) [b]	**[A]:** Primary health care system registry (SIDIAP)	Linked	ICD-10	**CRC** (A_1+A_4) [38];
	[B]: Spanish national hospital discharge database (RAE-CMBD)	Not Linked*	ICD-9-CM, ICD-10-PCS	**IBD** (B_1 or B_3) [39]; **Colo** (D_{10} or D_{11}) [39];
	[C]: Girona Cancer Registry	Not linked*	ICD10	**CCR** (A_{16}) [38];
Switzerland (Vaud) [c]	**[B]:** Swiss federal statistical office's database	Not Linked	CHOP, ICD-10-GM	?
	[C]: Vaud Cancer Registry	Not linked*	ICD10	**CCR** (A_1) [52];
The Netherlands	**[A]:** All-payer claims database (Vektis)	Not Linked*	?	?
	[B]: Dutch hospital discharge data	Not Linked	?	?
	Dutch Cancer Registry	Not linked*	ICD10	**CCR** (A_1) [53];
USA (Washington) [d]	[A]: All-payer Claims database (WA-APCD) ‡‡	Not Linked	CPT, HCPCS, ICD-9-CM, ICD-9-P	**CCR** (A_{15}) [40]; IBD (B_3) [40]; **Adenoma** (C_5) [40]; **Colo** (D_{14}, D_{15}, D_{16}) [40];
	[C]: Washington State Cancer Registry (WSCR)	Not linked	ICD10	**CCR** (A_2) [54];

Phenotypes

Phenotypes of CRC

A_1 (*ICD-10*: C18-C20); **A_2: (*ICD-10*:** C18.0, C18.2-C18.9, C19, C20); **A_3: (*ICD-10*:** D01.0, D01.1, D01.2); **A_4: (*ICD-10*:** C21, D01.3); **A_5: (*SNOMED CT:*** T6491, T65900, T65901, T65925, T65926, T660, T67 or T68 with morphology codes M8 or M9 with ≥3 in the fifth position (e.g., M8XXX3)); **A_6: (*ICD-9*:** 153, 154.0, 154.1); **A_7: (*ICD-9*:** 153.0-153.4, 153.6-153.9, 154.0, 154.1, 154.8); **A_8: (*ICD-9*:** 153.0-153.4, 153.6-153.9, 154.0, 154.1); **A_9: (*ICD-8*:** 153, 154.0, 154.1); **A_{10}: (*SNOMED CT*:** 126838000, 126845000); **A_{11}: (*ICD-10*:** C26.0); **A_{12}: (*ICD-9*:** 159.0); A_{12}: (***OHIP diagnostic***: 153, 154); **A_{13}: (*ICD-7:*** 153, 153.0, 153.4, 153.5, 154.9, 253.0, 253.1, 253.2, 253.3, 253.4, 453.0, 453.1, 453.2, 453.3, 453.4, 453.5, 453.8, 454.9, 853.0, 853.1, 853.2,



316

317 **Fig. 2: Executive summary of algorithms and schemes for selecting the eligible population in colorectal cancer screening programs in Type-A (i.e., Australia, Belgium, Catalonia in**
318 **Span, Czech Republic, Ontario in Canada, Portugal, Slovenia) and Type-B (i.e., Croatia, France, Italy, The Netherlands, Vaud in Swiss) programs. (C1 : Personal history of CRC , C2**
319 **: Family history of CRC, C3 : Personal history of IBD, C4 : Personal history of adenoma, C5 : Recent colonoscopy or sigmoidoscopy or CT colonography, C6 : Patient with**
320 **transient benign pathology, C7 : Patient with another serious disease, C8 : Terminal phase of a severe disease, C9 : High-risk genetic syndrome, C10 : Patient with symptoms of CRC,**
321 **C11 : Ohers criteria).**

322 **Table 4: Models for reporting program evaluation indicators according to the chronology of medical exclusions.**

323

Programs refining the eligible population before campaign invitation			**Programs (France, Catalonia, Slovenia) refining the eligible population mainly after campaign invitation**					
Indicators Reported	**Flanders (Belgium)**		***Indicators Reported***	**France***			***Slovenia***	**Catalonia (Spain)**
				National level	**Examples at departmental level**			
					Isère	***Haute-Savoie***		
Campaigns Year	**2019**	**2020**	**Campaigns Year**	2019-2020	2019-2020	2019-2020	2019 - 2020	2020-2021
Total [M] target population 50-74 years living in Flanders	**2,006,959 (100.0)**	**2,116,369 (100.0)**	**Total [N] target population 50-74**	**20,308,341**	369,470	236,941	674,637	500,000
Total [A] number of medical exclusions (% in M)	**817,094 (40.7)**	**976,543 (46.1)**	**Number [T] of people invited to the campaign (% in N)**	NA	344,973 (93.4)	180,624 (76.2)	614,439 (99.6)	474,716 (94.9)
Of which:			**Number [U] of exclusions for any cause (% in N)** †	2,828,436 (NA)	45,820 (12.4)	8173 (3.4)	27,011 (4.4)	22,625 (4.5)
Full colectomy: Total (% in A)	1,223 (0.1)	1,316 (0.1)	**Number [U] of exclusions for any cause (% in T)**	NA	45,820 (13.3)	8173 (4.5)	27,009 (4.4)	22,625 (4.8)
Full colonoscopy (only complete) < 10 years: Total (% in A)	484,919 (24.2)	520,397 (24.6)	**Number of medical exclusion (% in U)**	NA	23,241 (50.7)	4658 (57.0)	23,979 (88.8)	9,912 (43.8)
Virtual colonoscopy < 4 years: Total (% in A)	1,191 (0.1)	1,154 (0.1)	***Of which:***					
Diagnosed with CRC < 10 years: Total (% in A)	6,4385 (0.3)	6,5865 (0.3)	*Personal history of CRC (C1)*	NA	187	72	387	3,300
Screening outside the program (opportunistic) in past 2 years: Total (% in M)	42,675 (2.1)	41,573 (2.0)	*Family history of CRC (C2)*	NA	1496	598	CNA	1,558
Participation in screening program in year -1: Total (% in M)	280,648 (14.0)	405,517 (19.2)	*Personal history of IBD (C3)*	NA	240	448	406	993
Total target population after exclusions: Total (% in M)	**1,189,865** (59.3)	**1,139,826** (53.9)	*Personal history of adenoma (C4)*	NA	1871	739	5,293	1379
Invited in year X-1 (previous screening year)	341,638	412,929	*Recent colonoscopy or CT colonography (C5)*	NA	19,226	2694	17,714	2,641
Not willing to get invited (i.e., refusals)	12,869	12,732	*Patient with another serious disease (C7)*	NA	CNA	CNA	179	41

Group entitled to an invitation in year X [B]	**837,656 (100.0)**	**716,191 (100.0)**	*Terminal phase of a severe disease (C8)*	NA	CNA	CNA	CNA	16
Group effectively invited in year X: Total (% in B)	828,592 (98.9)	705,888 (98.6)	*High-risk genetic syndrome (C9)*	NA	231	107	CNA	CNA
Finally invited group [C]	**851,825 (100.0)**	**734,037 (100.0)**	*Patient with symptoms of CRC (C10)*	NA	CNA	CNA	CNA	CNA
Death or moved outside Flanders on moment of invitation	7,652	13,563	**Number of medical exclusions (% in T)[a]**	NA	23,241 (6.7)	4658 (2.6)	23,979 (3.9)	9,912 (2.1)
Eligible to invitation	844,173	720,474	**Number of tests completed (% in (T-U))[b]**	5,075,943	125,539 (42.0)	58354 (33.8)	355,169(60.1)	205,720(45.5)
Eligible to invitation and invited	828,596	705,887						
Invited although later an exclusion registered ‡	15,577	14,587						
Total coverage: Number (% in M)	**1,296,544 (64.6)**	**1,344,730 (63.5)**	**Total coverage: Number (% in N)**	NA	148,780 (40.3)	63,012 (26.6)	373901 (62.3)	222781 (44.5)
Coverage by medical exclusion: Number (% in M)	493,771 (24.6)	529,453 (25.0)	*Coverage by medical exclusion: Number (% in N)*	NA	23,241 (6.3)	4658 (2.0)	23,884 (4.0)	17,061 (3.4)
Coverage by screening in current year or year -1: Number (% in M)	721,056 (35.9)	745 680 (35.2)	*Coverage by screening in current campaign: Number (% in N)*	5,075,943(25.0)	125,539 (34.0)	58354 (24.6)	350,017(58.3)	205,720(41.1)
Coverage by opportunistic screening: Number (% in M)	68,140 (3.4)	59,692 (2.8)	*Coverage by opportunistic screening: Number (% in N)*	CNA	CNA	CNA	CNA	CNA
Coverage by colonoscopy: Number (% in M)	13,577 (0.7)	9,905 (0.5)	*Coverage by colonoscopy: Number (% in N)*	NA	NA	NA	?	?

***CNA**: Criteria not applicable; **NA:** Not available; **Two French department chosen: Isère has 3 partnerships (Cancer Registry, Primary Health Insurance Fund--CPAM-- and Medical Information Service of the Grenoble University Hospital) to optimize exclusion and Haute-Savoie, having no partnership, only makes its exclusions after the return of patients or their treating physicians; †The total exclusion for any cause is expressed by the sum of non-medical exclusions (refusal to participate, invitation letter not received due to incorrect address, moving, death) and exclusions for medical reasons (C1 to C11); [a] The exclusion rate is expressed by the ratio between the number of exclusions and the total number of invitations to the biannual campaign;. [b] The participation rate was expressed by the ratio between the number of tests carried out and the total number of invitations to the biannual campaign minus the total number of exclusions (medical + non-medical). ‡Administrative delay in dataflow, so on the moment of the exclusion, the screening managing structure did not know and an invitation was sent.*

324

.

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: a cancer journal for clinicians. 2021;71(3):209-49. doi: 10.3322/caac.21660. PMID: 33538338
2. Faivre J, Dancourt V, Lejeune C, Tazi MA, Lamour J, Gerard D, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. Gastroenterology. 2004;126(6):1674 - 80. doi: 10.1053/j.gastro.2004.02.018. PMID: 15609792
3. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet (London, England). 1996;348(9040):1472-7. doi: 10.1016/S0140-6736(96)03386-7. PMID: 8942775
4. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49. doi: 10.1136/gutjnl-2014-309086. PMID: 26041752
5. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA: a cancer journal for clinicians. 2009;59(6):366-78. Epub 2009/11/10. doi: 10.3322/caac.20038. PMID: 19897840
6. Klabunde C, Blom J, Bulliard JL, Garcia M, Hagoel L, Mai V, et al. Participation rates for organized colorectal cancer screening programmes: an international comparison. Journal of medical screening. 2015;22(3):119-26. doi: 10.1177/0969141315584694. PMID: 25967088
7. Bulliard JL, Garcia M, Blom J, Senore C, Mai V, Klabunde C. Sorting out measures and definitions of screening participation to improve comparability: the example of colorectal cancer. Eur J Cancer. 2014;50(2):434-46. doi: 10.1016/j.ejca.2013.09.015. PMID: 24144735
8. Leuraud K, Jezewski-Serra D, Viguier J, Salines E. Colorectal cancer screening by guaiac faecal occult blood test in France: Evaluation of the programme two years after launching. Cancer epidemiology. 2013;37(6):959-67. doi: 10.1016/j.canep.2013.07.008. PMID: 24035240
9. Wang A, Lee B, Patel S, Whitaker E, Issaka RB, Somsouk M. Selection of patients for large mailed fecal immunochemical test colorectal cancer screening outreach programs: A systematic review. Journal of medical screening. 2021;28(4):379-88. doi: 10.1177/0969141321997482. PMID: 33683155
10. de Klerk CM, Gupta S, Dekker E, Essink-Bot ML, Expert Working Group 'Coalition to reduce inequities in colorectal cancer screening' of the World Endoscopy O. Socioeconomic and ethnic inequities within organised colorectal cancer screening programmes worldwide. Gut. 2018;67(4):679-87. doi: 10.1136/gutjnl-2016-313311. PMID: 28073892
11. Vale DB, Anttila A, Ponti A, Senore C, Sankaranaryanan R, Ronco G, et al. Invitation strategies and coverage in the population-based cancer screening programmes in the European Union. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2019;28(2):131-40. doi: 10.1097/CEJ.0000000000000426. PMID: 29570103

12. Seitz JF, Lapalus D, Arlotto S, Gentile S, Ettori F, Rinaldi Y, et al. Colorectal cancer screening by fecal immunochemical test or colonoscopy in France: how many people are actually covered? Focus on the Provence-Alpes-Cote d'Azur region. European journal of gastroenterology & hepatology. 2022;34(4):405-10. doi: 10.1097/MEG.0000000000002338. PMID: 34882643
13. Randel KR, Schult AL, Botteri E, Hoff G, Bretthauer M, Ursin G, et al. Colorectal Cancer Screening With Repeated Fecal Immunochemical Test Versus Sigmoidoscopy: Baseline Results From a Randomized Trial. Gastroenterology. 2021;160(4):1085-96 e5. doi: 10.1053/j.gastro.2020.11.037. PMID: 33227280
14. Verma M, Sarfaty M, Brooks D, Wender RC. Population-based programs for increasing colorectal cancer screening in the United States. CA: a cancer journal for clinicians. 2015;65(6):497-510. doi: 10.3322/caac.21295. PMID: 26331705.
15. Coronado GD, Johnson ES, Leo MC, Schneider JL, Smith D, Mummadi R, et al. Patient randomized trial of a targeted navigation program to improve rates of follow-up colonoscopy in community health centers. Contemp Clin Trials. 2020;89:105920. doi: 10.1016/j.cct.2019.105920. PubMed PMID: 31881390
16. Larsen PT, Jorgensen SF, Njor SH. Colorectal cancer screening participation among citizens not recommended to be screened: a cohort study. BMC Gastroenterol. 2022;22(1):256. doi: 10.1186/s12876-022-02331-9. PMID: 35596148
17. Goldsbury D, Weber M, Yap S, Banks E, O'Connell DL, Canfell K. Identifying incident colorectal and lung cancer cases in health service utilisation databases in Australia: a validation study. BMC medical informatics and decision making. 2017;17(1):23. doi: 10.1186/s12911-017-0417-5. PMID: 28241763
18. Worthington J, He E, Lew JB, St John J, Horn C, Grogan P, et al. Colonoscopies in Australia - how much does the National Bowel Cancer Screening Program contribute to colonoscopy use? Public Health Res Pract. 2022. doi: 10.17061/phrp32232216. PMID: 36477980
19. Kelty E, Ward SV, Cadby G, McCarthy NS, O'Leary P, Moses EK, et al. Familial and non-familial risk factors associated with incidence of colorectal cancer in young and middle-aged persons in Western Australia. Cancer epidemiology. 2019;62:101591. doi: 10.1016/j.canep.2019.101591. PMID: 31494463
20. Macken E, Van Dongen S, De Brabander I, Francque S, Driessen A, Van Hal G. Post-colonoscopy colorectal cancer in Belgium: characteristics and influencing factors. Endosc Int Open. 2019;7(5):E717-E27. doi: 10.1055/a-0751-2660. PMID: 31073539
21. Tinmouth J, Sutradhar R, Liu N, Baxter NN, Paszat L, Rabeneck L. Validation of 5 key colonoscopy-related data elements from Ontario health administrative databases compared to the clinical record: a cross-sectional study. CMAJ Open. 2018;6(3):E330-E8. doi: 10.9778/cmajo.20180013. PubMed PMID: 30104417
22. Castelo M, Paszat L, Hansen BE, Scheer AS, Faught N, Nguyen L, et al. Measurement of clinical delay intervals among younger adults with colorectal cancer using health administrative data: a population-based analysis. BMJ Open Gastroenterol. 2022;9(1). doi: 10.1136/bmjgast-2022-001022. PMID: 36410773

391 23. Hol L, Sutradhar R, Gu S, Baxter NN, Rabeneck L, Tinmouth JM, et al. Repeat colonoscopy after a colonoscopy with a
392 negative result in Ontario: a population-based cohort study. CMAJ Open. 2015;3(2):E244-50. doi:
393 10.9778/cmajo.20140063. PMID: 26389103
394 24. Benchimol EI, Kuenzig ME, Bernstein CN, Nguyen GC, Guttmann A, Jones JL, et al. Rural and urban disparities in the
395 care of Canadian patients with inflammatory bowel disease: a population-based study. Clinical epidemiology.
396 2018;10:1613-26. doi: 10.2147/CLEP.S178056. PMID: 30519110
397 25. Paszat LF, Sutradhar R, Luo J, Baxter NN, Tinmouth J, Rabeneck L. Morbidity and mortality after major large bowel
398 resection of non-malignant polyp among participants in a population-based screening program. Journal of medical
399 screening. 2021;28(3):261-7. doi: 10.1177/0969141320967960. PMID: 33153368
400 26. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient
401 Registry: a review of content, data quality, and research potential. Clinical epidemiology. 2015;7:449-90. doi:
402 10.2147/CLEP.S91125. PubMed PMID: 26604824
403 27. Lash TL, Riis AH, Ostenfeld EB, Erichsen R, Vyberg M, Thorlacius-Ussing O. A validated algorithm to ascertain
404 colorectal cancer recurrence using registry resources in Denmark. Int J Cancer. 2015;136(9):2210-5. doi:
405 10.1002/ijc.29267. PMID: 25307704
406 28. Albaek Jacobsen H, Jess T, Larsen L. Validity of Inflammatory Bowel Disease Diagnoses in the Danish National
407 Patient Registry: A Population-Based Study from the North Denmark Region. Clinical epidemiology. 2022;14:1099-
408 109. doi: 10.2147/CLEP.S378003. PMID: 36226162
409 29. Kirkegard J, Farkas DK, Jorgensen JOL, Cronin-Fenton D. Hyper- and hypothyroidism and gastrointestinal cancer risk:
410 a Danish nationwide cohort study. Endocr Connect. 2018;7(11):1129-35. doi: 10.1530/EC-18-0258. PMID: 30352404
411 30. Czech Hospitalization yearbook, 2017. Institute of Health Information and Statistics of the Czech Republic, Prague.
412 ISSN 1210- 8731, (0862-5654) Available at: https://www.uzis.cz/index-en.php?pg=publications--library&id=238.
413 31. Melac AT, Lesuffleur T, Bousquet PJ, Fagot-Campagna A, Gastaldi-Menager C, Tuppin P. Cancer and end of life: the
414 management provided during the year and the month preceding death in 2015 and causes of death in France. Support
415 Care Cancer. 2020;28(8):3877-87. doi: 10.1007/s00520-019-05188-z. PMID: 31845006
416 32. Bousquet PJ, Lefeuvre D, Tuppin P, BenDiane MK, Rocchi M, Bouée-Benhamiche E, et al. Cancer care and public
417 health policy evaluations in France: Usefulness of the national cancer cohort. PLoS One. 2018;13(10):e0206448. Epub
418 2018/11/01. doi: 10.1371/journal.pone.0206448. PMID: 30379946
419 33. Laanani M, Coste J, Blotière PO, Carbonnel F, Weill A. Patient, Procedure, and Endoscopist Risk Factors for
420 Perforation, Bleeding, and Splenic Injury After Colonoscopies. Clin Gastroenterol Hepatol. 2019;17(4):719-27.e13.
421 Epub 2018/08/14. doi: 10.1016/j.cgh.2018.08.005. PMID: 30099110
422 34. Friedman MY, Leventer-Roberts M, Rosenblum J, Zigman N, Goren I, Mourad V, et al. Development and validation of
423 novel algorithms to identify patients with inflammatory bowel diseases in Israel: an epi-IIRN group study. Clinical
424 epidemiology. 2018;10:671-81. doi: 10.2147/CLEP.S151339. PMID: 29922093

35. Zorzi M, Gennaro N, Capodaglio G, Urso EDL, Pucciarelli S, Memo L, et al. Colorectal cancer screening: The surgery rates they are a-changing. A nationwide study on surgical resections in Italy. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2019;51(2):304-9. doi: 10.1016/j.dld.2018.10.011. PMID: 30449608
36. Meregaglia M, Banks H, Fattore G. Hospital Burden and Gastrointestinal Surgery in Inflammatory Bowel Disease Patients in Italy: A Retrospective Observational Study. J Crohns Colitis. 2015;9(10):853-62. doi: 10.1093/ecco-jcc/jjv104. PMID: 26069194
37. Santiago M, Magro F, Correia L, Portela F, Ministro P, Lago P, et al. Inflammatory Bowel Disease Reoperation Rate Has Decreased Over Time If Corrected by Prevalence. Clin Transl Gastroenterol. 2020;11(9):e00227. doi: 10.14309/ctg.0000000000000227. PMID: 33094955
38. Recalde M, Manzano-Salgado CB, Diaz Y, Puente D, Garcia-Gil MDM, Marcos-Gragera R, et al. Validation Of Cancer Diagnoses In Electronic Health Records: Results From The Information System For Research In Primary Care (SIDIAP) In Northeast Spain. Clinical epidemiology. 2019;11:1015-24. doi: 10.2147/CLEP.S225568. PMID: 31819655
39. Mendez-Bailon M, Jimenez-Garcia R, Munoz-Rivas N, Hernandez-Barrera V, de Miguel-Yanes JM, de Miguel-Diez J, et al. Trends and Clinical Impact of Gastrointestinal Endoscopic Procedures on Acute Heart Failure in Spain (2002-2017). J Clin Med. 2021;10(3). doi: 10.3390/jcm10030546. PMID: 33540753
40. Ko CW, Doria-Rose VP, Barrett MJ, Kamineni A, Enewold L, Weiss NS. Screening colonoscopy and flexible sigmoidoscopy for reduction of colorectal cancer incidence: A case-control study. PLoS One. 2019;14(12):e0226027. doi: 10.1371/journal.pone.0226027. PMID: 31805156
41. Yan MK, Adler NR, Heriot N, Shang C, Zalcberg JR, Evans S, et al. Opportunities and barriers for the use of Australian cancer registries as platforms for randomized clinical trials. Asia Pac J Clin Oncol. 2022;18(4):344-52. doi: 10.1111/ajco.13670. PMID: 34811922.
42. Decker KM, Lambert P, Nugent Z, Biswanger N, Samadder J, Singh H. Time Trends in the Diagnosis of Colorectal Cancer With Obstruction, Perforation, and Emergency Admission After the Introduction of Population-Based Organized Screening. JAMA Netw Open. 2020;3(5):e205741. doi: 10.1001/jamanetworkopen.2020.5741. PMID: 32453385
43. Kirac I, Sekerija M, Simunovic I, Zgaga L, Velimir Vrdoljak D, Kovacevic D, et al. Incidence and mortality trends of gastric and colorectal cancers in Croatia, 1988-2008. Croat Med J. 2012;53(2):124-34. doi: 10.3325/cmj.2012.53.124. PMID: 22522990
44. Christensen J, Hojsgaard Schmidt LK, Kejs AMT, Sogaard J, Rasted MC, Andersen O, et al. Agreement between the Danish Cancer Registry and the Danish Colorectal Cancer Group Database. Acta Oncol. 2020;59(1):116-23. doi: 10.1080/0284186X.2019.1669815. PMID: 31559881
45. Pehalova L, Krejci D, Snajdrova L, Dusek L. Cancer incidence trends in the Czech Republic. Cancer epidemiology. 2021;74:101975. doi: 10.1016/j.canep.2021.101975. PMID: 34247066

459 46. Tron L, Belot A, Fauvernier M, Remontet L, Bossard N, Launay L, et al. Socioeconomic environment and disparities in
460 cancer survival for 19 solid tumor sites: An analysis of the French Network of Cancer Registries (FRANCIM) data. Int
461 J Cancer. 2019;144(6):1262-74. doi: 10.1002/ijc.31951. PMID: 30367459
462 47. Magid A, Robinson E, Keinan-Boker L. Preventable Cancer in Israel. Int J Environ Res Public Health. 2022;19(17).
463 doi: 10.3390/ijerph191710521. PMID: 36078236
464 48. Giorgi Rossi P, Carretta E, Mangone L, Baracco S, Serraino D, Zorzi M, et al. Incidence of interval cancers in faecal
465 immunochemical test colorectal screening programmes in Italy. Journal of medical screening. 2018;25(1):32-9. doi:
466 10.1177/0969141316686391. PMID: 28195508
467 49. Gauci D, Allemani C, Woods L. Population-level cure of colorectal cancer in Malta: An analysis of patients diagnosed
468 between 1995 and 2004. Cancer epidemiology. 2016;42:32-8. doi: 10.1016/j.canep.2016.03.001. PMID: 26999401
469 50. Morais S, Antunes L, Rodrigues J, Fontes F, Bento MJ, Lunet N. The impact of the COVID-19 pandemic on the short-
470 term survival of patients with cancer in Northern Portugal. Int J Cancer. 2021;149(2):287-96. doi: 10.1002/ijc.33532.
471 PMID: 33634852
472 51. Boltezar L, Novakovic BJ, Moltara ME. Trends in specialized palliative care referrals at an oncology center from 2007
473 to 2019. BMC Palliat Care. 2021;20(1):135. doi: 10.1186/s12904-021-00828-w. PMID: 34479512
474 52. Feller A, Matthes KL, Bordoni A, Bouchardy C, Bulliard JL, Herrmann C, et al. The relative risk of second primary
475 cancers in Switzerland: a population-based retrospective cohort study. BMC cancer. 2020;20(1):51. doi:
476 10.1186/s12885-019-6452-0. PMID: 31964352
477 53. Schulpen M, van den Brandt PA. Mediterranean diet adherence and risk of colorectal cancer: the prospective
478 Netherlands Cohort Study. European journal of epidemiology. 2020;35(1):25-35. doi: 10.1007/s10654-019-00549-8.
479 PMID: 31494792
480 54. Orlich MJ, Singh PN, Sabate J, Fan J, Sveen L, Bennett H, et al. Vegetarian dietary patterns and the risk of colorectal
481 cancers. JAMA Intern Med. 2015;175(5):767-76. doi: 10.1001/jamainternmed.2015.59. PMID: 25751512
482 55. Monteiro H, Tavares F, Reis J, Ferreira G, Campos MJ, Costa S, et al. Colorectal Screening Program in Northern
483 Portugal: First Findings. Acta Med Port. 2022;35(3):164-9. doi: 10.20344/amp.15904. PMID: 34283017
484 56. Altobelli E, Lattanzi A, Paduano R, Varassi G, di Orio F. Colorectal cancer prevention in Europe: burden of disease and
485 status of screening programs. Preventive medicine. 2014;62:132-41. doi: 10.1016/j.ypmed.2014.02.010. PMID:
486 24530610
487 57. Valori R, Rey JF, Atkin WS, Bretthauer M, Senore C, Hoff G, et al. European guidelines for quality assurance in
488 colorectal cancer screening and diagnosis. First Edition--Quality assurance in endoscopy in colorectal cancer screening
489 and diagnosis. Endoscopy. 2012;44 Suppl 3:SE88-105. doi: 10.1055/s-0032-1309795. PMID: 23012124
490 58. Force USPST, Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, et al. Screening for Colorectal
491 Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2021;325(19):1965-77. doi:
492 10.1001/jama.2021.6238. PMID: 34003218

59. Anttila A, Lonnberg S, Ponti A, Suonio E, Villain P, Coebergh JW, et al. Towards better implementation of cancer screening in Europe through improved monitoring and evaluation and greater engagement of cancer registries. Eur J Cancer. 2015;51(2):241-51. doi: 10.1016/j.ejca.2014.10.022. PMID: 25483785

60. Stellmach C, Muzoora MR, Thun S. Digitalization of Health Data: Interoperability of the Proposed European Health Data Space. Studies in health technology and informatics. 2022;298:132-6. doi: 10.3233/SHTI220922. PMID: 36073471

Medical Exclusion from CRC Screening Program Project Collaborators (MECSP_Co) :

Jie-Bin Lew[1], Eleonora Feletto[1], Joachim Worthington[1], Emily He[1], Mark Short[2], Gerald Haidinger[3], Roselien Pas[4], Bronwen McCurdy[5], Christine Stogios[5], Sydney Liang[5], Health Educators[6], Nataša Antoljak[7], Ondřej Májek[8], Ondřej Ngo[8], Mette Bach Larsen[9], Mette Kielsholm Thomsen[10], Karen Emery-Downing[11], Maija Jäntti[12], Ágúst Ingi Ágústsson[13], Nocon Marc[14] , Dana Frost[15], Gad Rennert[16], Carlo Senore[17], Paola Mantellini[17], Audrius Dulskas[18], Tomas Poškus[18], Johan Boucher[19], Sylvia Camilleri[20], Bartlomiej Krzeczewski[21], Ricardo Araújo Cardoso[22], Anna Lisa Schult[23], Paula Berstad[23], Isabel Portillo Villares[24], Montse Garcia[25], Stina Fuentes[26], Jean-Luc Bulliard[27], Cyril Ducros[28], Monique van Wieren[29], Petrik, Amanda F[30]

[1]**Australia:** The Daffodil Centre, The University of Sydney, a joint-venture with Cancer Council New South Wales. [2]**Australia:** Australian Cancer Database, Australian Institute of Health and Welfare, Bruce. [3]**Austria**: Department for Social- and Preventive Medicine Centre for Public Health, Medical University of Vienna Kinderspitalgasse 15, 1090 Vienna. [4]**Belgium:** Belgian Cancer Registry, Brussel. [5]**Canada:** ColonCancerCheck, GI Endoscopy Program & Ontario Cervical Screening Program, Toronto. [6]**Canada** : CancerCare Manitoba Prevention & Screening, Winnipeg MB. [7]**Croatia:** Croatian Institute of Public Health, Zagreb. [8]**Czech Republic:** Institute of Biostatistics and Analyses, Praha. [9]**Denmark:** University Research Clinic for Cancer Screening, Department of Public Health Programmes, Randers Regional Hospital, Skovlyvej 15, DK-8930 Randers NO. [10]**Denmark**: Aarhus University Hospital, Department of Clinical Epidemiology, Aarhus N. [11]**England:** National Programme Manager Bowel Cancer Screening, **NHS**. [12]**Finland**: Finnish Cancer Registry, Helsinki. [13]**Iceland** : Icelandic Cancer Screening Coordination Center. [14]**Germany:** Federal Joint Committee (Healthcare), Medical Consultancy Department, Gutenbergstr. 13, 10587 Berlin. [15]**Israël:** Israel Cancer Association, Information and Health Promotion Department, Tel Aviv. [16]**Israel:** Department of Community Medicine and Epidemiology, Carmel Medical Center and B. Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa. [17]**Italy**: Istituto per lo Studio, la Prevenzione e la Rete Oncologica – ISPRO-Firenze, Firenze. [18]**Lituanie**: Vilnius University Hospital Santaros clinics, Santariškių. [19]**Luxembourg**: Ministère de la Santé, Direction de la santé, Coordination PDOCCR. [20]**Malta**: National Screening Programme, Health - Primary HealthCare, Valletta. [21]**Poland:** Department of Corporate Finance, Faculty of Economics and Sociology University of Lodz. [22]**Portugal:** Department of Gastroenterology, Centro Hospitalar Tondela-Viseu, E.P.E., Viseu. [23]**Norway**: Cancer Registry of Norway, Oslo. [24]**Spain**: Programas de Detección del Cáncer Colorrectal y Cribado Prenatal Osakidetza, Bilbao. [25]**Spain:** Catalan Institute of Oncology, Prevention and Control Programme, Cancer Screening Unit, L'Hospitalet de Llobregat, Barcelona. [26]**Sweden:** Medical screening developer, Stockholm Gotland. [27]**Switzerland:** Centre for primary care and public health, Sector Epidemiology and Data Science, University of Lausanne, Lausanne. [28]**Switzerland**: Programmes vaudois de dépistage du cancer, Lausanne. [29]**The Netherlands**: Rijksinstituut voor Volksgezondheid en Milieu, Centrum voor Bevolkingsonderzoek, Bilthoven. [30]**USA**: Kaiser Permanente Center for Health Research, Portland, Oregon.

Extended data Form 1: Colorectal cancer screening Qualitative Data Collection form

General instruction

The questionnaire relates to the program in progress in 2021 (pilot program or redeployment completed). Thanks for your support and collaboration to this project. The instructions and definitions for answering each of the questions are those made available to you during a previous survey organized by the IARC (see the IARC questionnaire attached). if you still require assistance in filling any of the form, please contact us by email at xxxxxxxxxx@xxxxxxxxxx.fr

1 - Organization of screening

1.1 Country : _____

1.2 Reporting for: ☐ National ☐Sub-national

1.3 If not reporting for entire country, please give the name of the geographic area(s): _____
_____

1.4 Please provide name of team/institution responsible for management/coordination of the colorectal cancer screening activities : _____

1.5 Please provide year screening programme was initiated(9999 if no programme): _____

2 - Screening protocol

2.1 Is there a screening protocol or guideline? ☐No ☐Yes

2.2 If yes, year when the current protocol was developed/last updated (9999 if no documented protocol): _____

2.3 Please provide (table below) information on the primary screening test :

Primary screening test	Target age (00 to 00)	Screening interval (months)
FIT		
gFOBT		
Stool DNA		
Flexible sigmoidoscopy		
Colonoscopy		
Virtual colonoscopy		
Other		

3 - Information system and data collection

3.1 Is there a computerized information system that collects screening-related data on individual basis?

☐No ☐Yes

3.1.1 If yes, for which purpose (multiple responses)?

☐ Identification of eligible population ☐ Screening participation ☐ Screening test results ☐ Further assessment ☐ Final pathology diagnosis ☐ Cancer staging ☐ CRC treatment

3.1.2 If yes, at which level does this computerized information system exist?

☐National ☐Sub-national ☐Both national and sub-national

3.2 Is there a system (computerized or paper-based) that gathers aggregated data on screening activities?

☐No ☐Yes

3.2.1 If yes, this system collects aggregated data on? (multiple responses)

☐Screened individuals ☐Screening test results ☐Further assessment ☐Final pathology diagnosis ☐Cancer staging ☐Cancer treatment

3.2.2 If yes, at which level does this computerized or paper-based information system exist?

☐National ☐Sub-national ☐Both national and sub-national

3.3 Is there a system to collect information on screening outside the programme (opportunistic screening/private sector)?

☐No ☐Yes ☐Cannot differentiate

3.4 Are screening data linked with population-based cancer registries (PBCR)?

☐Yes ☐PBCR exists but is not linked ☐No PBCR

3.5 Please provide if exists regional or national cancer registry name or internet link _____

3.6 Are screening data linked with the regional or national health *insurance claims database (NHICDB)* ?

☐Yes (1) ☐NHICDB exists but is not linked(2) ☐No NHICDB

3.7 Please provide if exists regional or national health insurance name or internet link _____

4. Eligible population and Invitations for screening

4.1 What is the current definition of your eligible population?

☐Anyone in the target age group without any exclusion criteria (1), provide age group.....

☐Anyone in the target age group with at less one exclusion criteria (2) , provide age group.....

4.1.1 If target age group with at less one exclusion criteria (2), please provide (table below) the exclusion criteria applicable to your current definition of eligible population (multiple responses)

Exclusions criteria from screening campaign	**Put X if yes**	**Temporary(T) or Permanent (P) exclusion** ***Please provide the exclusion delay (in year) if temporary***
Personal history of CRC		
Family history of CRC		
History of inflammatory bowel or Chron's disease		
Known adenomas		
Recent colonoscopy		
Patients with transient benign pathology		
Patients with serious illness/disability that limit participation in screening		
Patients with terminal illness or severe disabling condition		
High-risk genetic syndrome		
CRC symptoms		
Others		

4.1.2 If (2), what means are available to optimize this exclusion? (multiple responses)

☐Exclusion made from the PBCR ☐Exclusion made from the NHICDB

☐Exclusion based on information provided by the patients themselves

☐Exclusion based on information provided by GP or primary health centers (PHC)

4.1.3 When does this exclusion take place?

☐Before sending campaign invitations, using data from the cancer registry

☐Before sending campaign invitations, using data from the the NHICDB

☐Response from invitees or GP or PHC, after sending campaign invitations,

4.2 Is there a system to send individual invitations to the eligible population?

☐No ☐Yes

4.2.1 If yes, from which source are the eligible individuals identified? (multiple responses)

☐Population register ☐List of an insurance company ☐Electoral roll
☐List form GP or PHC ☐Other_____

4.2.2 If yes, what is the method of inviting the eligible individuals? (multiple responses)

☐Letter ☐Email ☐SMS ☐Phone calls ☐Home visits by health workers ☐Other_____

4.2.3 If yes, does the invitation include screening kit? ☐ No ☐Yes

4.2.4 If no system to send individual invitations to the eligible population, how do you call people to participate in a campaign? Free comment: _____

_____

4.3 Is there a system to invite specifi or selected populations only? (multiple responses)

☐Not screened in the round ☐High risk populations only (criteria for high risk: _____) ☐ other: _____

Conclusion

Please provide any feedback you find useful about your colorectal cancer screening program: _____

_____

Would you agree to have your name and address listed in our investigation report? ☐No ☐Yes

If yes, please write it here:_____

Extended data Form 2: Colorectal cancer screening Database information Collection form

like some information on the data retention period in the databases:

1- What is the legal data retention period for health data collected in the database
"NAME"
……
……
……

2 - After this legal data retention period, are the data destroyed or archived?

……
……
……

3 - If the data is archived, are the archives permanently available for research?

……
……
……

4- To access these databases, how should the actors of other national health programs such as the National colorectal Cancer Screening Program proceed?
……
……
……

5 - Could I have some references of studies carried out on these databases? In particular studies validating search algorithms for colorectal cancer in databases?
PubMed ID : ……, ……., ……..,
Others : ……

6 - Which version (8,9 or 10) of the WHO International Classification of Diseases do you use ?
……
Apart from this international classification, do you use other care/procedure classification codes ? Such as SNOMED, CPT, HCPCS or others?
……
……
……

Extended data Form 3: Colorectal cancer screening: Campaign data Collection form

Dear …,
Our survey in which you participated is almost at the end. Before showing you the results, we need to deepen the discussion which requires us to ask you for the umpteenth time, we are very sorry about that!
Our survey reveals that 10 CRC screening programs/countries (including yours) refine their target population for screening by exclusions before or after invitation to the campaign. We plan to make a comparison of exclusion figures according to exclusion strategies in these programs, would it be possible to include your data in this comparison?
I put in the table (see Attachment) the example of two French programs to better illustrate. Not necessarily the 2019-2020 biannual campaign, you can choose another biannual campaign in example 2018-2018 or 2020-2021.
The comparison will relate to the biannual campaigns exclusion rate and will include: the number of people invited to this biannual campaign, the number of exclusions made according to each of the criteria listed in point 4.1.1 of the questionnaire, the number of people who made a primary screening test.
Kind regards,
Interpretation of the French results:

In the ***Isère*** department, 344,973 people were invited to the 2019-2020 biannual campaign. Among them, 45,820 or 13.3% had been excluded for all causes of exclusion combined. 187 were excluded from the invited population because they had declared to have a personal history of cancer. 19,226 people were excluded because they had a colonoscopy less than 5 years old. the participation rate was estimated at 42% in Isère, while it was 33.8% in the Haute-Savoie department. These two departments of metropolitan France do not have the same means to perfect medical exclusions, hence the significant difference in rates and proportions.

Country or Programs	Number of people invited to the biannual campaign (Proportion of exclusion for any cause in %) †	Total exclusion for any cause (Proportion of medical exclusions in %)	Number (n) of people excluded from the campaign by current medical exclusion criteria									Total number of medical exclusions (Exclusion rate in %) **	Nb of tests completed (Biannual campaign participation rate in %) ***
			C1	C2	C3	C4	C5	C7	C8	C9	C10		
France*													
Isère	344,973 (13.3)	45,820(50.7)	187	1496	240	1871	19,226	NA	NA	231	NA	23,241 (6.7)	125539 (42.0)
Haute-Savoie	180,624 (4.5)	8173 (57.0)	72	598	448	739	2694	NA	NA	107	NA	4658 (2.6)	58354 (33.8)
Australia													
…													

Exclusion criteria: ***C1****: Personal history of CRC,* ***C2:*** *Family history of CRC,* ***C3:*** *History of chronic inflammatory bowel or Chron's disease -IBD-,* ***C4:*** *Personal history of adenoma,* ***C5:*** *Recent colonoscopy/sigmoidoscopy/CT colonoscopy ,* ***C7:*** *Patient with another serious disease that may limit participation in screening,* ***C8:*** *Patient in the terminal phase of a severe disease ,* ***C9:*** *High-risk genetic syndrome,* ***C10:*** *Patient with symptoms of CRC,* ***NA****: Not applicable;* † *The total exclusion for any cause is expressed by the sum of non-medical exclusions (refusal to participate, invitation letter not received due to incorrect address, moving, death) and exclusions for medical reasons (C1 to C10);* *******Two French department chosen: Isère has 3 partnerships (Cancer Registry, Primary Health Insurance Fund--CPAM-- and Medical Information Service of the Grenoble University Hospital) to optimize exclusion and Haute-Savoie, having no partnership, only makes its exclusions after the return of patients or their treating physicians.*
********The exclusion rate is expressed by the ratio between the number of exclusions and the total number of invitations to the biannual campaign. *********The participation rate was expressed by the ratio between the number of tests carried out and the total number of invitations to the biannual campaign minus the total number of exclusions (medical + non-medical).*

35

Extended data Form: Table 1: Referenced databases collecting information on morbid situations which define the consensual criteria for medical exclusion.

Country (Region or Program)	Claims-DB [A] and example of Other-MADB [B] referenced and collecting data on consensual medical exclusion criteria		Cancer registries referenced and collecting colorectal cancer data		
	[Database type]: Database name (Acronym)	Geographical Coverage (Initiation date, Data Retention period)	Number of region or country's cancer registries (Geographical Coverage, National coverage rate in %)	Example of Cancer Registry Name (Acronym)	Cancer Registry initiation date (Data Retention period)
Australia	**[A]**: Medical benefits schedule (MBS) **[B]:** National hospital morbidity database (NHMD)	National (1984, Permanent) National (1994, Permanent)	9 (National, 99%) [a]	Australian Cancer Database (ACD)	1982 (Permanent)
Belgium (Flanders)	**[A]:** The Intermutualistic agency (IMA) **[B]:** Minimum hospital data set (MHD-MZG-RHM)	National (2002, Permanent) National (2007, Permanent)	1(National, 100%)	Belgian Cancer Registry	2004 (Permanent since 2004)
Canada (Manitoba[a] & Ontario)	**[A]:** Ontario health insurance plan (OHIP) claims database **[B]:** Hospital morbidity database (HMDB)	Provincial (1991, Permanent) National (1994, Permanent)	1 (Province, 100%) 1 (Province, 100%)	Ontario Cancer Registry Manitoba Cancer Registry and treatment	1964 (Permanent) 1937(Permanent since 1984)
Croatia	**[A]:** Croatian primary health care database **[B]:** Croatian hospital discharge Database	National (?) National (1991, Permanent from 2001)	1 (National, 100%)	Croatian National Cancer registry	1959 (Permanent since 1988)
Denmark [a]	**[B]:** Danish national patient registry (DNPR)	National (1977, Permanent)	2 (National, 100%)	The Danish Cancer Registry	1943 (Permanent since 1943)
Czech Republic	**[A]:** Health insurance database **[B]:** National register of hospitalized patients (NRHOSP)	? National (1991, Permanent from 1981)	1 (National, 100%)	Czech National Cancer Registry	1950 (Permanent since 1988)
France	**[A]/[B]:** Système National des données de santé (SNDS)	National (2016, 19y (retention) & 10y (archive))	31 (Department, 20%) [b]	Isere Cancer Registry	1978 (?)
Israel [a]	**[B]:** CLALIT Health Services (CHS) database	?	1 (National, 100%)	The Israel National Cancer Registry (INCR)	1960 (Permanent since 1982)
Italy [a]	**[B]:** National hospital discharge database	National (1991, Permanent from 2001)	1 (Region, 100%) [g]	Tuscany Region Tumor	1984 (Permanent

36

				Registry	since 1985)
Malta [a]	**[B]:** National hospitals information system (NHIS)	?	1(National, 100%)	Malta National Cancer Registry	1985(Permanent since 1995)
Portugal (Northern)	**[A]:** National patient database **[B]:** National hospital morbidity database	National (?) ?	1(National, 100%)	Registo Oncológico Nacional (RON)‡	2018 (?)
Slovenia	**[A]:** Health insurance institute data (ZZZS) **[B]:** Hospital discharge database	National (?) National (2013, Permanent from 1997)	1 (National)	Cancer Registry of Republic of Slovenia	1950 (Permanent since 1961)
Spain (Catalonia) [b]	**[B]:** Primary health care system registry (SIDIAP) **[B]:** Spanish national hospital discharge database (RAE-CMBD)	National (2016, Permanent from 1997)	2 (Province, 20%) [h]	Girona Cancer Registry Tarragona Cancer Registry	1994 (Permanent) 1980 (Permanent)
Switzerland (Vaud) [c]	**[B]:** Swiss federal statistical office's database	?	23 (Cantonal, 100%) [g]	Vaud Cancer Registry	1974(?)
The Netherlands	**[A]:** All-payer claims database (Vektis) **[B]:** Dutch hospital discharge data	National (2006, Permanent) ?	1 (National, 100%)	Dutch Cancer Registry	1989 (?)
USA (Washington) [d]	[A]: All-payer Claims database (WA-APCD) †‡	75% State-wide (2014, Permanent)	1(Statewide, 100%)	Washington State Cancer Registry (WSCR)	1995 (Permanent)

†The link is established through the Belgian Cancer Registry; *No automated connection is established but it is possible to do if necessary; [a] No Claims-DB in the country according to survey; [b] No Claims-DB in Catalonia, to obtain the screening target population, the most comprehensive population register of the Catalan Health Service (Central register of insured persons: Primary Health Care System Registry) is used; [c] In Switzerland, each insurance company has its own claims-DB; [d] WA-APCD database is supplied by several other Claims-DB such as the Centers for Medicare & Medicaid service databases; [e] Australia has six state and two territory cancer registries. Legislation in each jurisdiction requires cancer data to be reported to each jurisdiction's respective cancer registry. These data are annually supplied by each jurisdiction to the Australian Institute of Health and Welfare, which produces the Australian Cancer Database [42]; [f] Only 20% of the national territory is covered by the 31 listed cancer registries (all cancer sites combined); [g] In the country, there is at least one cancer registry per region or canton; [h] The two cancer registries listed only cover 20% of the region; ‡On January 2018, by law, a new stage for epidemiology and cancer registration began in Portugal, the four regional cancer registries (RORENO, RORCentro and ROR-Sul) became only one, national and global – RON;

683

684 **Methods**

685 ***Study design***

686 The descriptive study carried out (i) a cross-sectional survey and (ii) a non-systematic review of the literature. To
687 describe the strategies used by CRCSPs management structures (MS-CRCSP) to consider exclusion criteria, a
688 cross sectional survey was conducted. The survey included, all the programs of countries in the WHO Europe
689 zone (opportunistic or population-based, pilot, complete or incomplete rollout) which were included in previous
690 surveys [4, 11]. The study included non-EU programs that were population-based with complete rollout [4].
691 Once a MADB was used in the strategies, the study investigated whether the definition of selection criteria was
692 published in the literature in terms of lists of codes from medical terminologies. A non-systematic review of the
693 literature was conducted to compile these published definitions (phenotypes) of exclusion criteria to be applied
694 on MADB and cancer registries, in the countries having a CRCSP in 2021 and practicing exclusion. The
695 description of these selection criteria was compared between programs to show the degree of consensus.

696 ***Conduct of the survey***

697 The survey was carried out in three main steps between February 2022 and August 2022. The first step of the
698 survey focused on the practice of exclusion in the programs. The standardized form used for the regular
699 monitoring of screening programs [11] was readapted and used, with the agreement of the screening service of
700 the International Agency for Research on Cancer (IARC). A first email was sent to potential studies participants,
701 requesting them to take part in the survey. This was sent to contacts who had taken part in the previous surveys
702 identified and to the authors of recent articles (>2010) describing or evaluating the screening campaigns in the
703 various programs listed by IARC.

704 A follow-up email included a form (***Extended data: Form 1***) for collecting information related to: i) the
705 screening approach (opportunistic or population-based) in progress in 2021, with reference to the CRCSP's
706 definition [4]; ii) the exclusion criteria in force in 2021: in accordance with the criteria listed in the literature [6,
707 7], a list of 11 potential medical exclusion criteria (1-Personal history of CRC, 2-Family history of CRC, 3-
708 Personal history of IBD, 4-Personal history of adenoma, 5-Recent colonoscopy/sigmoidoscopy/CT-
709 colonography, 6-Patient with transient benign pathology, 7-Patient with another serious disease, 8-Terminal
710 phase of a severe disease, 9-High-risk genetic syndrome, 10-Patient with symptoms of CRC, 11-Ohers criteria)

711 had been proposed; iii) the contribution of Claims-DB and cancer registries in exclusion. At this stage, the
712 survey also collected data from the articles or reports sent by the contacts, whether they responded to the survey.

713 The second step of the survey focused on the availability/dissemination of exclusion data (***Extended data: Form***
714 ***2***). This step involved (second email) only programs refining their target population by medical exclusions
715 according to the answers provided in the first step. Evaluation data from the biannual campaign (2019-2020
716 /2020-2021) have been collected.

717 The third step of the survey (third email) involved collection of information on the use of other databases to
718 optimize exclusions (**Extended data: Form 3*)***. For referenced databases listed during the first step of the
719 survey, whether or not they were connected to the screening database, the survey collected: i) the Type: Claims-
720 DB (Reimbursement database for medical/paramedical care and benefits from health insurance companies and
721 schemes), Other-MADB (i.e. Hospital discharge/morbidity database) and cancer registries; ii) the start date of
722 data collection, iii) the geographical area covered, iv) the availability of the data (Permanent or limited retention
723 period) (**Extended data: Table 1*)***.

724 ***Conduct of the literature search***

725 The non-systematic literature review sought a published definition (codes and terminology source) of each
726 exclusion criterion applied to each database listed in the previous step. First, for each referenced database, the
727 articles sent by the survey contact persons were used. Then, an email was sent to the contact persons
728 recommended in the previous steps as well as to other contacts found on the official sites of the institutions in
729 charge of the referenced databases.

730 Finally, the review was completed by a search on PubMed and Google Scholar. The search equation used the
731 "AND" and "OR" operators. The names of each listed MADBs/cancer registries were combined with each of the
732 keywords (Colon, colorectal, Rectum, colon sigmoid, Colonoscopy, Adenoma, Polyp, Polypectomy, IBD,
733 Familial Adenomatous Polyposis, Hereditary Non-Polyposis CRC, Lynch Syndrome, Ulcerative Colitis,
734 Colorectal Cancer) and publication date ≥2010.

735 Regardless of the language of publication, articles whose title or abstract had at least one of the terminologies
736 defining a MABD and at least one of the keywords, were reviewed. The study was selected only if a list of codes
737 in a reference terminology (Codes of diagnostic or treatment procedure, anatomopathological or biological
738 examination code, drug code) was described. In cases where multiple studies were performed on the same

739 morbid situation in the same database, the most recent study was preferred. Studies describing several
740 phenotypes of several morbid situation that can justify exclusion from the CRCSP were preferred to those
741 describing a single phenotype or a single morbid situation.

742 ***Data synthesis and analysis plan***

743 ***- CRCSP exclusion approach:*** The practice of medical exclusion has been defined by subtracting from the
744 campaign target population, people affected by a morbid situation justifying exclusion from campaign target
745 population. Whatever the means/strategy used (Exclusion carried out using cancer registries and other databases,
746 exclusion carried out using data provided by patients or their attending physicians), the practice of exclusion was
747 retained only if this exclusion was quantifiable (Number of people excluded from a campaign known with
748 precision), whether it was carried out before or after the campaign invitations. The frequency of each
749 means/strategy of exclusion was described.

750 ***- Consensual definition of the person to be excluded from the CRCSP:*** The exclusion criterion was qualified as
751 consensual when it was in force in the majority (>50%) of the programs practicing exclusion, and non-
752 consensual if ≤50% of the programs. The exclusion was qualified as temporary if the excluded person could
753 return to the program after a time (T) during which it was not invited to take part in a screening campaign, and
754 definitive if not.

755 ***- Analysis of data from a biannual screening campaign:*** For each program whose exclusion data were
756 accessible, the indicators assessed in the program were first described according to the means and strategies of
757 exclusion. The total target population was the number of people in CRC screening target age. The number of
758 exclusions for any reason includes medical exclusions and non-medical exclusions such as obvious refusal to
759 take part in a campaign. The medical exclusion rate (MER) in the campaign was estimated by the ratio between
760 the number of people excluded for medical reasons and the target population size (or population invited to the
761 campaign). The campaign participation rate (CPR) was estimated by the ratio between the number of people
762 having completed a primary screening test and the target population size minus the total of exclusions. The
763 coverage rate of the target population was estimated by the sum of MER + CPR. A comparative description of
764 these rates was made according to the means used for exclusion.

765 ***- Operational formulation of the person to be excluded and possibility of querying from listed databases:*** The
766 database was qualified as national/regional only if, in addition to the fact that its coverage area was

767 national/regional, it was exhaustive on all sections of the national/regional population. The connection between
768 the screening database and other databases was qualified as established if there was a systematic process for
769 extracting (not anonymized) or refining the eligible population upstream of the campaign invitations, using these
770 connected databases. For each consensual exclusion criteria, list of codes from terminologies used to search in
771 the databases have been described.

772

3.2 Performance des algorithmes de recherche dans les bases de données médico-administratives des facteurs morbides pouvant justifier d'une exclusion médicale des campagnes de DOCCR, Revue de la littérature: Étude 2.

3.2.1 – CONTEXTE

Une meilleure formulation opérationnelle de la personne à exclure des campagnes pour pouvoir requêter et l'exclure à partir du SNDS, passe nécessairement par une compilation et une synthèse des algorithmes existants. Cette 2ème étape permet de comprendre comment les requêtes sur les BDMA ont été formalisées pour être capable de rechercher chacune des situations morbides. En permettant de quantifier l'imprécision des algorithmes existants, cette étape pourrait prévenir des potentielles erreurs d'extraction des données du SNDS.

3.2.2 – METHODES

Une revue de la littérature systématique réalisée sur *PubMed et google Scholar*, a compilé des études publiées entre 2010 à 2022 et rapportant une validation d'un phénotype ou d'un algorithme de recherche, dans les BDMA, des situations morbides pouvant justifier d'une exclusion médicale du DOCCR. La performance des algorithmes a été a analysé en termes de Sensibilité -se-, spécificité -sp-, valeur prédictive positive-VPP-, valeur prédictive négative-VPN-, coefficient kappa). Pour analyser cette performance, la revue a collecté :

i) Le modèle de phénotype : En admettant que les modèles de phénotype sont régis par leur structure et leur syntaxe de définitions, Chapman et al. [139] ont distingué 5 modèles/formats de phénotypes (1-***Une liste de codes***; 2-***des éléments de données simples***; 3-***des éléments de données complexes***; 4-des éléments ***temporel***, 5-***un classificateur formé***).

ii) Pour tout phénotype défini par une liste de codes dans une base de données structurée, ont été décrits les types de code (acte de diagnostic ou acte de soin, code médicament) cibles de l'algorithme et la terminologie de codage utilisée.

iii) Le mécanisme de validation de l'algorithme d'extraction : Bien que la comparaison avec un référentiel standard soit le moyen courant de validation d'un phénotype extrait, Chapman et al. [139] ont décrit 6 autres mécanismes de validation d'un phénotype extrait par un algorithme : a-Comparaison avec les ***données d'un registre*** (Comparaison de la cohorte extraite par le phénotype avec celle présente dans un registre de maladies, exemple registre des cancers) ; b-Comparaison de la cohorte extraite par le phénotype avec celle obtenue à l'aide d'une recherche manuelle/humaine dans les ***dossiers médicaux*** ; c-Analyse de la ***concordance inter-enregistrements*** (Comparez le pourcentage de cas identifiés par le phénotype dans différentes sources d'enregistrement dématérialisées et identifiez tout chevauchement) ; d-Analyse de ***facteurs de risque*** (Comparer l'ampleur de la cohorte extraite par le phénotype avec

les résultats d'un modèle d'estimation de risque standard) ; e-Analyse de ***pronostic*** (Comparer l'ampleur de la cohorte phénotypique avec des modèles de pronostic externes) ; f-Association génétique (Comparer si la présence d'un patient dans une cohorte extraite par un phénotype est cohérente avec son profil génétique).

3.2.3 – RESULTATS

En total, 33 études portant sur la validation d'algorithmes d'extraction ont été collectés par le biais de *Pubmed et google Scholar et* incluses dans cette analyse [54, 55, 62, 81, 82, 88, 90, 91, 95, 96, 98, 100, 107, 114, 117, 118, 126, 140-155]. Pour la plupart des facteurs d'exclusion médicale, il existe des algorithmes de recherche de phénotypes dans les BDMA validés et reproductibles. Ces études de phénotypes ont été réalisées dans plusieurs pays ayant un DOCCR et pratiquant une exclusion médicale, notamment en Amériques [81, 95, 107, 117, 118, 140, 141, 143, 146, 152, 155], en Australie [90, 100], en Europe [54, 55, 62, 82, 88, 90, 96, 126, 148, 150, 151, 153], en Corée du Sud [144, 147], au Japon [91, 114, 142, 145, 149, 154]. A l'exception de quelques études qui ont développé d'autres formats de cibles (phénotypes) des algorithmes, la quasi-totalité des algorithmes ciblait une liste de codes (***Tableau-3.2.3.1***).

Le mécanisme de validation privilégié était la comparaison des données extraites par l'algorithme à celles présentes dans un registre. Pour ce mécanisme de validation, les auteurs ont le plus souvent défini la sensibilité par la probabilité d'être identifié comme cas de cancer par l'algorithme lorsque le cancer était présent dans les données de référence. La spécificité par la probabilité de ne pas être identifié comme cas de cancer lorsque le cancer n'était pas présent dans les données de référence ; la VPP par la probabilité d'avoir un cancer lorsqu'il est identifié comme cas de cancer par l'algorithme ; et le VPN par la probabilité de ne pas avoir de cancer lorsqu'il n'est pas identifié comme cas de cancer par l'algorithme[90].

En plus de ces indicateurs, plusieurs études ont utilisé le coefficient kappa (Cohen's κ) pour valider la concordance entre divers méthodes [54, 83, 88, 95]. Earles et al. [95] ont défini ce coefficient par le rapport $(P_o - P_e)/(1-P_e)$, où P_o est la proportion observée d'individus pour lesquels deux méthodes concordent et P_e est la probabilité que les deux méthodes concordent par hasard (attendue), sur la base de la classification des cas et des témoins de chaque méthode. Le coefficient $\kappa > 0$ si la concordance observée dépasse la proportion attendue et atteint sa valeur maximale de 1 lorsque deux méthodes parviennent à une concordance parfaite. Il ressort de cette synthèse de la littérature qu'aucun algorithme d'extraction n'a une précision à 100%.

Tableau-3.2.3.1: Exemples d'études de validation des phénotypes de CCR, coloscopie et IBD dans des BDMA.

Auteur (Année de Publication) [Référence]	Cible	**Terminologie (Codes)**	**Modèle de phénotype (Mécanisme de validation)**	**Meilleure valeur de mesure des indicateurs**	**BDMA, PAYS/REGION**
Albaek Jacobsen et al. (2022) [55]	MICI	CIM-10 (K50, K51)	1 (a, b)	Vpp (98)	The Danish National Patient Registry, DANEMARK
Ananthakrishnan et al. (2013) [140]	MICI	CIM-9 (555 or 556)	3 (d)	Se (79), Sp (97), Vpp (97)	EMR data from Massachusetts General Hospital and Brigham and Women's Hospital, USA
Benchimol et al (2014) [81]	MICI	CIM-9 (555 or 556); ICD10(K50 or K51)	1 (d)	Se(76,8), Sp (96,2), Vpp (81,4), Vpn (95,0)	Ontario Health Insurance Plan (OHIP) Claims Database, Hospital discharge abstract data, CA
Broe et al. (2020) [82]	CCR	Code local médicament chimio	1 (b)	Se (90), Vpp (95)	The Danish National Patient Registry, DANEÙARK
Chawla N et al. (2014) [141]	CCR	CIM-9-CM (métastase: 196.2, 197.5...)	1 (a)	Se (55), Sp (96,8), Vpp (95,2), vpn (65)	Medicare claims database, USA
Cozzolino et al. (2018) [88]	CCR	CIM-9(153.x, 154.x)	1 (b)	Se(99), Sp (87), Vpp (81), Vpn(99), k (0,9)	Hospital discharge database, Napoli, Venezia, ITALIE
Daniels et al. (2019) [90]	CCR	CIM-10, médicament (ATC)	1 (b)	Se(22,6), Sp (93,8), Vpp(82,7), Vpn (48,2)	Australian Pharmaceutical Benefits Scheme data, AUSTRALIE
De Luise et al. (2021) [91]	CCR	CIM-10 (C18-C20; D01.0-D01.2)	1 (b)	Se (41,2), Sp (99,9), Vpp (84,4)	Hospital Claim data, JAPON
Earles et al. (2018) [95]	CCR	CIM-9 (153.xx, 154.xx), ICD10-AM(C18-C20), CPT	1 (a, b)	Vpp (58), k(0,87)	VA: administrative claims data, USA
Escriba et al. (2020) [96]	CCR	CIM-10 (?)	1 (b)	Se(83,9), Vpp(94,9)	The Catalan minimum basic data set (MBDS) of hospital discharges, ESPAGNE
Fujiwara et al. (2021) [142]	CCR	CIM-10 (C182, C184, C185, C186, C187, C189, C19, C20)	1 (b)	Se 100), Sp (99,9), Vpp (80,5), Vpn (100)	Japanese health administrative real-world database (RWD), JAPON
Gil et al. (2019) [98]	CCR	CIM (C18-C21), texte	1 (b, d)	Vpp (87,3)	The Spanish primary care database BIFAP, ESPAGNE
Goldsbury et al. (2019) [100]	CCR	CIM-10-AM (C18-C20)	1 (b)	Se (94), Vpp (93)	Admitted Patient Data Collection, AUSTRALIE
Hawang et al. (2018) [144]	CCR	CIM-10 (C18-C21)	1 (b)	Se(100), Sp (98,9), Vpp (99,7), Vpn(100)	Seoul National University Bundang Hospital EMR, COREE DU SUD
Hou et al. (2014) [143]	MICI	CIM-9 (555 or 556)	1 (b)	Vpp (94)	Veterans Affairs Health Care Systems (VA) administrative datasets, USA
Ihira et al. (2022) [145]	CCR	CIM-10 (C18-C20) Code Local traitement	1 (a)	Se (91.0), Sp(99.5), Vpp (32.6), Vpn (99.9)	National Health Insurance System data/Medical Care System data, JAPON

Ko et al. (2014) [107]	COLO	CPT, HCPCS (Coloscopie)	2 (d)	Se(56), Sp(96), Vpp(84), Vpn(87)	Medicare claims database, USA
Lamont et al. (2014) [146]	CCR	CPT, HCPCS (Chimiothérapie)	1 (a)	Se (78)	Medicare claims database, USA
Lash et al. (2015) [54]	CCR	CIM-10 (métastase: C76–C80)), SNOMED, Code local traitement	1, 2 (c)	Se (95), Sp (95), Vpp (86), Vpn (99), k (0,57)	The Danish National Patient Registry, The Danisc cancer registry, DANEMARK
Lee CK et al. (2020) [147]	MICI	CIM-10 (K50.x, K51.x), Code local médicament	1 (b)	Se (93,1), Sp (98,1), Vpp (97,5), Vpn (94,5)	National Health Insurance System (NHIS) database, COREE DU SUD
Lo et al. (2020) [148]	MICI	CIM-10 (K50, K51)	1 (b)	Se (95), Sp (96)	The Danish National Patient Registry, DANEMARK
Morikubo et al. (2021) [114]	MICI	CIM-10 (K50.x, K51.x), Code local médicament	1 (b, c)	Se (100), Sp (98,5), Vpp (97), Vpn (100), k 0,99)	Japanese health administrative data, JAPON
Nordstrom et al. (2016) [117]	CCR	CIM-10 (C18.x, C19.9, C20.9, C21.0–C21.2, C21.8)	1 (a, b)	Se (42,4), Sp (92,7), Vpp (43,8), Vpn (92,3)	Geisinger Health Plan claims databases, USA
Nordstrom et al. (2012) [118]	CCR	CIM-9 (153.xx, 154.xx)	1 (a)	Se (60), Sp (88), Vpp (81), Vpn (72)	Surveillance Data, Inc. (SDI) Health claims databases, USA
Ogino et al. (2022) [149]	MICI	CIM-10 (K51), Code local	1 (b)	Vpp (90,6)	Hospital Claim data (Kitasato Institute Hosp Tokyo & Kyushu Univ Hosp Fukuoka), JAPON
Quantin et al. (2014) [151]	CCR	CIM-10 (C18-C20), CCAM	1 (a, c, d)	Se(87,5), Vpp(77,9),	SNDS (PMSI), FRANCE
Quantin et al. (2012) [150]	CCR	CIM-10, CCAM	1 (a, c, d)	Se (87,5), Vpp (77,9),	SNDS (PMSI), FRANCE
Recalde et al. (2019) [62]	CCR	CIM-10(C18-C21)	1 (d)	Se (81,1), Vpp (70,1)	Information System for Research in primary health care(SIDIAP), Catalogne, ESPAGNE
Rezaie et al. (2012) [152]	MICI	CIM-9(555 or 556); ICD10(K50 or K51)	1 (a, b)	Se (83,4), Sp (99,8), Vpp (97,4), Vpn (98,5)	Physician billing claims (PC) database, Hospital Discharge Abstract Database (DAD), CANADA
Rye et al. (2021) [153]	MICI	CIM-10 (K50, K51)	1 (b)	Vpp (98)	The Danish National Patient Registry, DANEMARK
Takahashi et al. (2022) [154]	MICI	CIM-10 (K50, K51)	1 (a)	Se (75,6), Sp (99,8), Vpp (94,8), Vpn (98,7)	Tohoku University HospitaL database, JAPON
Thirumurthi et al, (2010) [155]	MICI	CIM-9-CM(555 or 556)	1 (b)	Se (84), Sp (99), Vpp (82)	VA administrative data, USA
Vesteghem et al. (2021) [126]	CCR	CIM-10 (C18-C20) Code Local traitement, ATC	1 (b, c)	Se (98,4), Vpp (98,8)	The Danish National Patient Registry, DANEMARK

***ATC** : Anatomical Therapeutic Chemical Classification System ; **CCAM** : Classification Commune des Actes Médicaux ; **CIM** : Classification Internationale des Maladies ; **CPT** : Current Procedural Terminology ; **PMSI :** Programme de médicalisation des systèmes d'information ; **HCPCS** : Healthcare Common Procedure Coding System ; **Se :** Sensibilité, **SNDS :** Système national des données de santé ; **SNOMED :** Systemized Nomenclature of Medicine ; **Sp** : Spécificité ; **Vpp :** Valeur prédictive positive ; **Vpn :** Valeur prédictive négative. **Modèle de phénotype :** 1-Une liste de codes; 2-des éléments de données simples; 3-des éléments de données complexes; 4-des éléments temporel. **Mécanisme de validation** : **a**-Comparaison avec les données d'un registre ; **b**- Comparaison avec dossiers médicaux ; **c**-Analyse de la concordance inter-enregistrements ; **d**-Analyse de facteurs de risque; e-Analyse de pronostic.*

3.3 Synthèse des résultats sur la définition de la personne cible.

L'enquête d'exclusion médicale avait ciblé 41 pays et une réponse a été obtenue de 31 programmes dans 30 pays. Parmi les 30 pays, 24 avaient un DOCCR à l'échelle régionale ou nationale. Sur les 25 DOCCR répertoriés, 18 réalisaient des exclusions médicales et 2 programmes (Finlande, Norvège) prévoyaient d'en réaliser en 2022. Parmi les 11 critères d'exclusion répertoriés, seulement 5 (Antécédent personnel de CCR, de MICI, d'adénome, Coloscopie/coloscanner récente, risque génétique) étaient consensuels (appliqués par >60% des programmes) dans les programmes affinant leurs populations cibles par des exclusions médicales.

Il faut cependant noter que dans la plupart des programmes, il existe une différence importante entre les recommandations du cahier des charges et ce qui est réellement appliqué. Les exclusions sont très dépendantes des moyens à la disposition des structures, en illustration, en République tchèque, les critères sont théoriquement définis conformément à la définition du risque moyen mais l'exclusion médicale est à la discrétion du médecin référent (généraliste, gynécologue) qui peut décider d'exclure ou pas un patient et surtout de la durée d'une exclusion.

L'enquête et la revue de la littérature révèlent qu'il existe au moins une BDMA [43-63, 156-158] ou un registre de cancers[43, 46, 49, 62, 64-77], couvrant totalement ou partiellement, chacun des Etat/pays/régions ayant un déploiement complet du DOCCR et pratiquant une exclusion médicale. Cependant, quelques rares programmes avaient une interconnexion régulière entre la base de données du DOCCR et une BDMA. Aucun des algorithmes répertoriés n'avait une précision de 100%, notamment l'algorithme de Quintin et al. [151] validé en France dans la base du PMSI.

4. AXE 2 : DEFINIR UN PATIENT A EXCLURE DU DEPISTAGE A PARTIR DES DONNEES DU SNDS

4.1 Pratiques actuelles des structures de gestion du dépistage en termes d'exclusion médicale aux campagnes de dépistage: Étude 3.

4.1.1 – CONTEXTE

Le cahier de charge du DOCCR définit les critères d'exclusions médicales du DOCCR sans être explicite sur les moyens à mettre à disposition des SGDO pour optimiser l'exclusion médicale. Lors des premières évaluations des campagnes, SPF admettait que le nombre de personnes exclues était fourni par les structures de gestion à partir de diverses sources de données exploitables (réponses des personnes invitées, fichiers des actes de coloscopies remboursés par les caisses primaires d'assurance maladie, fichiers des actes diagnostiques et thérapeutiques, comptes rendus standardisés des résultats de coloscopie et des examens anatomo-cytopathologiques…), mais dont l'accès variait d'une structure de gestion du dépistage (SGDO) à l'autre[40]. En amont de toute procédure de standardisation de la démarche d'exclusion médicale, un état des lieux sur les pratiques empiriques fournira les variables d'ajustement des recommandations.

4.1.2 – METHODES

Il s'agissait d'une enquête transversale et descriptive portant sur les pratiques d'exclusion en vigueur dans les SGDO. Un questionnaire a été adressé aux SGDO pour connaitre les stratégies et moyens utilisés pour les exclusions médicales dans les départements français. Les informations recueillies par ce questionnaire étaient : L'existence d'un partenariat avec les structures pourvoyeuses notamment les CPAM, les registres de cancers et les centres hospitaliers, l'année de début du partenariat, le type de données fournies, la fréquence annuelle de transmission des données et l'avis sur une demande collective des SGDO pour le requêtage des exclusions à partir du SNDS.

Revue d'Epidémiologie et de Santé Publique

Enquête transversale sur les pratiques d'exclusion médicale lors des campagnes de dépistage du cancer colorectal en France

--Projet de manuscrit--

Numéro du manuscrit:	
Type d'article:	Article original / Original article
Mots-clés:	Dépistage du cancer colorectal; Exclusion médicale; Taux de participation
Auteur correspondant:	AKOI KOIVOGUI, M.D., M.P.H. Laboratory of Medical Informatics and Knowledge Engineering in e-Health BONDY, FRANCE
Premier auteur:	AKOI KOIVOGUI, M.D., M.P.H.
Ordre des auteurs:	AKOI KOIVOGUI, M.D., M.P.H.
	CHRISTIAN BALAMOU
	ROBERT BENAMOUZIG
	CATHERINE DUCLOS
Résumé:	Objectif : Décrire l'évolution des taux bruts d'exclusion (TBE) et de participation (TBP) aux campagnes en fonction des pratiques d'exclusion médicale dans le programme français de dépistage organisé du cancer colorectal (DOCCR). Méthode : L'enquête transversale conduite en septembre 2021 a collecté les informations sur l'utilisation de données reçues des structures pourvoyeuses de données médicales (SPDM) pour qualifier l'exclusion médicale. La population éligible à une campagne de DOCCR, était la population cible extraite des fichiers des régimes d'assurance maladie, dont sont soustraites, toutes les personnes affectées par une situation morbide justifiant d'une exclusion médicale. Les données (exclusion médicale participation) analysées sont celles collectées par Santé Publique France (SPF) auprès des structures de gestion du DOCCR, compilées et diffusées en libre accès. Les TBE (Proportion de personnes exclues de la population cible pour raison médicale) et TBP (Proportion de personnes ayant réalisé un test dans la population éligible) ont été décrits en fonction de l'utilisation des données des SPDM entre 2010 et 2021. Résultats : 60/101 départements (59,4%) ont participé à l'enquête. Parmi eux, 22 (36,7%) recevaient des données d'au moins une des trois principales structures pourvoyeuses (Caisse primaire d'assurance maladie -CPAM-, Registres de Cancers -RC-, Centres hospitaliers -CH-). Le TBE (Campagne 2016-2017 : 12,0% dans les départements ne recevant aucune donnée des SPDM -AUCUN-; 16,1% dans ceux recevant les données des CPAM ; 14,0% dans ceux utilisant les données RC; et 13,8% dans ceux utilisant les données hospitalières; p<0,05. Campagne 2021-2022, ces proportions étaient respectivement de : 12,5%; 16,0%; 13,1% et 12,7%; p<0,05) et le TBP (Campagnes 2016-2017 : Aucun : 31,2%; CPAM : 31,7%; RC : 33,5%; CH :31,8%; p<0,05. En 2021-2022: Aucun: 33,1%; CPAM: 29,7% ; RC: 37,2% ; CH: 32,0% ; p<0,05) variaient significativement en fonction des moyens d'exclusion médicale. Conclusion : Il existe une disparité dans l'accès aux moyens supplémentaires d'exclusion médicale qui n'a pu être corrigée par la régionalisation du DOCCR. Les départements ayant eu accès à au moins l'un des trois moyens supplémentaires d'exclusion médicale, avaient relativement un taux de participation supérieur à la moyenne globale.

Activé par Editorial Manager® et ProduXion Manager® depuis Aries Systems Corporation

Enquête transversale sur les pratiques d'exclusion médicale lors des campagnes de dépistage du cancer colorectal en France

Cross-sectional survey on medical exclusion practices during colorectal cancer screening campaigns in France

Akoï Koïvogui[1,2], Christian Balamou[3], Robert Benamouzig[4], Catherine Duclos[1,5]

[1]Université Sorbonne Paris Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, LIMICS, Sorbonne Université, INSERM, 93000, Bobigny, France.

[2]CRCDC-IDF, Site de Seine-Saint-Denis, 93146, Bondy, France.

[3]CRCDC-AURA, Site de l'Ain Bourg-en-Bresse, 01000, Bourg-en-Bresse, France.

[4]AP-HP, Hôpital Avicenne, Service d'Hépato Gastro Entérologie, 93000, Bobigny, France.

[5]AP-HP, Hôpital Avicenne, Service d'Information médicale, 93000, Bobigny, France.

Correspondence to: Dr Akoï KOÏVOGUI, Université Sorbonne Paris Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances en e-Santé, LIMICS, Sorbonne Université, INSERM, 93000, Bobigny, France.
mail: akoi.koivogui@inserm.fr

Titre courant: Pratique d'exclusion médicale et dépistage du CCR

Nombre de mots: manuscript (3783), résumé (332/342)

Enquête transversale sur les pratiques d'exclusion médicale lors des campagnes de dépistage du cancer colorectal en France

Summary

Objective : To describe the evolution of the campaign exclusion rate (CER) and the campaign participation rate (CPR), according to medical exclusion practices in the French population-based colorectal cancer screening program (CRCSP).

Method : The cross-sectional survey conducted in September 2021 collected information on the use of data received from structures providing medical data (SPMD) to qualify medical exclusion. The eligible population for a CRCSP campaign was the target population extracted from the health insurance plan data, from which people affected by a morbid situation justifying medical exclusion were subtracted. The data (medical exclusion and participation data) analyzed were those collected by French public health agency (SPF) from the CRCSP management structures, compiled and published as open access data. The CER (Proportion of people excluded from the target population for medical reasons) and CPR (Proportion of people who completed a screening test in the eligible population) were described based on the use of SPMM data between 2010 and 2021.

Results : 60 French districts out of 101 (59.4%) participated in the survey. Among them, 22 (36.7%) received data from at least one of the three main SPMD (Primary Health Insurance Agency -CPAM-, Cancer Registries -CR-, Hospital). The CER (In 2016-2017 campaign: 12.0% in districts receiving no data from SPDM -NONE-; 16.1% in those receiving data from CPAM; 14.0% in those using RC data; and 13.8% in those using Hospital discharge morbidity data -HDMD-; $p<0.05$. In 2021-2022 campaign, these proportions were respectively: 12.5%; 16.0%; 13.1% and 12.7%; $p< 0.05$) and CPR (In 2016-2017 campaign: NONE: 31.2%; CPAM: 31.7%; CR: 33.5%; HDMD: 31.8%; $p<0.05$. In 2021- 2022 campaign: NONE: 33.1%; CPAM: 29.7%; CR: 37.2%; HDMD: 32.0%; $p<0.05$) varied significantly depending on the use of SPDM's data.

Conclusion: There is a disparity in access to additional means of medical exclusion which could not be corrected by the new organization of the CRCSP. Districts that had access to at least one of the three SPDM's data had a relatively higher participation rate than the overall average.

Keywords: Colorectal cancer screening; Medical exclusion; Participation rate

Résumé

Objectif : Décrire l'évolution des taux bruts d'exclusion (TBE) et de participation (TBP) aux campagnes en fonction des pratiques d'exclusion médicale dans le programme français de dépistage organisé du cancer colorectal (DOCCR).

Méthode : L'enquête transversale conduite en septembre 2021 a collecté les informations sur l'utilisation de données reçues des structures pourvoyeuses de données médicales (SPDM) pour qualifier l'exclusion médicale. La population éligible à une campagne de DOCCR, était la population cible extraite des fichiers des régimes d'assurance maladie, dont sont soustraites, toutes les personnes affectées par une situation morbide justifiant d'une exclusion médicale. Les données (exclusion médicale participation) analysées sont celles collectées par Santé Publique France (SPF) auprès des structures de gestion du DOCCR, compilées et diffusées en libre accès. Les TBE (Proportion de personnes exclues de la population cible pour raison médicale) et TBP (Proportion de personnes ayant réalisé un test dans la population éligible) ont été décrits en fonction de l'utilisation des données des SPDM entre 2010 et 2021.

Résultats : 60/101 départements (59,4%) ont participé à l'enquête. Parmi eux, 22 (36,7%) recevaient des données d'au moins une des trois principales structures pourvoyeuses (Caisse primaire d'assurance maladie -CPAM-, Registres de Cancers -RC-, Centres hospitaliers -CH-). Le TBE (Campagne 2016-2017 : 12,0% dans les départements ne recevant aucune donnée des SPDM -AUCUN-; 16,1% dans ceux recevant les données des CPAM ; 14,0% dans ceux utilisant les données RC; et 13,8% dans ceux utilisant les données hospitalières; $p<0,05$. Campagne 2021-2022, ces proportions étaient respectivement de : 12,5%; 16,0%; 13,1% et 12,7%; $p<0,05$) et le TBP (Campagnes 2016-2017 : Aucun : 31,2%; CPAM : 31,7%; RC : 33,5%; CH :31,8%; $p<0,05$. En 2021-2022: Aucun: 33,1%; CPAM: 29,7% ; RC: 37,2% ; CH: 32,0% ; $p<0,05$) variaient significativement en fonction des moyens d'exclusion médicale.

Conclusion : Il existe une disparité dans l'accès aux moyens supplémentaires d'exclusion médicale qui n'a pu être corrigée par la régionalisation du DOCCR. Les départements ayant eu accès à au moins l'un des trois moyens supplémentaires d'exclusion médicale, avaient relativement un taux de participation supérieur à la moyenne globale.

Mots clés : Dépistage du cancer colorectal, Exclusion médicale, Taux de participation

Introduction

La plupart des politiques de dépistage mises en place dans l'Union européenne (UE), ciblaient une population à risque moyen de cancer colorectal (CCR), définie sur le critère âge [1] avec une variabilité considérable entre programmes. Le dépistage était proposé aux hommes et femmes âgés de 40 à 60 ans en Bulgarie, de 50 à 74ans en France, en Angleterre ou en Belgique ; de 50 à 70ans en Italie et de 55 à 75ans en Hollande [2].

En France, depuis 2009, les campagnes étaient organisées conformément à un cahier des charges du dépistage organisé du CCR (DOCCR)[3]. En amont de chaque campagne, une mise à jour des fichiers de personnes cibles était faite après la transmission des données individuelles par les différents Régimes obligatoires d'Assurance Maladie (RAM) sur la base de leur âge et lieu de résidence. Les personnes étaient ainsi invitées à leur 50ème anniversaire (primo-invitation) et tous les deux ans après un test de dépistage (Invitation-subséquente) jusqu'à l'âge de 74ans révolus.

Malgré une preuve d'efficience des campagnes de dépistage en termes de diminution de la mortalité par cancer colorectal, le taux de participation aux campagnes en France, reste très inférieur à l'objectif européen minimal de 45 % [4-6]**.** Certes, plusieurs raisons ont déjà été évoquées pour expliquer cette faible participation [5, 7-10] mais le biais lié à la méthode d'estimation du taux de participation n'est pas souvent discuté. Dans son protocole d'évaluation du DOCCR, Santé Publique France (SPF) a défini le taux de participation par le rapport entre le nombre de personnes dépistées et la population éligible, soit la population cible du programme (hommes et femmes de 50 à 74 ans) dont sont soustraites, les personnes exclues du programme pour raisons médicales et non médicales (refus par exemple)[11]. La définition de la population éligible était basée sur une notion de risque moyen caractérisé par l'âge (50-74 ans) et l'absence de risque personnel ou familial connu[12]. Ainsi, les personnes ayant un risque élevé (antécédents personnels ou familiaux d'adénomes colorectaux, de cancers ou de maladies inflammatoires chroniques de l'intestin (MICI) : colite ulcéreuse et maladie de Crohn) ou très élevé (polypose adénomateuse familiale (PAF), syndrome de Lynch) de développer un CCR et les personnes atteintes de pathologies extra-intestinales sévères, n'étaient pas éligibles aux campagnes de DOCCR. De même, les personnes ayant réalisé une coloscopie complète dans les 5 ans ou une coloscanner dans les 2 ans, étaient temporairement exclues si le résultat de cette coloscopie était normal [4].

Selon SPF, le taux d'exclusion était apprécié par les structures de gestion à partir de diverses sources de données exploitables (fichiers des actes de coloscopies remboursés par les caisses primaires d'assurance maladie, fichiers des actes diagnostiques et thérapeutiques, comptes rendus standardisés des résultats de coloscopie et des

examens anatomo-cytopathologiques…), mais dont l'accès variait d'une structure de gestion du dépistage (SGDO) à l'autre[11]. Les structures de gestion avaient recours à ces sources de données complémentaires car le recueil de déclarations volontaires des sujets (questionnaire joint au courrier d'invitation) et des médecins traitant (questionnaire structuré) n'était pas exhaustif. Cette étude s'est fixée pour objectif de décrire les pratiques d'exclusion médicale en fonction des accès des SGDO aux sources de données complémentaires et leurs impacts en termes d'évolution des taux d'exclusion médicale et taux de participation aux campagnes biannuelles.

Méthodologie

Type d'étude

L'étude descriptive a analysé l'évolution du taux brut d'exclusion médicale (TBE) et du taux brut de participation (TBP) aux campagnes de dépistage en fonction des sources de données utilisées par les structures de gestion du dépistage (SGDO) en France entre 2010 et 2022. Les données d'exclusion et de participation sont celles publiées par SPF (https://www.santepubliquefrance.fr/media/files/01-maladies-et-traumatismes/cancers/cancer-du-colon-rectum/taux-de-participation-au-programme-de-depistage-2021-2022/doccr_participation_evolution_2010_2022) après transmission des données par les SGDO.

Les informations en rapport avec la transmission des données aux SGDO par les structures pourvoyeuses ont été collectées lors d'une enquête nationale réalisée en septembre 2021 et portant sur les pratiques d'exclusion médicale en vigueur dans les SGDO. L'enquête transversale a été menée à l'aide d'un questionnaire (**Document_supplémentaire**) adressé par courriel, aux 17 structures régionales pour connaitre les moyens utilisés dans chaque département français avant et après la régionalisation du programme, la fréquence d'utilisation de chaque moyen.

Contextes des exclusions médicales dans le DOCCR

En France, les personnes sont invitées à participer aux campagnes de DOCCR à partir de leur 50ème anniversaire (primo-invitation) et tous les deux ans (Invitation-subséquente) jusqu'à l'âge de 74ans révolus. Pour l'organisation du DOCCR, environ une fois par trimestre, les listes de populations cibles (définies sur le critère d'âge) sont mises à la disposition des SGDO par chaque RAM. De ces listes sont extraites dans chaque SGDO, une population éligible à une campagne (C_n) constituée de personnes indemnes de toute situation morbide pouvant justifier d'une exclusion médicale (Haut risque ou très haut risque de CCR, coloscopie <5ans, coloscanner de <2ans) conformément aux cahier des charges du DOCCR[3].

Toute personne invitée à C_n, est réinvitée deux ans après à la campagne suivante C_{n+1} seulement si le test réalisé à C_n était négatif ou si elle n'avait pas réalisé un test de dépistage à C_n. Si le résultat du test réalisé à C_n était positif, la C_{n+1} dépendra des résultats de la coloscopie de dépistage réalisée suite à ce test positif. La personne est réinvitée pour la C_{n+1} cinq ans après une coloscopie de dépistage négative (exclusion temporaire). La personne est définitivement exclue du DOCCR après une coloscopie de dépistage positive (Adénome/polype à risque de CCR, MICI, CCR).

Pour constituer la liste de population éligible lors d'une prochaine campagne C_{n+1}, la SGDO affine la population cible en amont des invitations par la soustraction des personnes exclues temporairement ou définitivement à cause de cette coloscopie de dépistage réalisée à C_n. Les personnes ayant réalisé une coloscopie ou un coloscanner de diagnostic (hors test de dépistage positif) sont exclues temporairement (en cas de résultat négatif de la coloscopie/coloscanner) ou définitivement (cas de diagnostic d'affection colorectale justifiant d'une exclusion médicale) seulement si l'information est parvenue à la SGDO quelle que soit la source de l'information (la personne elle-même, un médecin, un registre…).

Lors d'une C_n, toute personne invitée à participer (son Médecin traitant ou autre), pouvait adresser une réponse postale (RP) à la SGDO, ce qui permettait à la SGDO de réaliser des exclusions médicales post-invitations. Ces exclusions post-invitations étaient définitives si la RP rapportait une situation morbide justifiant d'une exclusion médicale, notamment : un diagnostic d'une pathologie chronique colorectale (CCR, MICI, adénome à risque, PAF, Syndrome de Lynch), des antécédents familiaux de CCR. Une exclusion temporaire était décidée dans tous les cas où la réponse postale évoquait une coloscopie de moins de cinq ans, une coloscopie incomplète de moins de 2 ans ou un coloscanner de moins de deux ans ou une recherche de sang dans les selles de moins de deux ans. Ces exclusions post-invitations à C_n étaient comptabilisées dans toute analyse portant sur C_n et permettaient d'affiner la population cible en amont des invitations pour toute campagne ultérieure.

Données recueillies et définition des variables d'étude

L'étude a collecté dans l'archive de données de SPF, les effectifs pour chaque département de France et par période de deux années glissantes entre 2010 et 2022: i) la population cible estimée par L'Institut national de la statistique et des études économiques (INSEE), ii) le nombre d'exclusions médicales et iii) le nombre de test de dépistage réalisés. Les informations recueillies par le questionnaire d'enquête étaient : L'existence (année de début et de fin) d'un partenariat entre la SGDO et chacune des potentielles structures pourvoyeuses de données d'exclusion médicale répertoriées (Caisse Primaire d'Assurance Maladie -CPAM-, Registre de cancers -RC-, Département d'information médicale d'un Centre hospitalier -CH-, Autres), le type de données transmises (Acte de coloscopie; Date de l'acte de coloscopie ; Diagnostic de CCR ; Date de diagnostic du CCR ; Diagnostic autre cancer ; Diagnostic de MICI ; Diagnostic d'adénome/Polype), la fréquence annuelle de transmission des données (1fois, 2fois, 3fois, 4fois, ≥5 fois) et l'avis sur une demande collective des SGDO pour le requêtage des exclusions à partir du SNDS.

Les départements pilotes du DOCCR (départements ayant été choisi pour préfigurer les campagnes de DOCCR) ont été distingués des autres départements ayant commencé les campagnes seulement après la généralisation du programme[4].

Pour l'analyse de l'impact des données d'exclusion reçues des structures pourvoyeuses, 3 périodes ont été définies pour chaque département ayant reçu des données et pour chacune des principales sources listées dans le questionnaire d'enquête : i) La période comprenant toutes les campagnes antérieures à la date de début de transmission des données par la source considérée : "AV_ACCES" ; ii) La période comprenant toutes les campagnes au cours desquelles il y a eu transmission des données par la source considérée : "PENDANT_ACCES" ; iii) La période comprenant toutes les campagnes postérieures à la date de fin de transmission des données par la source considérée : "AP_ACCES". Pour les périodes "AV_ACCES" et "PENDANT_ACCES", les taux (TBE et TBP) les plus élevés ont été choisis comme taux de référence de la période. Pour la période "AP_ACCES", le taux de référence était le taux enregistré lors de la dernière campagne 2021-2022.

Huit types d'accès aux moyens d'exclusion médicale lors d'une campagne ont été distingués: i) La SGDO n'avait aucun autre moyen d'exclusion médicale en dehors des RP (AUCUN) ; ii) La SGDO a utilisé les données de la CPAM en plus des RP (CPAM) ; iii)les données d'un registre de cancers en plus des RP (RC) ; iv)les données hospitalières en plus des RP (CH) ; v)à la fois, les données des CPAM et RC (CPAM/RC) ; v)à la fois, les données des CPAM et CH en plus des RP (CPAM/CH) ; v)à la fois données des RC et CH (RC/CH) ; v)à la fois, les données des CPAM, RC et CH en plus des RP (CPAM/RC/CH).

Analyse des données

Le taux brut d'exclusion (TBE) a été estimé par le rapport entre le nombre de personnes exclues temporairement ou définitivement des campagnes de DOCCR et la population cible du dépistage (Taille de la population cible estimée par l'INSEE). Le TBP a été exprimé par le rapport entre le nombre de personnes dépistées et la population cible du dépistage (INSEE) dont ont été soustraites le nombre de personnes exclues du dépistage pour raisons médicales.

Un Chi-2 de tendance a été faite au seuil de 5 %, pour tester la variation des taux entre périodes (Avant vs. Pendant puis Pendant vs. Après). Un Chi2 de Pearson au seuil de 5% a testé la différence de taux en fonction des types d'accès aux données d'exclusion médicale lors de chaque campagne biannuelle.

Une courbe évolutive des taux (TBE et TBP) par département, a été tracée en fonction des modes d'exclusion médicale utilisés . Les analyses réalisées ont été faites avec la version 13 du logiciel STATA (College Station, Texas, USA).

Mention légale

Cette étude s'inscrit dans une démarche d'évaluation continue du DOCCR, à ce titre elle bénéficie de l'avis favorable de la Commission Nationale de l'Informatique et des Libertés (CNIL) à travers la délibération n° 2015-175 du 11 juin 2015 portant autorisation unique de traitements de données à caractère personnel ayant pour finalité le dépistage organisé[13]. Conformément à la législation française en vigueur, une étude qui ne change pas la prise en charge des patients ne nécessite pas l'avis d'un comité d'éthique.

Résultat

Le questionnaire a été retourné pour 60 des 101 départements français, soit un taux de retour de 59,4%. 13 départements pilotes sur 23 ont retourné le questionnaire. Parmi les 13 régions de la France métropolitaine, 8 ont eu au moins un département qui a retourné le questionnaire. Parmi les 5 DOM, seule la Réunion a retourné le questionnaire. Au total, 22 départements (dont 9 des 13 départements pilotes) sur les 60 départements ayant répondu à l'enquête avaient reçu des données d'au moins un des trois types de structures partenaires entre 2010 et 2022. Seulement 2 départements (Doubs, et Isère), ont reçu des données d'exclusion de tous les 3 types de structures pourvoyeuses de données individuelles (***Tableau-1***).

Dans le Doubs, sur la période "*AV_* ACCES_*CPAM*" (en 2017), le plus fort TBE (taux de référence) enregistré était de 12,3% et le plus fort TBP enregistré était de 34,3%. L'augmentation de ces taux entre les périodes "*AV_* ACCES_*CPAM*" vs. "PENDANT_ACCES_CPAM" (TBE :14,4% ; TBP :40,3%) était significative ($p<0,05$). De même, leur diminution observée entre les périodes "PENDAT_ ACCES_CPAM" vs. "AP_ ACCES_CPAM" était significative ($p<0,05$) (***Tableau-2***). Dans le Doubs, l'accès aux données du Registre des tumeurs de Doubs a aussi débuté en 2017 (***Tableau-3***) et l'accès aux données hospitalières a été accordé sur la campagne 2017-2018 ***Tableau-4***). L'usage des données de registre n'a pu infléchir la baisse du TBR amorcée dans ce département (Doubs) après l'arrêt de la transmission des données de la CPAM et centres hospitaliers **(Document_supplémentaire Fig_sup 1, 2 et 3).**

Globalement, les départements qui utilisaient les données mises à leur disposition par les structures pourvoyeuses avaient un TBE et un TBP significativement plus élevés comparativement à ceux qui n'utilisaient pas de données autres que les réponses postales. Quelle que soit la campagne, le TBE et le TBP variaient significativement en fonction des moyens d'exclusion médicale. A la campagne 2016-2017, le TBE était de 12,0% dans les départements ne recevant aucune donnée des structures pourvoyeuses -AUCUN-; de 16,1% dans ceux recevant les données des CPAM ; de 14,0% dans ceux utilisant les données RC; et de 13,8% dans ceux utilisant les données CH; $p<0,05$. A la campagne 2021-2022, ces proportions étaient respectivement de : 12,5%; 16,0%; 13,1% et 12,7%; $p<0,05$. Le TBP (Campagnes 2016-2017 : Aucun :31,2%; CPAM : 31,7%; RC :33,5%; CH :31,8%; $p<0,05$. Campagne 2021-2022 : Aucun :33,1%; CPAM :29,7% ; RC :37,2% ; CH :32,0% ; $p<0,05$) a varié dans le même sens que le TBE (***Tableau-5***).

La totalité des départements ayant participé à l'enquête étaient favorables à une demande collective des CRCDC pour l'accès aux données de la CNAM pour l'optimisation des exclusions médicales. Quelle que soit la

campagne, les TBE et TBP étaient plus élevés que la moyenne (Global) dans les départements qui recevaient les données des structures pourvoyeuses et dans les départements ayant été pilotes du DOCCR (**Figure-1**).

Seuls 7 SGDO (31,8% des 22 recevant des données) recevaient les données dans le cadre d'une convention formalisée : l'Indre-et-Loire-37 et l'Essonne-91 avaient une convention avec les CPAM (**Tableau-2**) ; le Doubs-25 et la Saone-et-Loire-71 avec un registre de cancer (**Tableau-3**) ; l'Ardennes-08, la Moselle-57, la Saone-et-Loire-71 et La Réunion-974 avec un centre hospitalier (**Tableau-4**). Les contextes et fréquences de transmission des données variaient d'un département à l'autre. Dans l'Essonne où la fréquence de transmission des données était la plus élevée, il s'agissait d'un échange de données entre la SGDO et la CPAM : « *Nous avons avec la CPAM un protocole : chaque mois nous leur déposons sur leur plateforme d'échange sécurise de fichiers "Pétra", un fichier des assurés invités et relancés du département affiliés au régime général afin qu'ils puissent envoyer un mail d'invitation renforcé aux assurés après avoir vérifiés (en matchant avec leur fichier des actes) qu'ils n'avaient pas eu de coloscopie dans les 2 ans. Ils nous déposent en retour un fichier avec tous les invités ayant eu une coloscopie dans les 2 ans avec la date de la coloscopie* ». En Indre-et-Loire, la SGDO recevait 4 fois par an, entre 2016 et 2020, les données de la CPAM, de la MSA(Mutualité sociale agricole) et du RSI (Régime sociale des indépendants), dans le cadre d'un projet de recherche. Dans les 4 départements de la Franche-Comté (Doubs, Jura, Haute-Saône, Territoire-de-Belfort) les données hospitalières avaient été transmises dans le cadre d'une étude expérimentale de suivi des personnes à risque élevé de CCR.

Discussion

Cette étude met en évidence la disparité dans l'accès des SGDO à des données supplémentaires pour qualifier les exclusions médicales. Bien qu'il ait existé dans chaque région au moins un département ayant accès à au moins une des sources de données, la régionalisation du DOCCR n'a pas corrigé cette disparité intrarégionale dans l'accès aux données des principales structures pourvoyeuses de données individuelles. Il n'y a aucune région dans laquelle tous les départements recevaient des données des structures pourvoyeuses, pour l'optimisation des exclusions médicales. Les SGDO recevaient des données des structures (CPAM, Registres de cancers, Centre Hospitalier) sans qu'un partenariat n'ait été formalisé pour la plupart. La fréquence de transmission des données était faible quand la demande d'accès avait été faite pour des besoins d'étude, plus forte quand une convention d'échange avait été formalisée entre la SGDO et la structure pourvoyeuse de données individuelles.

L'enquête met en évidence, l'irrégularité dans la transmission des données d'exclusion médicale. A l'exception de trois départements (Isère, Rhône, Réunion) l'accès entre 2010 et 2022 a été accordé aux SGDO sur une durée limitée, tardivement (après 2015 pour la plupart) et très souvent dans le cadre d'un projet d'étude. Malgré cette irrégularité dans l'accès, les départements ayant reçu les données individuelles d'au moins l'une des trois structures pourvoyeuses ciblées dans cette étude, avaient un taux de participation supérieur à la moyenne globale, à l'exception de 5 départements (Hérault, Nord, Somme, 2-Corse, Réunion).

Il est avéré qu'un taux de participation estimé à 45% dans une population cible sans exclure du dénominateur, des potentiels cas d'inéligibilité, passerait à 50% (gain de 5 points) si 10% de la population était classée inéligible et soustraite du dénominateur [14]. Malgré le faible nombre de départements ayant accès aux sources de données et l'irrégularité dans l'accès, l'augmentation du taux de participation est corrélée positivement avec le taux d'exclusion médicale dans cette étude. En effet en douze années de campagnes, les taux d'exclusion et de participation ont augmenté respectivement de 2 points et 3 points.

Après plus de dix ans de pratique du DOCCR en France, la collecte des informations permettant l'exclusion médicale se fait exclusivement à l'aide des réponses postales dans la majorité des départements. Pour les rares départements recevant des données des structures pourvoyeuses de données individuelles, la faible fréquence de mise à disposition des données d'exclusion ne leur permettait pas d'atteindre un niveau optimal dans les exclusions médicales. En rappel, le taux d'exclusion médicale (France entière, campagne 2016-2017: 12,9%) était largement sous-estimé car 20% des personnes âgées de 50-74 ans avaient eu une coloscopie dans les 5 dernières années[15]. Le nombre de coloscopies réalisées sur une année civile en France, a été estimé à 1

402 047 (en 2016) dont 76,9% chez les personnes âgées de ≥50 ans[16]. Cette proportion de coloscopie dans la tranche d'âge cible du dépistage, converge avec le fort TBE dans les trois départements ayant une forte fréquence de transmission des données d'actes de coloscopies par leurs CPAM (Hérault, Indre-et-Loire) ou recevant pendant longtemps (Isère) les données d'actes de coloscopie de leur CPAM.

L'enquête révèle qu' aucun département n'utilisait les données du SNDS, bien qu'ils soient tous favorables à leur utilisation pour affiner la population cible des campagnes. Rappelons que les données des CPAM n'ont qu'une profondeur de 5 ans et ne sont pas informatives sur l'ensemble des facteurs morbides pouvant justifier d'une exclusion médicale, notamment les diagnostics (CCR, MICI). Dans le contexte actuel du DOCCR, l'usage du SNDS permettrait d'attendre un taux minimum d'exclusion nettement supérieur au taux actuel car le SNDS contiendrait la totalité des situations morbides pouvant justifier d'une exclusion médicale. Cet apport potentiel du SNDS est largement justifié par le fort TBE (25%) observé en Indre-et-Loire, département ayant accès aux données d'actes de coloscopies.

A l'exception des actes de coloscopie, aucune SGDO n'avait reçu des hôpitaux les informations en rapport avec les autres facteurs morbides justifiant d'une exclusion médicale (diagnostics, coloscanner…) qui sont pourtant codés dans les bases hospitalières du PMSI (Programme de médicalisation des systèmes d'information). La collecte de données sur tous les ces facteurs morbides pouvant justifier d'une exclusion médicale du DOCCR permettrait d'atteindre un taux d'exclusion minimal largement supérieur au taux actuel. En rappel en France, la proportion de personnes prises en charge pour MICI était de 4‰ en 2018 ; celle pour CCR (actif ou sous surveillance) était de 5,3‰ [17].

L'usage du SNDS dans le DOCCR permettrait aussi d'amorcer la véritable réflexion sur l'évolution du DOCCR. Il a été argumenté que l'utilisation d'un scores de risque de CCR tenant compte de certains facteurs permettrait d'améliorer les pratiques de prévention par l'adaptation des modalités du dépistage selon le niveau de risque [18]. En plus des facteurs de risque personnels et familiaux actuellement pris en compte pour les exclusions du DOCCR, la plupart des facteurs de risque pouvant faciliter cette réadaptation des modalités du DOCCR sont accessibles par le SNDS (traitement médicamenteux d'affection chroniques, obésité morbide, traitement hormonal de la ménopause…). L'usage des données du SNDS permettrait ainsi d'élaborer ce score de risque afin de l'expérimenter dans le DOCCR. En exemple, la consommation régulière d'aspirine [19] et le traitement hormonal substitutif CCR[20] sont associés à une diminution du risque de CCR. Une expérimentation de la

fréquence de réalisation d'un test de dépistage modulée par la consommation de ces médicaments ou d'autres, est envisageable si un accès aux données de consommation médicamenteuse (SNDS) est accordé aux SGDO.

L'enquête met en évidence une faible utilisation des registres de cancers comme moyens supplémentaires (seulement 5 départements). Le manque d'implication des registres de cancer est certes un obstacle majeur à l'amélioration de l'efficacité du programme mais ce n'est pas un fait nouveau car il a été révélé qu'ils ne sont pas impliqués dans l'évaluation du dépistage dans plusieurs pays européens[21]. Il faut cependant rappeler que les registres de cancer ne couvrent que 20% de la population française[22] avec de grandes régions (l'Ile-de-France) n'ayant aucun registre de cancer.

Conclusion

Il existe une disparité dans l'accès des SGDO à des données supplémentaires permettant de qualifier les exclusions médicales. Cette disparité n'a pu être corrigée par la régionalisation du DOCCR car il n'y a aucune région dans laquelle tous les départements avaient accès à un moyen supplémentaire de repérage des potentiels cas justifiant d'une exclusion médicale. Les SGDO recevaient des données des structures (CPAM, Registres de cancers, Centre Hospitalier) sans qu'un partenariat n'ait été formalisé pour la plupart.

Exception pour trois départements (Isère, Rhône, Réunion), les SGDO ont reçu les données entre 2010 et 2022 sur une durée limitée, La démarche de transmission a été mise en place tardivement (après 2015 pour la plupart) et le plus souvent dans le cadre d'un projet d'étude. Malgré cette irrégularité, les départements qui utilisaient les données fournies par au moins l'une des trois structures pourvoyeuses de données individuelles, avaient des taux d'exclusion et de participation supérieurs aux moyennes globales. L'usage du SNDS permettrait d'attendre un taux minimum d'exclusion qui soit nettement supérieur au taux actuel.

Remerciements

Les auteurs remercient les médecins coordonnateurs et tous les correspondants ayant répondu à l'enquête d'exclusion médicale.

Conflit d'intérêt

Les auteurs n'ont aucun conflit d'intérêt à déclarer

Références

1. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA: a cancer journal for clinicians. 2009;59(6):366-78. Epub 2009/11/10. doi: 10.3322/caac.20038. PMID: 19897840.
2. Altobelli E, Lattanzi A, Paduano R, Varassi G, di Orio F. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. Preventive medicine. 2014;62:132-41. doi: 10.1016/j.ypmed.2014.02.010. PMID: 24530610.
3. JORF. Arrêté du 19 mars 2018 modifiant l'arrêté du 29 septembre 2006 relatif aux programmes de dépistage organisé des cancers et portant modification du cahier des charges du dépistage organisé du cancer colorectal. JORF n° 0068 du 22 mars 2018 texte n° 13. Journal Officiel de la République Française (JORF). 2018;0068:NOR: SSAP1807442A
4. Leuraud K, Jezewski-Serra D, Viguier J, Salines E. Colorectal cancer screening by guaiac faecal occult blood test in France: Evaluation of the programme two years after launching. Cancer epidemiology. 2013;37(6):959-67. doi: 10.1016/j.canep.2013.07.008. PMID: 24035240.
5. Tazi MA, Faivre J, Dassonville F, Lamour J, Milan C, Durand G. Participation in faecal occult blood screening for colorectal cancer in a well defined French population: results of five screening rounds from 1988 to 1996. Journal of medical screening. 1997;4(3):147-51.
6. Moss SM, Campbell C, Melia J, Coleman D, Smith S, Parker R, et al. Performance measures in three rounds of the English bowel cancer screening pilot. Gut. 2012;61(1):101-7. doi: 10.1136/gut.2010.236430. PMID: 21561880.
7. Pornet C, Denis B, Perrin P, Gendre I, Launoy G. Predictors of adherence to repeat fecal occult blood test in a population-based colorectal cancer screening program. British journal of cancer. 2014;111(11):2152-5. doi: 10.1038/bjc.2014.507. PMID: 25314056.
8. Pollack CE, Mallya G, Polsky D. The impact of consumer-directed health plans and patient socioeconomic status on physician recommendations for colorectal cancer screening. J Gen Intern Med. 2008;23(10):1595-601. doi: 10.1007/s11606-008-0714-x. PMID: 18629590.
9. Guignot C. Cancer colorectal. Vers un dépistage mieux accepté. Science et Santé. 2015;26:36-7.
10. Le Pimpec F, Moutel G, Piette C, Lievre A, Bretagne JF. Fecal immunological blood test is more appealing than the guaiac-based test for colorectal cancer screening. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2017;49(11):1267-72. doi: 10.1016/j.dld.2017.08.018. PMID: 28867474.

11. SPF. Taux de participation au programme de dépistage organisé du cancer colorectal 2018-2019. Saint Maurice, France: Santé Publique France(SPF), 2020.
12. JORF. Arrêté du 29 septembre 2006 relatif aux programmes du dépistage des cancers. Journal Officiel de la République Française. Annexe au n° 295.32003-48. 2006.
13. Délibération n° 2017-215 du 13 juillet 2017 portant adoption d'une norme destinée à simplifier l'obligation de déclaration des traitements de données à caractère personnel ayant pour finalité le dépistage organisé du cancer du sein, du cancer colorectal et du cancer du col de l'utérus mis en œuvre par les structures de gestion conventionnées, et abrogeant la délibération n° 2015-175 du 11 juin 2015 (décision d'autorisation unique n° AU-043) (NS-059) NOR: CNIL1724568X (2017).
14. Bulliard JL, Garcia M, Blom J, Senore C, Mai V, Klabunde C. Sorting out measures and definitions of screening participation to improve comparability: the example of colorectal cancer. Eur J Cancer. 2014;50(2):434-46. doi: 10.1016/j.ejca.2013.09.015. PMID: 24144735.
15. Seitz JF, Lapalus D, Arlotto S, Gentile S, Ettori F, Rinaldi Y, et al. Colorectal cancer screening by fecal immunochemical test or colonoscopy in France: how many people are actually covered? Focus on the Provence-Alpes-Cote d'Azur region. European journal of gastroenterology & hepatology. 2022;34(4):405-10. doi: 10.1097/MEG.0000000000002338. PMID: 34882643.
16. Bernardini D, Bulois P, Barthet M, Chaussade S, Gronier O, Lecomte T, et al. « Une semaine de coloscopie en France » : résultats 2017 de l'enquête annuelle de la Société française d'endoscopie digestive. Acta Endoscopica. 2017;47(5):242-51. doi: 10.1007/s10190-017-0613-2.
17. CNAM. Cartographie des pathologies et des dépenses: Prévalence départementale par pathologie (2018). Caisse Nationale d'Assurance Maladie(CNAM): Direction de la Stratégie, des Etudes et des Statistiques-Département des Etudes sur les Pathologies et les Patients. 2020.
18. Robaszkiewicz M. Intérêt des scores pour la prédiction du risque de cancer colorectal. In: Le Grand Metier, eds. Dépistage et prévention des cancers colorectaux. ROUEN, FRANCE 2018. p. 99-101.
19. Benamouzig R, Uzzan B. Aspirin to prevent colorectal cancer: time to act? Lancet 2010;376(9754):1713-4. doi: 10.1016/S0140-6736(10)61509-7. PMID: 20970845.
20. Delluc C, Benamouzig R. [Hormone replacement therapy and colorectal cancer]. Gastroenterol Clin Biol. 2008;32(1 Pt. 1):19-28. doi: 10.1016/j.gcb.2007.12.008. PMID: 18341973.

21. Anttila A, Lonnberg S, Ponti A, Suonio E, Villain P, Coebergh JW, et al. Towards better implementation of cancer screening in Europe through improved monitoring and evaluation and greater engagement of cancer registries. Eur J Cancer. 2015;51(2):241-51. doi: 10.1016/j.ejca.2014.10.022. PMID: 25483785.

22. Tron L, Belot A, Fauvernier M, Remontet L, Bossard N, Launay L, et al. Socioeconomic environment and disparities in cancer survival for 19 solid tumor sites: An analysis of the French Network of Cancer Registries (FRANCIM) data. Int J Cancer. 2019;144(6):1262-74. doi: 10.1002/ijc.31951. PMID: 30367459.

Figure

Click here to access/download;Figure;FIGURE_1_ENQUETE_EXCLUSION.tif

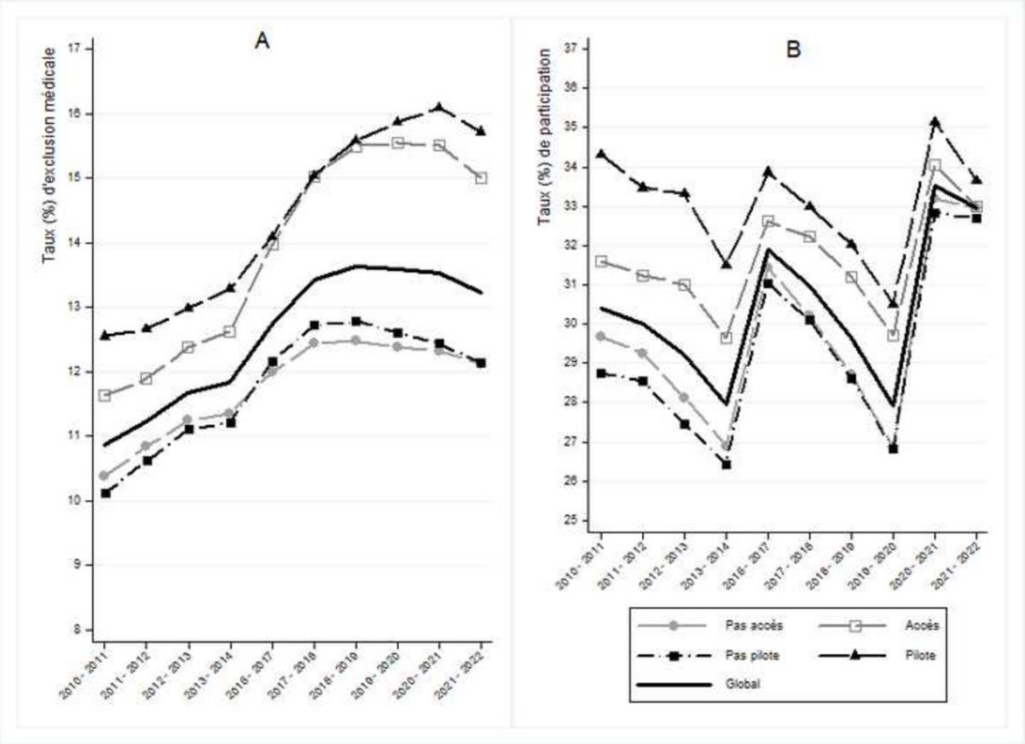


Figure 1 : Evolution du taux brut d'exclusion médicale (A) et du taux brut de participation (B) selon l'accès à un moyen lors de la campagne (Pas accès vs. Accès) et selon l'antériorité du département dans le DOCCR (Pas Pilote vs. Pilote) , entre 2010 et 2022. La modalité « Global » est la moyenne des 60 départements .

Tableau 1 : Taux d'exclusion et taux de participation à la campagne biannuelle 2021-2022 dans les départements inclus.

Département Nom (Numéro)	Sources de données d'exclusion médicale				Taux d'exclusion médicale à la campagne biannuelle 2021-2022		Taux de participation à la campagne biannuelle 2021-2022	
	Nb de sources	Principaux types de source			Taux brut (%)†	Gain (en point) en comparaison à 2013-2014 †	Taux brut (%)†	Gain (en point) en comparaison à 2013-2014†
		CPAM	RC	CH				
Ain (01)	0				11,8	**-0,3**	35,6	5,4
Aisne (02)	0				13,6	0,0	29,7	**-2,0**
Allier (03)*	1	1			13,8	**-1,4**	36,0	**-0,9**
Ardèche (07)	0				11,8	1,9	31,5	3,8
Ardennes (08)*	1			1	12,8	0,2	34,5	**-4,5**
Ariège (09)	0				10,0	0,5	28,5	3,4
Aude (11)	0				12,3	2,3	30,6	9,3
Aveyron (12)	0				14,6	1,0	32,5	6,7
Cantal (15)	0‡				9,0	-2,0	32,9	3,6
Cher (18)	0				21,0	6,3	36,2	**-1,8**
Doubs (25)	3	1	1	1	13,1	0,8	37,2	2,9
Drôme (26)	0				11,4	2,0	35,1	6,3
Eure (27)	0				12,2	0,9	36,3	5,1
Eure-et-Loir (28)	0				13,1	-0,7	33,8	1,5
Gard (30)	0				12,5	1,5	30,6	8,2
Haute-Garonne (31)	0				13,7	1,6	32,1	9,9
Gers (32)	0				15,6	1,6	35,3	7,1
Hérault (34)*	2	1	1		19,9	10,3	29,4	7,7
Indre (36)	0				15,8	0,5	33,0	2,8
Indre-et-Loire (37)*	1	1			25,2	6,9	41,0	3,9
Isère (38)*	3	1	1	1	18,8	3,3	40,8	**-0,8**
Jura (39)	2	1		1	11,7	1,2	36,4	1,5
Loir-et-Cher (41)	0				20,1	0,6	39,1	**-2,9**
Loire (42)	0				11,8	**-1,5**	32,9	1,1
Haute-Loire (43)	0‡				11,3	**-3,1**	33,1	3,6
Loiret (45)	0				12,9	0,5	35,9	0,7
Lot (46)	0				15,3	1,6	34,4	0,7
Lozère (48)	0				11,7	1,3	28,4	8,5
Marne (51)*	0‡‡				21,7	3,4	38,8	0,6
Haute-Marne (52)	0				10,1	2,0	37,3	7,0
Meuse (55)	1			1	14,2	1,2	38,0	6,5
Moselle (57)*	1			1	10,9	0,4	34,5	3,7
Nord (59)*	1	1			14,8	1,5	28,2	2,0
Oise (60)	1	1			12,5	**-0,3**	27,5	1,2
Pas-de-Calais (62)	0				12,6	0,9	30,1	2,8
Puy-de-Dôme (63)*	0‡				12,6	**-1,3**	36,1	2,8
Hautes-Pyrénées (65)	0				10,5	1,4	32,2	2,9
Pyrénées-Orientales (66)*	0				15,2	0,8	37,0	**-0,3**
Rhône (69)	1‡			1	13,1	0,0	32,7	5,8
Haute-Saône (70)	2	1		1	12,3	0,4	33,1	1,5
Saône-et-Loire (71)*	2‡		1	1	14,6	**-1,3**	41,6	**-6,0**
Savoie (73)	0				14,3	1,6	42,8	5,6
Haute-Savoie (74)	0‡‡				10,8	2,8	41,1	15,2
Paris (75)	0				6,6	**-1,3**	27,3	13,5
Seine-et-Marne (77)	0				11,7	**-1,3**	35,2	6,2
Yvelines (78)	0				11,5	0,4	33,3	11,3
Somme (80)	1	1			14,2	0,0	30,8	1,8
Tarn (81)	1		1		15	1,8	34,9	2,7
Tarn-et-Garonne (82)	1	1			12,4	**-1,2**	35,5	1,7
Haute-Vienne (87)	0				13,3	2,8	34,1	3,3
Yonne (89)	0				11,2	1,2	35,0	**-5,2**
Territoire-de-Belfort (90)	2	1		1	12,3	0,9	34,7	3,8
Essonne (91)*	1	1			16,9	4,6	34,9	9,2
Hauts-de-Seine (92)	0				9,6	2,3	30,3	12,3
Seine-Saint-Denis (93)*	0				10,1	**-0,3**	27,1	**-0,3**
Val-de-Marne (94)	0				11,4	2,0	32,3	8,0
Val-d'Oise (95)	0				11,9	1,0	33,5	8,0
La Réunion (974)	1			1	11,8	5.0	29,7	6,1
Corse-du-Sud (2A/2B)**	1	1			13,2	8,4	18,1	11,9

†Les taux ont été estimés à partir des données officielles publiées par Santé Publique France ; *Départements pilotes du DOCCR en France ; **SPF regroupe les données des deux départements de la Corse ‡Compte-rendu d'anatomo-cyto-pathologie sans rapport avec le dépistage ; ‡‡Compte-rendu de coloscopie de diagnostic (hors test de dépistage positif) ; CH : Centre hospitalier ; CPAM : Caisse primaire d'assurance maladie ; Nb : Nombre ; RC : Registre de cancers.

Tableau

Tableau 2 : Analyse de l'impact de l'accès aux données des CPAM pour qualifier l'exclusion médicale dans les 14 départements ayant reçu les données d'une CPAM sur au moins une campagne entre 2010 et 2022.

Département (N°)	Accès aux données de la CPAM					Évolution des indicateurs entre périodes selon l'accès aux données des CPAM							
	Début	Fin	Fréq.	Type	Autres sources	Avant l'accès		Pendant l'accès		Après l'accès		Impact	
						TBE (%)	TBP (%)	TBE (%)	TBP (%)	TBE (%)	TBP (%)	TBE (p1/p2)	TBP (p1/p1)
Allier (03)	≤2015	2018	NC	A , B	-	†	†	16,9	38,0	13,8	36,0	NA/**	NA/**
Doubs (25)	2017	2018	1	A , B	RC, CH	12,3	34,3	14,4	40,2	13,1	37,2	**/**	**/**
Hérault (34)*	2018	†	≥5	A , B	RC	7,4	22,1	19.9	29,4	†	†	**/NA	**/NA
Indre-et-Loire (37)*	2016	2020	4 ‡	A , B	-	18,3	40,2	25,0	40,3	25,2	41,0	**/NS	NS/NS
Isère (38)*	≤2015	†	NC	A , B	RC, CH	†	†	19,7	45,9	†	†	NA/NA	NA/NA
Jura (39)	2017	2018	1	A , B	CH	10,5	38,2	12,5	39,3	11,7	36,4	**/NA	**/**
Nord (59)*	2016	†	2	A , B	-	13,3	28,2	15,8	29,4	†	†	**/NA	**/NA
Oise (60)	2018	†	4	A , B	-	13,3	27,5	14,0	29,7	†	†	NS/NA	**/NA
Haute-Saône (70)	2017	2018	1	A , B	CH	11,9	34,5	14,0	35,8	12,3	33,1	**/**	**/**
Somme (80)	≤2015	2018	2	A , B	-	†	†	17,3	31,6	14,2	30,8	NA/**	NA/NS
Tarn-et-Garonne (82)	≤2015	2017	3	A , B	-	†	†	14,7	37,2	12,4	35,5	NA/**	NA/**
Territoire-de-Belfort (90)	2017	2018	1	A , B	CH	11,4	33,5	13,9	36,1	12,3	34,7	**/**	**/**
Essonne (91)*	2019	†	≥5‡	A , B	-	14,7	33,4	17,6	37,5	†	†	**/NA	**/NA
Corse (2AB)	2019	2020	1	A , B	-	8,8	15,5	14,1	17,2	13,2	18,1	**/NS	**/NS

*Département pilote ; ‡Partenariat formalisé avec la CPAM ; †Accès en cours au moment de l'enquête en 2021 ou sur la période analysée ; **A** : Acte de coloscopie ; B : Date de l'acte ; **CH :** Centre hospitalier ; **Fréq**. : Fréquence ; **NA** : Non applicable; **NC** : Fréquence indéfinie ou non communiquée ; **NS** : Test du Khi2 non significatif ; **RC** : Registre de cancers ; **TBE** : Taux brut d'exclusion médicale ; **TBP** : Taux brut de participation ; **Test du Khi2 de tendance** : p1(Avant vs. Pendant), p2(Pendant vs. Après), **p<0,05.

Tableau

Tableau 3 : Analyse de l'impact de l'accès aux données des registres de cancers (RC) pour qualifier l'exclusion médicale dans les 5 départements ayant reçu les données d'un registre de cancers sur au moins une campagne entre 2010 et 2022.

Département (N°)	Accès aux données d'un registre de cancers						Évolution des indicateurs entre périodes selon l'accès aux données d'un registre de cancers							
	Registre	Début	Fin	Fréq.	Type	Autre source	Avant l'accès		Pendant l'accès		Après l'accès		Impact	
							TBE (%)	TBP (%)	TBE (%)	TBP (%)	TBE (%)	TBP (%)	TBE (p1/p2)	TBP (p1/p1)
Doubs (25)	R1	2017	†	≥5‡	A, B, E	CPAM, CH	12,3	34,3	14,4	40,2	†	†	**/NA	**/NA
Hérault (34)*	R2	≤2015	2019	NC	A , B	CPAM	†	†	15,5	26,0	19,9	29,4	NA/**	NA/**
Isère (38)*	R3	≤2015	†	NC	A , B	CPAM, CH	†	†	19,7	45,9	†	†	NA/NA	NA/NA
*Saône-et-Loire (71)**	R4	≤2015	NC	NC ‡	A , B	CH	†	†	17,7	52,4	†	†	NA/NA	NA/NA
Tarn (81)	R5	2016	2018	1	A , B	-	13,2	34,2	15,5	33,5	15,0	34,9	**/NS	**/**

*Département pilote ; ‡Partenariat formalisé avec un registre ; †Au moment de l'enquête en 2021; **A :** Diagnostic CCR ; **B :** Date de diagnostic CCR ; **CH :** Centre hospitalier ; **CPAM** : Caisse primaire d'assurance maladie ; **E :** Diagnostic autres cancers ; **Fréq. :** Fréquence annuelle ; **NA** : Non applicable; **NC** : Fréquence indéfinie ou non communiquée ; **NS** : Test du Khi2 non significatif ; **R1 :** Registre des tumeurs du Doubs ; **R2 :** Registre général des tumeurs de l'Hérault ; **R3 :** Registre général des cancers de l'Isère ; **R4 :** Registre Bourguignon des cancers digestifs ; **R5 :** Registre général des cancers du Tarn ; **TBE** : Taux brut d'exclusion médicale ; **TBP** : Taux brut de participation ; **Test du Khi2 de tendance** : **p1**(Avant vs. Pendant), **p2**(Pendant vs. Après), **p<0,05.

Tableau

Tableau 4 : Analyse de l'impact de l'accès aux données des Centres Hospitaliers pour qualifier l'exclusion médicale dans les 11 départements ayant reçu les données d'un Centre Hospitalier sur au moins une campagne entre 2010 et 2022.

Département (N°)	Accès aux données des Centre Hospitaliers					Évolution des indicateurs entre périodes selon l'accès aux données des Centres Hospitaliers							
	Début	Fin	Fréq.	Type	Autres sources	Avant l'accès		Pendant l'accès		Après l'accès		Impact	
						TBE (%)	TBP (%)	TBE (%)	TBP (%)	TBE (%)	TBP (%)	TBE (p1/p2)	TBP (p1/p1)
Ardennes (08)	2020	†	4‡	A , B	-	12;6	43,2	12,8	34,5	†	†	NS/NA	**/ NA
Doubs (25)	2017	2018	NC	A , B	CPAM, RC	12,3	34,3	14,4	40,2	13,1	37,2	**/**	**/**
Isère (38)*	≤2015	†	NC	A , B	CPAM, RC	†	†	19,7	45,9	†	†	NA/NA	NA/NA
Jura (39)	2017	2018	NC	A , B	CPAM	10,5	38,2	12,5	39,3	11,7	36,4	**/NS	**/**
Meuse (55)	≤2015	2018	2	A , B	-	†	†	17,0	37,1	14,2	38,0	NA/**	NA/NS
Moselle (57)	≤2015	2018	≥5‡	A , B	-	†	†	12,1	35,0	10,9	34,5	NA/**	NA/NS
Rhône (69)	≤2015	†	NC	A , B	-	†	†	16,6	32,7	†	†	NA/NA	NA/NA
Haute-Saône (70)	2017	2018	NC	A , B	CPAM	11,9	34,5	14,0	35,8	12,3	33,1	**/**	**/**
Saône-et-Loire (71)*	≤2015	NC‡	NC	A , B	RC	†	†	17,7	52,4	†	†	NA/NA	NA/NA
Territoire-de-Belfort (90)	2017	2018	NC	A , B	CPAM	11,4	33,5	13,9	36,1	12,3	34,7	**/**	**/**
La Réunion (974)	≤2015	†	1‡	A , B		†	†	12,9	29,7	†	†	NA/NA	NA/NA

*Département pilote ; ‡Partenariat formalisé avec un centre hospitalier ; †Accès en cours au moment de l'enquête en 2021 ou sur la période analysée ; **A** : Acte de coloscopie ; B : Date de l'acte ; **Fréq.** : Fréquence ; **NA** : Non applicable; **NC** : Fréquence indéfinie ou non communiquée ; **NS** : Test du Khi2 non significatif ; **RC** : Registre de cancers ; **TBE** : Taux brut d'exclusion médicale ; **TBP** : Taux brut de participation ; **Test du Khi2 de tendance** : p1(Avant vs. Pendant), p2(Pendant vs. Après), **p<0,05.

Tableau

Tableau 5 : Évolution des taux d'exclusion médicale et taux de participation en fonction des moyens d'exclusion médicale, campagnes biannuelle 2010 à 2022.

Campagne	Aucun	CPAM	RC	CH	CPAM/RC	CPAM/CH	RC/CH	CPAM/RC/CH	p	Global
2010 – 2011 : Nb (N°) de Dép.	**49**	**3 (03, 80, 82)**	**1 (34)**	**4 (55, 57, 69, 974)**	**0**	**0**	**1 (71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	911803 (10,6)	40074 (11,9)	26704 (9,0)	97155 (10,3)	0	0	26328 (15,2)	49164 (15,6)	*	1151228 (10,9)
Nb (TBP†) de participation	2272161 (29,8)	99553 (33,6)	62883 (23,2)	244709 (28,9)	0	0	77306 (52,4)	118123 (44,3)	*	2874735 (30,4)
2011 - 2012 : Nb (N°) de Dép.	**49**	**3 (03, 80, 82)**	**1 (34)**	**4 (55, 57, 69, 974)**	**0**	**0**	**1 (71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	961033 (11,1)	43821 (12,9)	28508 (9,4)	100538 (10,4)	0	0	27253 (15,5)	48950 (15,2)	*	1210103 (11,2)
Nb (TBP†) de participation	2265909 (29,4)	98586 (33,2)	59690 (21,6)	250679 (29,0)	0	0	76154 (51,3)	118689 (43,5)	*	2869707 (30,0)
2012 - 2013 : Nb (N°) de Dép.	**49**	**3 (03, 80, 82)**	**1 (34)**	**4 (55, 57, 69, 974)**	**0**	**0**	**1 (71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	1014997 (11,5)	48221 (13,9)	29267 (9,4)	107121 (10,9)	0	0	28280 (15,9)	50216 (15,3)	*	1278102 (11,7)
Nb (TBP†) de participation	2213637 (28,4)	99669 (33,5)	65995 (23,3)	252253 (28,8)	0	0	74901 (50,1)	118745 (42,7)	*	2825200 (29,2)
2013 - 2014 : Nb (N°) de Dép.	**49**	**3 (03, 80, 82)**	**1 (34)**	**4 (55, 57, 69, 974)**	**0**	**0**	**1 (71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	1045088 (11,7)	50551 (14,4)	30663 (9,6)	111112 (11,1)	0	0	28685 (15,9)	51994 (15,5)	*	1318093 (11,8)
Nb (TBP†) de participation	2145772 (27,2)	97743 (32,6)	62667 (21,7)	247260 (27,7)	0	0	72004 (47,6)	117819 (41,6)	*	2743265 (27,9)
2016 - 2017 : Nb (N°) de Dép.	**42**	**5 (03, 37, 59, 80, 82)**	**2 (34, 81)**	**4 (55,57,69,974)**	**0**	**3 (39, 70, 90)**	**1 (71)**	**2 (25, 38)**		**59**
Nb (TBE†) d'exclusion	959475 (12,0)	201134 (16,1)	42907 (14,0)	147667 (13,8)	0	25 519 (12,5)	30680 (16,4)	81 750 (16,1)	*	1489132 (12,7)
Nb (TBP†) de participation	2199975 (31,2)	332459 (31,7)	60972 (33,5)	292686 (31,8)	0	66 681 (37,3)	74452 (47,8)	179 547 (42,2)	*	3252192 (31,9)
2017 - 2018 : Nb (N°) de Dép.	**41**	**6 (03,37,59,60,80, 82)**	**1 (81)**	**4 (55, 57, 69, 974)**	**1 (34)**	**3 (39, 70, 90)**	**1 (71)**	**2 (25, 38)**		**59**
Nb (TBE†) d'exclusion	982332 (12,5)	251825 (16,7)	19083 (14,9)	155915 (14,4)	32481 (9,3)	27 549 (13,3)	32853 (17,4)	88 213 (17,2)	*	1590251 (13,4)
Nb (TBP†) de participation	2078637 (30,2)	387004 (30,9)	35884 (32,8)	282950 (30,5)	77884 (24,7)	66 074 (36,9)	71254 (45,8)	176 092 (41,3)	*	3175779 (31,0)
2018 - 2019 : Nb (N°) de Dép.	**40**	**7 (03, 37, 59, 60,80,91,2AB)**	**1 (81)**	**4 (55, 57, 69, 974)**	**1 (34)**	**3 (39, 70, 90)**	**1 (71)**	**2 (25, 38)**		**59**
Nb (TBE†) d'exclusion	949 239 (12,5)	310433 (16,4)	20203 (15,5)	157359 (14,3)	43491 (12,3)	27 514 (13,2)	33 598 (17,7)	93275 (17,9)	*	1635112 (13,6)
Nb (TBP†) de participation	1 917 903 (28,8)	477132 (30,1)	36660 (33,4)	288864 (30,6)	80751 (26,0)	54 493 (30,2)	53 163 (34,1)	162 772 (38,0)	*	3071738 (29,6)
2019 - 2020 : Nb (N°) de Dép.	**47**	**5 (37,59,60,91,2AB)**	**1 (25)**	**3 (08,69,974)**	**1(34)**	**0**	**1 (71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	1 086 690 (12,6)	268873 (16,6)	22385 (14,1)	111257 (13,7)	55601 (15,5)	0	32968 (17,2)	72780 (19,7)	*	1650554 (13,6)
Nb (TBP†) de participation	2 073 243 (27,5)	377132 (27,9)	45685 (33,5)	185524 (26,5)	75168 (24,7)	0	58483 (37,0)	114246 (38,5)	*	2929481 (27,9)
2020 - 2021 : Nb (N°) de Dép.	**47**	**6 (34,37,59,60,91,2AB)**	**1 (25)**	**3 (08,69,974)**	**0**	**0**	**1(71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	1 085 048 (12,5)	341187 (16,5)	21 961 (13,7)	106 727 (12,9)	0	0	30 808 (16,0)	73 411 (19,6)	*	1 659 142 (13,5)
Nb (TBP†) de participation	2 548 372 (33,4)	509105 (31,0)	54 432 (39,4)	226 245 (31,5)	0	0	75 931 (47,1)	13 7821 (45,9)	*	3 551 906 (33,5)
2021 - 2022 : Nb (N°) de Dép.	**49**	**4 (34,59,60,91)**	**1 (25)**	**3 (08,69,974)**	**0**	**0**	**1(71)**	**1 (38)**		**59**
Nb (TBE†) d'exclusion	1 130 490 (12,5)	273352 (16,0)	21 083 (13,1)	105 544 (12,7)	0	0	27 942 (14,6)	71 141 (18,8)	*	1 629 552 (13,2)
Nb (TBP†) de participation	2 623 308 (33,1)	425309 (29,7)	52 013 (37,2)	232 986 (32,0)	0	0	68 156 (41,6)	125 102 (40,8)	*	3 526 874 (33,0)

CH : Centre hospitalier ; CPAM : Caisse primaire d'assurance maladie ; Dép. : Département ; Nb : Nombre ; N° : Numéro du département ; RC : Registres de cancers ; TBE : Taux (%) brut d'exclusion médicale ; TBP : Taux (%) brut de participation ; †Les taux ont été estimés à partir des données officielles publiées par Santé Publique France ; *Test du Chi2 de Pearson, p<0,05.

e-component

Enquête transversale sur les pratiques d'exclusion médicale lors des campagnes de dépistage du cancer colorectal en France.

Supplément 1 : Le questionnaire d'enquête

Contexte : Le fait que le taux d'exclusion soit estimé seulement après un recueil de déclarations volontaires des sujets (questionnaire joint au courrier d'invitation) et des médecins traitants pourrait compromettre un recueil exhaustif des cas potentiels d'exclusion médicale. A propos de la méthode de recueil, Santé Publique France (SPF) admet que le nombre de personnes exclues est fourni par les structures de gestion à partir de diverses sources de données exploitables (réponses des personnes invitées, fichiers des actes de coloscopies remboursés par les caisses primaires d'assurance maladie, fichiers des actes diagnostiques et thérapeutiques, comptes rendus standardisés des résultats de coloscopie et des examens anatomo-cytopathologiques…), dont l'accès varie selon les structures de gestion du dépistage. Notre hypothèse est que le taux d'exclusion semble être sous-estimé dans la plupart des départements français, ce qui pourrait expliquer en partie, le faible taux de participation dans certains départements.

A l'aide de modèles de simulation, nous envisageons estimer le taux de participation à une campagne de DOCCR en fonction d'un niveau optimal du taux d'exclusion médicale.

Cette analyse sera faite sur les données officielles de la campagne biannuelle 2018-2019, consolidées et publiées par SPF.

Pour y parvenir, nous avons besoins de plus d'informations sur votre stratégie de collecte des données d'exclusion médicale pour les départements couverts par votre structure régionale.

Les résultats de cette analyse feront l'objet d'un projet de communication (orale ou affichée) lors des JFHOD-2022 voire un article scientifique.

Vous trouvez ce projet pertinent ?

Merci de répondre pour chacun de vos départements ou d'adresser le fichier aux coordonnateurs de chaque site/antenne de votre CRCDC.

Le questionnaire est simple, il consiste à cocher dans chaque carré de bonne réponse, tout en oubliant pas de préciser le numéro du département concerné en entête.

QUESTIONNAIRE

Merci d'aller à la page 2 pour répondre au questionnaire

Département N° :……..

CPAM

1 – La CPAM du département vous a-t-elle transmis entre 2015 et 2021, des données individuelles pour vous faciliter l'exclusion du DOCCR ? Non☐ Oui☐

1a - Si oui, en quelle(s) année(s)? ≤2015☐ 2016☐ 2017☐ 2018☐ 2019☐ 2020☐ 2021☐

1b - Si oui, à quelle fréquence annuelle ? 1fois☐ 2fois☐ 3fois☐ 4fois☐ ≥5fois☐

1c - Si oui, existe-t-il un partenariat formalisé par écrit ? Non☐ Oui☐

1d - Si oui, quelles données individuelles sont/étaient transmises

Acte de Coloscopie☐ Date de l'acte Coloscopie☐ Diagnostic de CCR ☐ Diagnostic de maladies chroniques inflammatoire (Crohn, Lynch, PAF, RCH…) du colon☐ Diagnostic d'autres cancers☐

REGISTRES DES CANCERS

2 – Un registre de cancer vous a-t-il transmis entre 2015 et 2021, des données individuelles pour vous faciliter l'exclusion du DOCCR ? Non☐ Oui☐ le(s) quel(s)……

……

2a - Si oui, en quelle(s) année(s)? ≤2015☐ 2016☐ 2017☐ 2018☐ 2019☐ 2020☐ 2021☐

2b - Si oui, à quelle fréquence annuelle ? 1fois☐ 2fois☐ 3fois☐ 4fois☐ ≥5fois☐

2c - Si oui, existe-t-il un partenariat formalisé par écrit ? Non☐ Oui☐

2d - Si oui, quelles données individuelles sont/étaient transmises ?

Diagnostic de CCR ☐ Date de diagnostic de CCR ☐ Adénome/Polype du colon☐ Diagnostic autres cancers☐

CENTRES HOSPITALIERS

3 – Un DIM d'un CHU/CH vous a-t-il transmis entre 2015 et 2021, des données individuelles pour vous faciliter l'exclusion du DOCCR ? Non☐ Oui☐ le(s) quel(s)……

……

3a - Si oui, en quelle(s) année(s)? ≤2015☐ 2016☐ 2017☐ 2018☐ 2019☐ 2020☐ 2021☐

3b - Si oui, à quelle fréquence annuelle ? 1fois☐ 2fois☐ 3fois☐ 4fois☐ ≥5fois☐

3c - Si oui, existe-t-il un partenariat formalisé par écrit ? Non☐ Oui☐

3d - Si oui, quelles données individuelles sont/étaient transmises

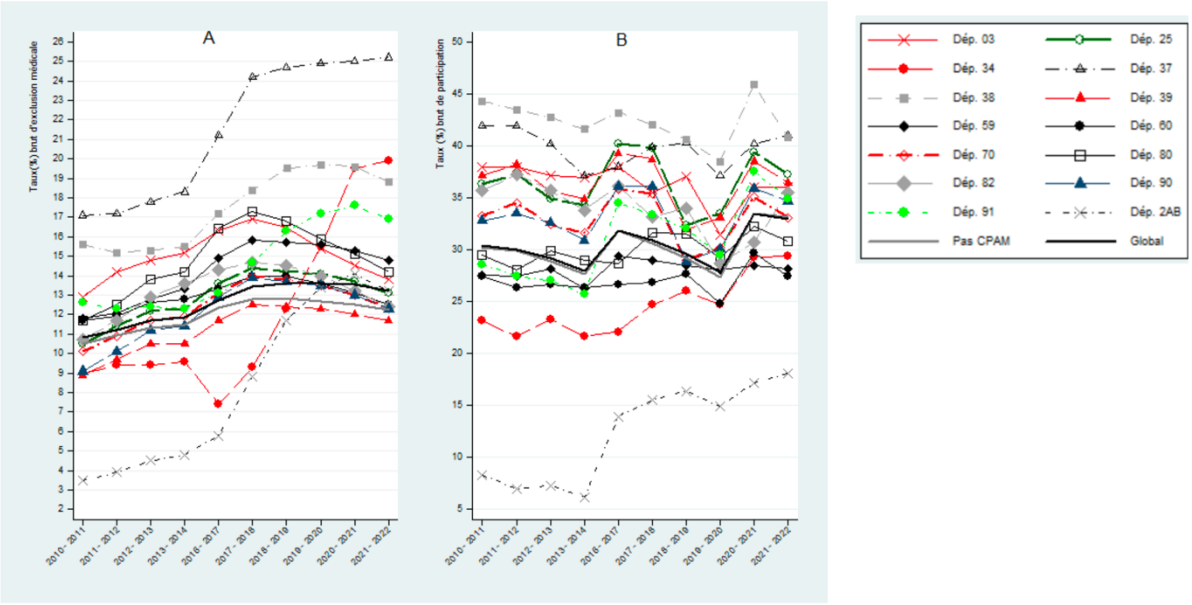
Acte de Coloscopie☐ Date de l'acte Coloscopie☐ Diagnostic de CCR ☐ Diagnostic de maladies chroniques inflammatoire (Crohn, Lynch, PAF, RCH…) du colon☐ Diagnostic d'autres cancers☐

CONCLUSION

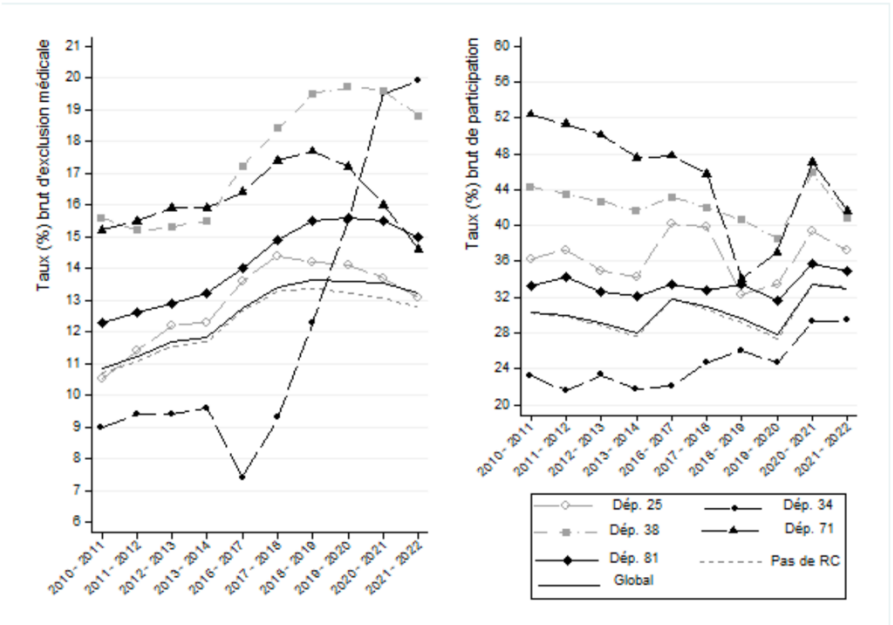
Saviez-vous que les données d'exclusions peuvent être obtenues par requête dans les bases de données médico-administratives (PMSI, ALD, CEPIDC, …) accessibles à la CNAM ? Non☐ Oui☐

Etes-vous favorables à une demande collective des CRCDC pour l'accès aux données de la CNAM ? Non☐ Oui☐

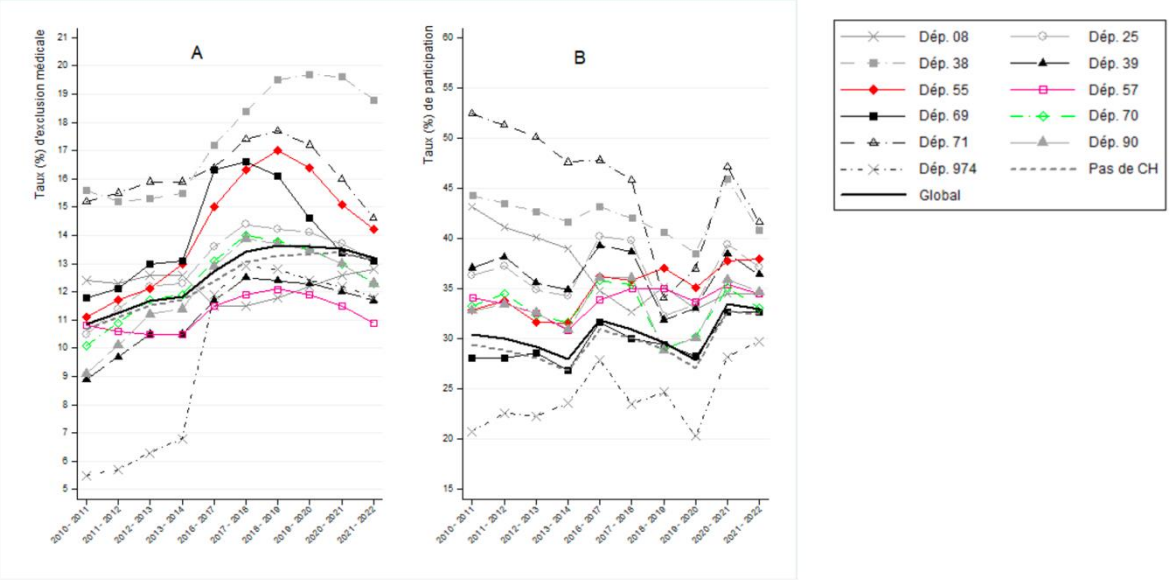
Merci de vous limiter ici si vous répondez pour un seul département. Continuez si vous répondez pour un autre département.



Fig_sup 1 : Évolution du taux brut d'exclusion médicale (A) et du taux brut de participation (B) dans les départements ayant eu accès aux données d'acte de la CPAM entre 2010 et 2022. La modalité « Global » est la moyenne des 60 départements ; La modalité « Pas CPAM» est la moyenne des 45 départements n'ayant jamais eu accès aux données d'acte de la CPAM entre 2010 et 2022.



Fig_sup 2 : Évolution du taux brut d'exclusion médicale (A) et du taux brut de participation (B) dans les départements ayant eu accès aux données des registres de cancers entre 2010 et 2022. La modalité « Global » est la moyenne des 60 départements ; La modalité « Pas de RC» est la moyenne des 55 départements n'ayant jamais eu accès aux données de registres de cancers entre 2010 et 2022.



Fig_sup 3 : Evolution du taux brut d'exclusion médicale (A) et du taux brut de participation (B) dans les départements ayant eu accès aux données du Programme de médicalisation des systèmes d'information (PMSI) entre 2010 et 2022. La modalité « Global » est la moyenne des 60 départements ; La modalité « Pas de CH» est la moyenne des 49 départements n'ayant jamais eu accès aux données du PMSI entre 2010 et 2022.

4.2 Filtrage de la population cible du dépistage du cancer colorectal en France, État de lieux en prélude de l'usage de la base médico-administrative de l'assurance maladie: Étude 4.

4.2.1 – CONTEXTE

A ce jour en France, bien que certains SGDO reçoivent des données d'exclusions des CPAM, il n'existe aucune interconnexion entre les bases de données du dépistage et le SNDS tel que possible dans plusieurs programmes comme décrit dans l'axe 1 de cette thèse. En vue d'argumenter toute démarche d'interconnexion, il nous a semblé opportun de faire un état des lieux des stratégies et moyens actuels d'exclusion médicale et leurs potentiels impacts sur l'estimation du taux d'exclusion médicale. L'imprécision des algorithmes d'extraction étant BDMA a été mise en évidence dans l'axe 1 de cette thèse. Une analyse a priori, du niveau d'amélioration du taux d'exclusion médicale que permettrait d'atteindre un filtrage en amont de la population cible, à l'aide des données du SNDS, est nécessaire.

4.2.2 – METHODES

Dans le programme français, il existe actuellement 11 critères majeurs d'exclusion et trois stratégies majeures d'exclusion médicale (***Figure-4.2.1***) : a) la stratégie d'exclusion en amont des invitations que réalisent les structures de gestion du dépistage par l'usage disparate des bases de données (Base de données du dépistage, Registre de cancer, Base de données hospitalières…) ou par l'intermédiaire des caisses primaires d'assurance maladie (CPAM : données de coloscopie). b) La réponse postale de la personne ou de son médecin, suite à une invitation de la structure de gestion; c)L'auto-exclusion de la personne sur le site internet de commande de kit de test ou lors d'un retrait de kit de test à la pharmacie. Toute stratégie d'amélioration du taux d'exclusion médicale par l'usage du SNDS serait cohérente avec la stratégie d'exclusion en amont.

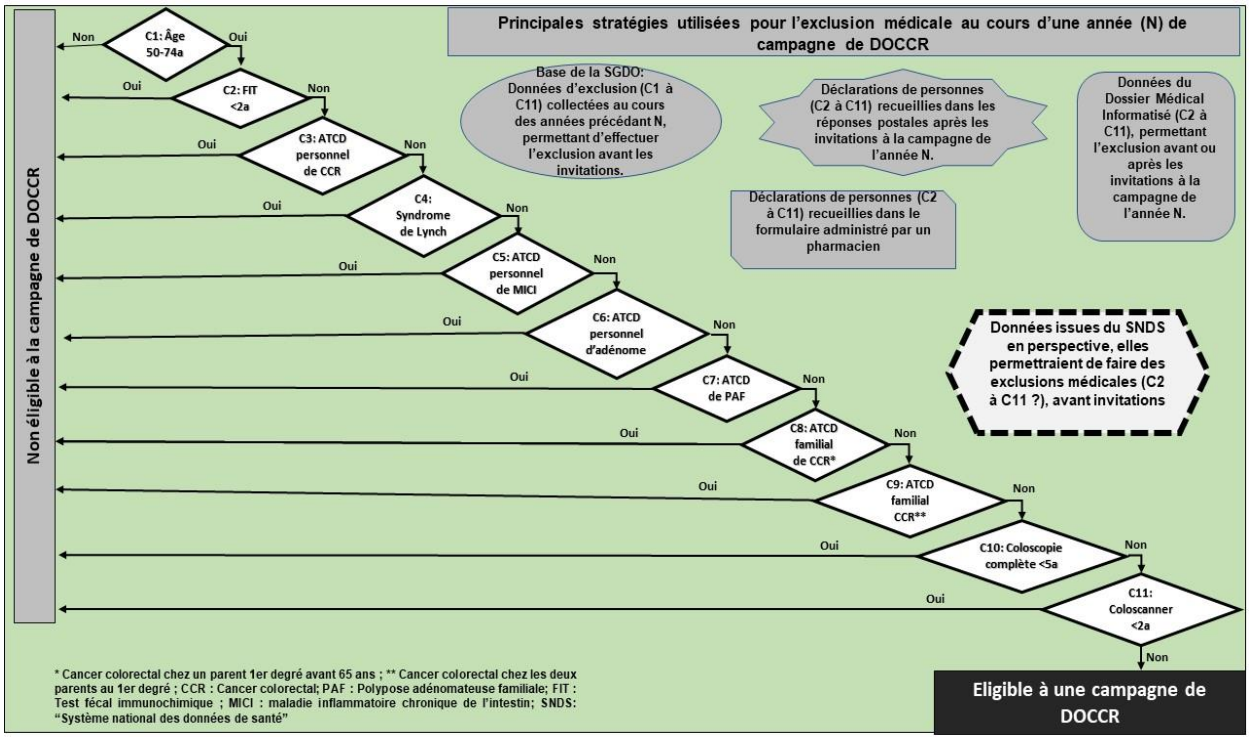


Figure-4.2.1 : Principales stratégies utilisées pour l'exclusion médicale au cours d'une année (N) de campagne de DOCCR.

Les moyens d'exclusion médicale inclus dans cette analyse étaient : i) Les bases de données départementale ou régionale du dépistage organisées : Elles contiennent des informations sur les tumeurs (Polypes/Adénomes, CCR) dépistés et antécédents morbides justifiant d'une exclusion(collectées lors des campagnes antérieures) depuis 2009 (année de lancement officiel du programme) ou avant (notamment dans les départements pilotes). Elles permettent une exclusion exclusivement en amont des invitations. ii) Le Dossier Patient (DP) du médecin traitant et le Dossier Patient Informatisé(DPI) des établissements de soin ont vocation à stocker l'ensemble des documents liés au parcours de soins du patient. Ils permettent une exclusion lors d'une consultation spontanée ou faisant suite à une invitation à participer à une campagne de dépistage. iii) Le SNDS est un moyen en perspective qui pourrait être utilisé seulement s'il est chaîné avec les bases de données du dépistage.

Analyse d'impact

L'impact des stratégies et moyens a été analysé en termes de proportion de personnes potentiellement éligibles au dépistage pouvant être incluses à raison (VP : vrai positif) dans les campagnes de dépistage ou exclues à tort (FN : faux négatif), en termes de proportion de personnes potentiellement inéligibles au dépistage et incluses à tort (FP : faux positif) ou exclues à raison(VN : vrai négatif). En rappel, le courrier d'invitation actuel invite les

personnes à retirer un kit de test chez le Médecin traitant ou chez le pharmacien (depuis juin 2022) ou d'en faire une commande en ligne sur un site dédié (depuis mars 2022).

Les FP, FN, VN et VP ont été simulés dans quatre scénarii d'inéligibilité au dépistage:

Scenario S1 : Le retrait du kit de test se fait chez un médecin (en cabinet, à l'hôpital ou dans la structure de Gestion du dépistage) à la suite d'une consultation spontanée ou après réception d'un courrier d'invitation de la structure de gestion du programme de dépistage . L'éligibilité est complètement vérifiée à l'aide du DP/DPI ou à l'aide de la base de données du dépistage, une personne éligible l'est donc réellement (VP).

Scenario S2 : Le retrait du kit de test se fait chez un pharmacien à la suite d'un passage spontané ou après réception d'un courrier d'invitation de la structure de gestion du programme de dépistage. Il existe un risque non nul que l'éligibilité soit partiellement établie par manque d'accès au DP/DPI ou à la base de données du dépistage, la personne jugée éligible peut ne pas l'être réellement (existence de FP).

Scenario S3 : Le kit de test est reçu à domicile suite à un envoi postal par la structure de gestion ou suite à une demande réalisée sur le site internet dédié, l'éligibilité est établi par questionnaire auto-administré avec un risque non nul qu'elle soit fausse selon la sincérité ou la mémoire du sujet qui remplit le questionnaire.

Scenario S4 : Le filtrage en amont de la population cible à l'aide des données du SNDS, scenario en perspective.

Caring is Sharing – Exploiting the Value in Data for Health and Innovation
M. Hägglund et al. (Eds.)
© 2023 European Federation for Medical Informatics (EFMI) and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SHTI230157

Refinement of the Target Population for Colorectal Cancer Screening in France, Inventory as a Prelude to the Use of Medico-Administrative Database

Akoi KOÏVOGUI[a,1], Robert BENAMOUZIG[b] and Catherine DUCLOS[a]
[a]*Université Sorbonne Paris Nord, Laboratoire d'Informatique Médicale et d'Ingénierie des connaissances en e-Santé, LIMICS, Sorbonne Université, UMLS 1142-INSERM, 93000, Bobigny, France*
[b]*AP-HP, Avicenne, Service d'Hépato Gastro Entérologie, 93000, Bobigny, France*

Abstract. The study describes the level of improvement in the risk of misclassification that would be achieved by refining the campaign target population using a query in the French medico-administrative database (SNDS). The SNDS's use requires other new strategies that can minimize the number of people wrongly included in the campaigns, because its accuracy is less than 100%.

Keywords. Medico-administrative database, Colorectal cancer screening

1. Introduction

The French colorectal cancer (*CRC*) screening program (*CRCSP*) targets people aged 50 to 74. The medical exclusion (*MEx*) rate and the participation rate in 2020-2021 (respectively: 13.2% 35.7%; *www.santepubliquefrance.fr*) are estimated in an eligible population (*EliP*) and underestimated. To obtain *EliP*, a few people (i.e., people at risk of *CRC*) are excluded from target population (*TaP*). Unlike some European programs [1], in France, the *MEx* is assessed after invitations (Response from patients/GPs/physicians), generating a detrimental bias. With a view to interconnecting the health insurance medico-administrative database (*SNDS*) and the *CRCSP* database (*CRC-DB*), we aim to describe the risk of misclassification (Eligible/Not eligible) induced by *MEx* strategies and the level of risk's improvement that, would be achieved by refining the *TaP* upstream of invitations using *SNDS*.

2. Methodology

The CRCSP's major *MEx* strategies are:

[1] Corresponding Auhtor: Dr KOÏVOGUI Akoï, CRCDC-IDF, Seine-Saint-Denis site, 41 avenue de Verdun, 93146 Bondy, France. E-mail: akoi.koivogui@inserm.fr Tel.: +33155 89 10 10 Fax+33148 02 06 80

i) *MEx* made exclusively upstream of the invitations by *CRCSP*'s management structures (*CRCSP-Manager*) through the disparate use of: *CRC-DB*, Cancer registries, Hospital discharge databases and Primary health insurance databases. ***ii)*** *MEx* made following correspondence between *CRCSP-Manager* and GPs/Physician through the *Computerized Medical Record* (*CMR*) of the physician/healthcare facility. ***iii)*** *MEx* carried out by pharmacist by answering the questions of eligibility when collecting a test kit at the pharmacy. ***iv)*** The self-exclusion carried out by people either by making a postal response, or by completing an online eligibility form. ***v)*** *MEx* made through the *SNDS* (In perspective). Each strategy's risk of misclassification was analyzed in terms of the proportion of people potentially eligible for screening and included appropriately (*TP*: True positive) or wrongly excluded (*FN*: False negative) and, in terms of the proportion of people potentially ineligible for screening and wrongly included (FP: False positive) or appropriately excluded (*TN*: True negative). These *FP*, *FN*, *TN*, and *TP* were simulated in five scenarios, based on the five *MEx* strategies.

3. Results

To obtain a test kit, people must: ***Scenario-1 (Doctor consultation)***: eligibility by strategies ***(i)*** and ***(ii)*** with zero risk of misclassification (*TP*=100%, *FP*=0%, *FN*=0%, *TN*=100%). ***Scenario-2(Kit collect in pharmacy)***: There is a non-zero risk that eligibility (Strategy ***iii)*** may be partially proved (*TP*<100%, *FP*>0%, *FN*=0%, *TN*=100%) due to pharmacist 's lack of access to the *CMR* or *CRC-DB*. ***Scenario-3(Kit request, no consultation)***: The kit is received at home following a website request, eligibility is established (Strategy ***iv)*** with a non-zero risk (*TP*<100%, *FP*>0%, *FN*>0%, *TN*<100%). ***Scenario-4 (SNDS perspective)*** risks are like those of ***Scenario-3,*** because SNDS accuracy <100%. The risk that few people will be wrongly included or not included, in the participation rate's estimation is >0 in scenarios 2, 3 and 4 and =0 in 1.

4. Discussion and Conclusions

The *SNDS*'s sensitivity (≤87.7%) and positive predictive value(≤58.9%) are less than 100 [2]. This inaccuracy feed the reflection with an essential question: To what extent refining upstream invitations, by a *SNDS* query will improve the exclusion rate?

To making *MEx* using the SNDS, other new strategies are needed to minimize the number of people not wrongly included in the campaigns. The definition of these new strategies will have to go through a *SNDS's* evaluation using *CRC-DB* as a benchmark.

References

[1] Bulliard JL, Garcia M, Blom J, et al. Sorting out measures and definitions of screening participation to improve comparability: the example of colorectal cancer. Eur J Cancer. 2014;50(2):434-46.

[2] Quantin C, Benzenine E, Hagi M, et al. [Evaluation of medical information systems as a mean of identification of incident cases of colorectal cancer]. Sante publique 2014;26(1):55-63.

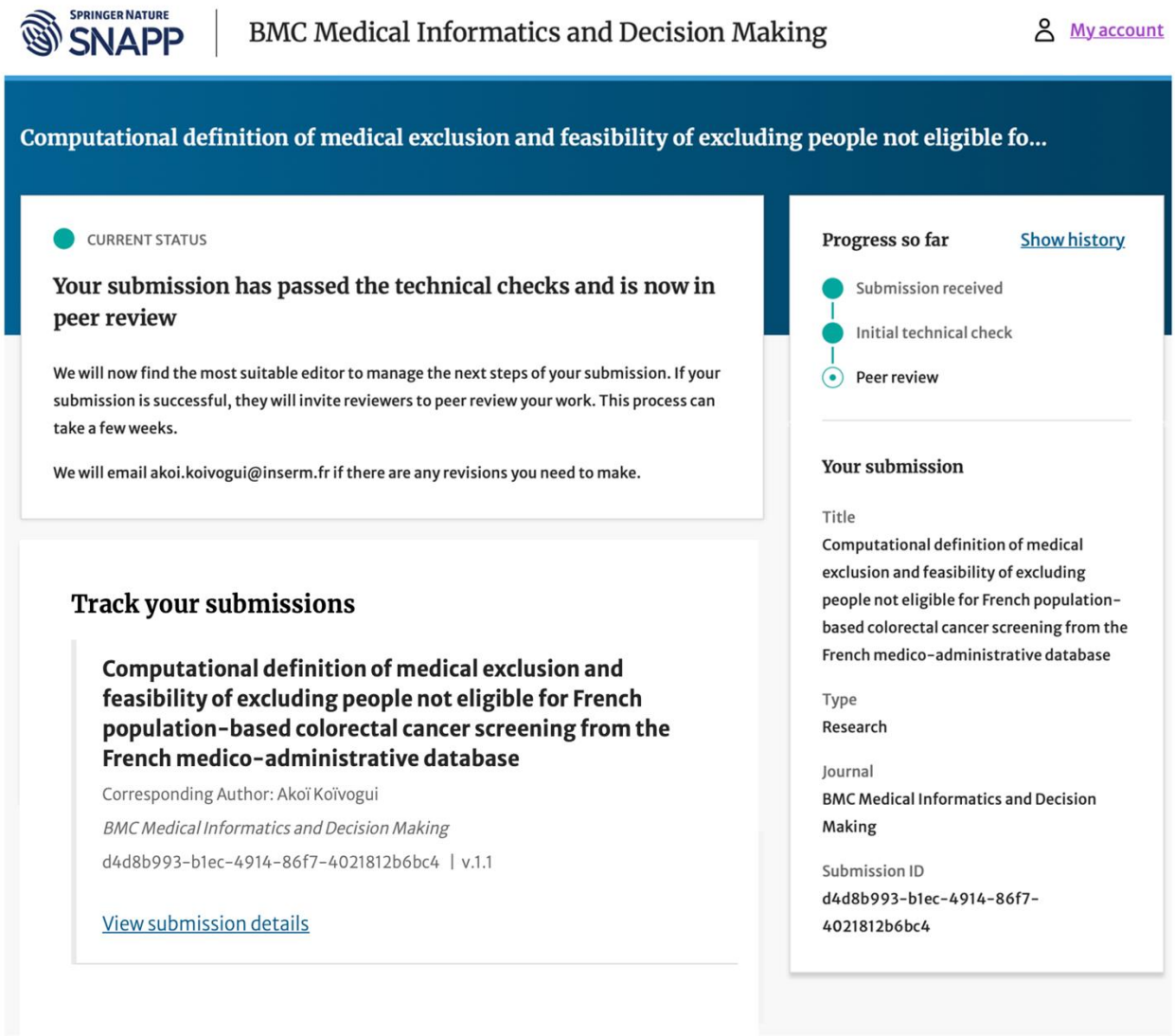
4.3 Algorithme d'exclusion médicale et faisabilité de l'exclusion des personnes non éligibles au dépistage à partir du SNDS : Étude 5.

4.3.1 – CONTEXTE

L'état des lieux a montré qu'aucune structure de gestion du dépistage n'utilise actuellement le SNDS. Cependant, l'optimisation du taux d'exclusion par un filtrage de la population cible à partir du SNDS ne peut être possible que si les situations morbides justifiant d'une exclusion médicale du DOCCR sont requêtables dans les bases du SNDS. En amont de toute exploitation du SNDS, il semblait évident d'investiguer leur présence dans les bases du SNDS. Une compilation d'algorithmes publiés faciliterait la définition opérationnelle de la personne à exclure du DOCCR en amont des invitations à partir du SNDS.

4.3.2 – METHODES

A l'aide d'une revue non systématique de la littérature, l'ensemble des codes définissant une situation morbide pouvant justifier d'une exclusion des campagnes de dépistages ont été compilés. La revue de la littérature a concerné la liste exhaustive de travaux réalisées sur les bases du SNDS et répertoriés sur le site de la CNAM (*https://assurance-maladie.ameli.fr/etudes-et-donnees/etudes-publications/references-publications-articles-assurance-maladie*). Il s'agit de travaux réalisés par l'Assurance Maladie ou d'autres équipes institutionnelles en partenariat avec la Cnam, publiés dans des revues à comité de lecture. Dans cette liste, ont été sélectionnés, les articles originaux publiés après 2010, quel que soit l'objectif principal et la langue de publication, traitant d'au moins une des situations morbides (Cancer colorectal, adénome ou polype colorectal, maladie inflammatoire chronique de l'intestin, polypose adénomateuse familiale ou syndrome de Lynch), des actes de diagnostic (coloscopie, coloscanner) ou de soins (polypectomie,….) pouvant justifier d'une exclusion médicale temporaire ou définitive du DOCCR.


SPRINGER NATURE
SNAPP
BMC Medical Informatics and Decision Making
My account
Computational definition of medical exclusion and feasibility of excluding people not eligible fo...
CURRENT STATUS
Your submission has passed the technical checks and is now in peer review
We will now find the most suitable editor to manage the next steps of your submission. If your submission is successful, they will invite reviewers to peer review your work. This process can take a few weeks.
We will email akoi.koivogui@inserm.fr if there are any revisions you need to make.
Progress so far
Show history
Submission received
Initial technical check
Peer review
Your submission
Title
Computational definition of medical exclusion and feasibility of excluding people not eligible for French population-based colorectal cancer screening from the French medico-administrative database
Type
Research
Journal
BMC Medical Informatics and Decision Making
Submission ID
d4d8b993-b1ec-4914-86f7-4021812b6bc4
Track your submissions
Computational definition of medical exclusion and feasibility of excluding people not eligible for French population-based colorectal cancer screening from the French medico-administrative database
Corresponding Author: Akoï Koïvogui
BMC Medical Informatics and Decision Making
d4d8b993-b1ec-4914-86f7-4021812b6bc4 | v.1.1
View submission details

Authors: Koïvogui Akoï[1,2], Balamou Christian[3], Benamouzig Robert[4], Duclos Catherine[1]

[1]Université Sorbonne Paris Nord, Laboratory in Medical Informatics and Knowledge Engineering in e-Health, LIMICS, Sorbonne Université, INSERM, F-93000, Bobigny, France

[2]CRCDC-IDF, Seine-Saint-Denis site, 93146 Bondy, France.

[3]CRCDC-AURA, Site de l'Ain Bourg-en-Bresse, 01000 Bourg-en-Bresse, France

[4]Hôpital Avicenne (AP-HP), Service d'Hépato Gastro Entérologie, 93000, Bobigny, France.

Correspondence to: Dr KOÏVOGUI Akoï, Université Sorbonne Paris Nord, Laboratory in Medical Informatics and Knowledge Engineering in e-Health, LIMICS, Sorbonne Université, INSERM, F-93000, Bobigny, France, mail: akoi.koivogui@inserm.fr

Title: Computational definition of medical exclusion and feasibility of excluding people not eligible for French population-based colorectal cancer screening from the French medico-administrative database

Abstract.

Background: In the French population-based colorectal cancer screening program (CRCSP), the fact that the medical-exclusion rate was estimated only after a collection of voluntary statements from subjects could compromise an exhaustive collection of potential cases of medical-exclusion. The health insurance medico-administrative database (SNDS) that contains medical and healthcare consumption information have to date never been used to refine the target population of the CRCSP.

Objective: To identify in the SNDS, from published and disparate algorithms, the computational definitions of morbid situations that could justify medical exclusion from the CRCSP.

Methods: The non-systematic review of the literature synthetised an exhaustive list of algorithms targeting in SNDS, the morbid situations (CCR, colorectal adenoma/polyp, chronic inflammatory bowel disease, familial adenomatous polyposis, or Lynch syndrome colonoscopy, coloscanner, polypectomy) which may justify temporary or permanent medical exclusion from the CRCSP campaigns. Secondly, the discovered codes of morbid situations were searched on statistical reports to estimate their frequencies of use in SNDS (in 2021), and their interest in the computational phenotypes' algorithm.

Results: The analysis of the literature (28 articles/studies) highlights the existence of diagnostic or therapeutic codes that can define in the SNDS database, the morbid situations justifying medical exclusion from the CRCSP. Except for personal or family history of CRC classifiable in the Z85.0 or Z80.0 codes of the ICD-10, almost all the morbid situations have a requestable definition in the SNDS. The target favoured by the search algorithms was the ICD-10 code (i.e., C18-C20, K50, K51). The definition codes listed were frequently used in SNDS in 2021, except for a few codes (D12.6+6, M07.5). From this definition of morbid situations by the only codes of the ICD-10 or the procedure codes emerges a feasibility and a decision-making algorithm for the choice of the person to be excluded from CRCSP campaign, using the SNDS. Age is the first discriminating variable in this decision-making algorithm because the CRCSP targeted people aged 50 to 74 years old and a restriction on age was made in several included SNDS's studies. The second discrimination based on diagnostic evidence derives its relevance from the quasi-systematic search for ICD-10 diagnostic codes in SNDS's studies.

Conclusion: In addition to being widely used in the context of medico-economic and epidemiological studies, the SNDS currently contains almost all the data essential for estimating the rate of medical-exclusion during colorectal cancer screening campaigns. While initiating the answer to the question of the choice of the most appropriate algorithm in each context, this review of the literature also emphasizes the need for validation studies because the quality of the algorithms used conditions the quality of the studies carried out in the medico-administrative databases.

Keywords. Medico-administrative database, Colorectal disease phenotype, Colorectal cancer screening, Participation rate, Medical exclusion rate

Introduction

All the conditions justifying the implementation of a screening policy are met by colorectal cancer (CRC) [1]. The French population-based colorectal cancer screening program (CRCSP) is organised in accordance with specifications[2]. Since 2009, the campaign participation rate remains well below the minimum European objective of 45% [3]**.** Several reasons have already been mentioned to explain this low participation [4, 5] but the bias related to the method of estimating the participation rate remains little discussed. The French public health agency ("*Santé Publique France*") defined the campaign participation rate by the ratio between the number of people screened and the eligible population, i.e. the population targeted by the program (men and women aged 50 to 74) from which are subtracted people excluded for medical reasons [6].

The definition of this eligible population was based on the CRC average risk, characterized by age (50-74 years) and the absence of personal or family CRC risk[2]. People at high risk (personal or family history of colorectal adenomas, cancers, or inflammatory bowel diseases -IBD-: ulcerative colitis and Crohn's disease) or very high risk (familial adenomatous polyposis or Lynch syndrome) of developing CRC and people with severe extra-intestinal pathologies must be excluded from this eligible population. Similarly, people who have undergone a colonoscopy within 5 years or a CT colonography within 2 years must be temporarily excluded if the result of this colonoscopy/colonography reports no colorectal neoplasia [3]. The fact that the exclusion rate was estimated only after a collection of voluntary statements from subjects (questionnaire attached to the invitation letter) and GPs or treating physicians (structured questionnaire) compromises an exhaustive collection of potential cases of medical exclusion. In addition to an interdepartmental disparity, the exclusion rate seems to be underestimated. In ten years, the exclusion rate has increased little, from 10.5% (2008-2009 First official CRCSP campaign) [3] to only 14.0% (2018-2019 campaign) [6] while in France, the annual number of colonoscopies exceeds 1 million. As a reminder, the number of colonoscopies performed in France was estimated at 1,200,529 in 2011[7] and 1,402,047 in 2016 (including 76.9% among people aged over 50)[8], this represents a 16.8% increase in five years.

In France, the health insurance medico-administrative database (usually called: "S*ystème National des Données de Santé*": SNDS) that contains medical and healthcare consumption information have to date never been used to refine the target population of the CRCSP. This study aimed to identify in the SNDS, from published and disparate algorithms, the computational definitions of morbid situations that could justify medical exclusion from the CRCSP.

Methods

Type of study

Using a non-systematic review of the literature, the study first listed the published computational phenotypes or algorithms. The study synthetised all algorithms searching in the SNDS for morbid situations that could justify medical exclusion from the CRCSP. Secondly, the discovered coded morbidities were searched on statistical reports to estimate their frequencies of use in SNDS, in 2021, and their interest in the computational phenotypes' algorithm.

Study context

Architecture of the SNDS database

The SNDS contains data from the health insurance Consumption Datamart ("*Datamart de Consommation Inter-Régimes*") [9]. Since 2016, the SNDS database has been fed with the national insurance information system database ("*Le système national d'information inter-régimes de l'assurance maladie*"), which contained the information on hospital care and city medicine, linked to the hospital information systems database ("*Programmme de Médicalisation des Systèmes d'Information*": PMSI) and the database of medical causes of death. These databases collecting individualized and anonymous data on all reimbursements of health expenditure from the public and private sectors for subjects affiliated to one of the compulsory health insurance schemes, i.e., nearly 99% (67 million people) of people residing in France. An unique anonymized identifier makes it possible to follow a person in the healthcare system. Diseases are identified according to the tenth revision of the WHO International Classification of Diseases (ICD-10) [9]. Diseases are recorded as diagnosis in hospital stays, as diagnosis for long term illness. Medical acts are coded according to the local classification of medical procedure ("*classification commune des actes médicaux*": CCAM). Medical acts are recorded as ambulatory acts or appears in hospital stays description. Records contain also dispensed drugs which are coded according to a national nomenclature, the anatomical therapeutic chemical (ATC) classification system and biological acts which are coded according to a national nomenclature. Exceptionally, in the reports of hospital stays (diagnoses covered, medical procedures, drugs administered during the hospital stay, homogeneous group of hospital stays, homogeneous group of patients -HGP-, etc.) contained in the hospital information systems data, the diagnoses are coded in main diagnosis which can be associated with a diagnosis related and associated diagnoses [9]. Aggregated data from SNDS are available through Open Data platforms [10, 11].

Organization of medical exclusions during CRCSP campaigns

The CRCSP management structures uses several means to quantify the number of people to be excluded and the duration (permanent or temporary) of medical exclusion. As part of the CRCSP, the lists of target populations (defined

104 on the age criterion) were made available to the CRCSP management structures by each Compulsory Health Insurance
105 Scheme. People were invited on their 50th birthday (primo-invitation) and every two years after a screening test
106 (Invitation-subsequent) until the age of 74. Anyone who completed a screening test was invited subsequently two years
107 after a negative result. In the event of a positive result, the person was invited subsequently five years after a negative
108 colonoscopy and excluded from the CRCSP after a positive colonoscopy (adenoma, CRC). For those who have not
109 completed a test but who have (or their representative) sent a postal response to the structure, a permanent exclusion
110 from the CRCSP was decided if a postal response reports a diagnosis of a pathology or morbid situation justifying a
111 medical exclusion (CRC, Crohn's disease, Ulcerative colitis, adenoma) or a family history of CRC (CRC in only one of
112 the two 1st degree parents before the age of 65, CRC in both 1st degree parents, Familial Adenomatous Polyposis). A
113 temporary exclusion was decided in the event of a colonoscopy less than five years old or a CT colonography less than
114 two years old or a search for blood in the stools less than two years old.

115 ***Literature search strategy and study selection criteria***

116 The literature review concerned the exhaustive list of studies carried out using the SNDS database and listed on the
117 health insurance ("*Caisse Nationale d'Assurance maladie*": CNAM) platform [12]. These studies were carried out by
118 the CNAM or other institutional teams in partnership with the CNAM, published in peer-reviewed journals. From this
119 list, were selected and read in full, the original articles published after 2010, whatever the main objective and the
120 language of publication, targeting with at least one of the morbid situations (Colorectal cancer, colorectal adenoma, IBD,
121 familial adenomatous polyposis-FAP-, Hereditary Non Polyposis Colorectal Cancer-HNPCC- or Lynch syndrome), or
122 diagnostic procedures (colonoscopy, CT colonography), or treatment (polypectomy) which may justify temporary or
123 permanent medical exclusion from CRCSP.
124 At the end of the documentary research, we excluded all articles published before 2010. Regardless of the language of
125 publication, publications whose title contained at least one of the keywords (defined as morbid situations) were selected
126 to read the summary and methodology. Whatever the search target of the algorithm (ICD-10 code, Code of diagnostic
127 or treatment procedure, anatomopathological or biological examination code, drug code, computer language), the study
128 was selected only if the algorithm was at least described in the methodology or at best accessible in the study or in a
129 previous study or on a dedicated digital directory.

Data collected in articles

The morbid situation (CRC, IBD, Colonoscopy, CT Colonography, family history) targeted by the search algorithm in the SNDS, their definitions (Type of codes and Terminology), their Type of assessment and indicators assessed were collected.

Analysis of the feasibility of CRCSP's medical exclusion from the SNDS

Feasibility has been described in terms of:

Possibility of extracting data from the SNDS: "*Scansante"* is a national platform for restoring data from French healthcare establishments. Through the statistical application[10] developed, the aggregated data from the SNDS are returned by year for the whole of France and by type of code (care act or diagnostic). Through this application, we searched for the existence in the SNDS of all the codes targeted by each of the algorithms found at the end of the literature review. For each code listed, the frequency of hospital stays using the codes was collected for the year 2021 and for the whole of France. After extracting the aggregated data from "*Scansante"*, a description of these frequencies was made.

Performance of code search algorithms: the algorithm evaluation indicators have been described. For each morbid situation justifying medical exclusion, the operational formulation was made by the synthesis of its different definitions in the studies listed by the literature review.

Results

Were included in the non-systematic review, 28 articles/studies [13-39] among the 339 listed by the CNAM. The targeted objectives (***Table-1***) were mainly to validate a diagnostic code search algorithm [19, 34, 35]; to estimate the incidence or prevalence of chronic colorectal diseases [13, 22]; to evaluate the prescription or the effectiveness of a treatment for chronic colorectal diseases [18, 28, 29] or to search for colonoscopy procedures [14, 23, 24]. Although a validation process was not carried out, two referenced algorithms, the INCA algorithm [40] and the health insurance mapping algorithm [41] were used respectively by three [16, 20, 21] and five[26, 36-39] cohort studies. Six studies analysed the performance of the search algorithms: either by a ratio of the standardized incidence rate between the rates observed in the SNDS and those expected (estimated on the basis of the French network of cancer registries)[13, 17, 22]; either by estimating a percentage of concordance (kappa coefficient) between the cases extracted from the PMSI

reference to the data cancer registries [34, 35].

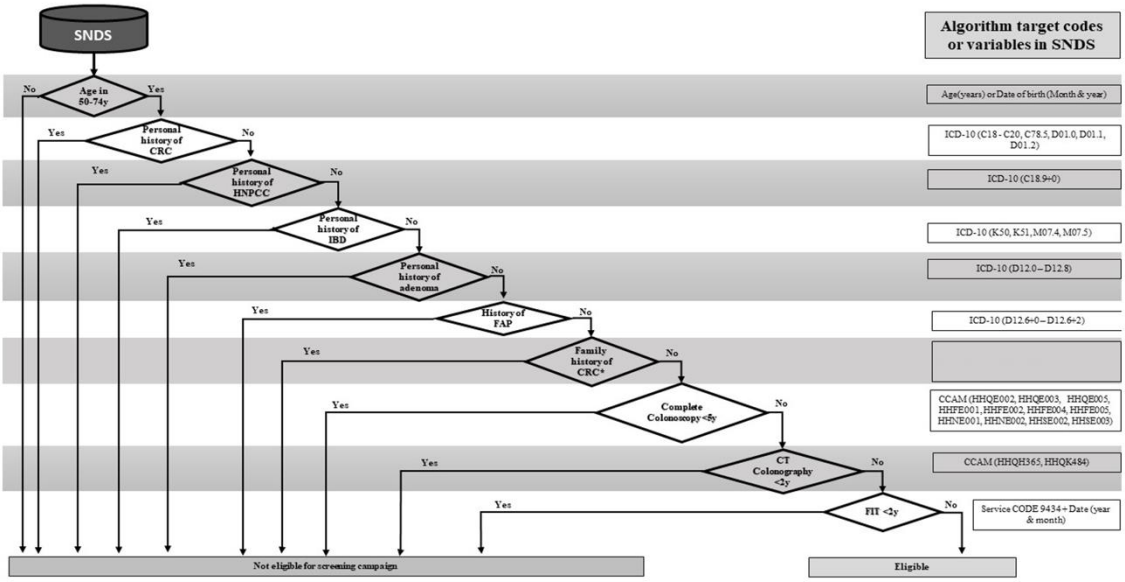
The analysis of the literature highlights the existence of diagnostic or therapeutic codes that can define in the SNDS database, the morbid situations justifying medical exclusion from the CRCSP. Except for personal or family history of CRC classifiable in the Z85.0 or Z80.0 codes of the ICD-10, almost all the morbid situations justifying medical exclusion from the CRCSP have a requestable definition in the SNDS (***Table-2***). The target favoured by the search algorithms was the ICD-10 code (i.e., C18, K51, M07.5) for any morbid situation evoking a pathological situation. The colonoscopies/CT colonoscopy for diagnostic purposes (i.e., HHQE002: "Total colonoscopy, with crossing of the ileocolic orifice") or therapeutic (i.e., HHSE003; "Haemostasis of colon lesion with laser, by total colonoscopy") were defined by CCAM codes.

The definition codes listed by the literature review were frequently used in SNDS database in 2021, except for a few codes (D12.6+6, M07.5). The ICD-10 code C18.x (colon cancer) was recorded as principal diagnosis in 48,595 hospital stays. The CCAM code HHQE005 (Total colonoscopy with visualization of the caecal lower fundus, without crossing the ileocolic orifice) had been used in 633,002 hospital stays in France in 2021 (***Table-2***).

From this definition of morbid situations by the only codes of the ICD-10 or the CCAM emerges a feasibility and a decision-making algorithm for the choice of the person to be excluded from screening campaign, using the SNDS. Age is the first discriminating variable in the decision-making algorithm for CRCSP's medical exclusion using the SNDS because the CRCSP targeted people aged 50 to 74 years old. (***Figure-1***). This first discrimination is feasible because a restriction on age was made in several included studies [17, 24, 37, 38]. As an illustration, in their study on healthcare reimbursements, Tuppin et al. [37] extracted from the SNDS the age group of the 25.7 million individuals who benefited from gastrointestinal treatment. The second group of discriminating variables must be based on diagnostic evidence: Whatever the year of diagnosis or treatment, any individual aged 50-74, identified in the SNDS with a diagnostic code (***Figure-1***) related to a chronic colorectal disease, falls within the criteria for definitive medical exclusion of CRCSP campaigns. The second discrimination based on diagnostic evidence derives its relevance from the quasi-systematic search for ICD-10 diagnostic codes in the studies included in this analysis [14, 16, 17, 19, 22, 25, 40]. The third group of discriminating variables must consist of colonoscopy procedures (Surveillance + Diagnosis + Therapeutics).

No code was explicit in the SNDS related to family history of colorectal cancer. Indeed, surveillance colonoscopies carried out in cases with a personal (Z85.0) or family history of CRC (Z80.0) are probably detectable in HGPs with a main diagnosis Z85.0 or Z80.0. As an illustration, in 2021, among a total of 90,541 hospital stays with main diagnosis coded Z80.0, a digestive endoscopy was performed in 81,583 (HGP code 06K04J). But there is no guarantee that it was a surveillance for CRC because the code "Z" was not the prerogative of any digestive cancerous pathology (***Table-3***).

190



191 CCAM: "Classification commune des actes médicaux"; CRC: Colorectal cancer; FAP: Familial adenomatous polyposis; FIT: Fecal immunochemical test; HNPCC: Hereditary Non Polyposis CRC; IBD: Inflammatory Bowel Disease

192 **Figure 1: Decision-making algorithm for the choice of the person to be excluded from screening campaigns based on the SNDDS databases.**

193 **Table 1:** ***Articles included in the non-systematic review.***

194

First Author et al. Journal and year of publication [Reference]	Transcription of all or part of the Objective consistent with this review	Terminology used (Codes of interest targeted by the algorithm and justifying inclusion in this study), Algorithm [Reference] if described elsewhere.	Number of algorithms assessed, Type of assessment (Indicators assessed)
Ajrouche A, et al. Pharmacoepidemiol Drug Saf 2017 [9]	To estimate cancer incidence (including **CRC**)	ICD-10 (C00-C97, Z85)	15, Comparative data from a Cancer Registry (SIR)
Blotiere PO, et al. Dig Liv Dis 2016 [10]	Search for drug prescription (treatment of **IBD**)	ICD-10 (K50, K51)	None
Blotiere PO, et al. Clin Res Hepatol Gastroenterol 2014 [11]	To estimate the perforation and haemorrhage rate after **COLONOSCOPY**	CCAM (HHFC001, HHFE001, HHFE002, HHFE004, HHFE005, HHNE001, HHNE002, HHQE002, HHQE004, HHQE005)	None
Bousquet PJ, et al. PLoS One 2018 [12]	To describe the French national cancer cohort (including **CRC patients**).	ICD-10 (C18-C20, D01.0, D01.1, D01.2, D01.4, D01.9 D12.0-D12.8, D13.9, D37.3, D37.4, D37.5, D37.7, D37.9, Z80.0, Z80.9, Z85.0, Z85.9), INCA Algorithm [36]	None
Doat S, et al. Rev Epidemiol Sante Publique 2016 [13]	To estimate **CRC** incidence	ICD-10 (C18, C19, C20)	5, Comparative data from a Cancer Registry (SIR)
Doat S, et al. Eur J Cancer 2014 [14]	To identify the **CRC**'s treatment modalities	ICD-10 (C18, C19, C20)	None
Gouverneur A, et al. Rev Epidemiol Sante Publique 2017 [15]	To identify cases of metastatic **CRC**	ICD-10 (C00-C75, C77.x, C78.X, C79.x)	Comparative data from a medical record (Coef. kappa)
Janah A, et al. BMC Health Serv Res 2020 [16]	To measure the prevalence of access to inpatient palliative care in **CRC** patients	ICD-10 (See list of codes above (INCA algorithm)) [36]	None
Janah A, et al. Clin Epidemiol 2019 [17]	To measure the prevalence of access to inpatient palliative care in **CRC** patients	ICD-10 (See list of codes above (INCA algorithm)) [36]	None
Kudjawu Y, et al. BEH 2013 [18]	To estimate **CRC** incidence	ICD-10 (C18-C21)	1, Comparative data from a Cancer Registry (IR*)
Laanani M, et al. Am J Gastroenterol 2020 [19]	To identify the systemic serious adverse event after **COLONOSCOPY**	CCAM (HHFC001, HHFE001, HHFE002, HHFE004, HHFE005, HHNE001, HHNE002, HHQE002, HHQE004, HHQE005)	None
Laanani M, et al. Clin Gastroenterol Hepatol 2019 [20]	To identify the systemic serious adverse event after **COLONOSCOPY**	CCAM (HHFC001, HHFE001, HHFE002, HHFE004, HHFE005, HHNE001, HHNE002, HHQE002, HHQE004, HHQE005)	None
Lamarche-Vadel A, et al. BMC Med Inform Decis Mak 2014 [21]	To identify the **CRC** cases in the causes of death	ICD-10 (C18, C20)	None
Melac AT, et al. Support Care Cancer 2020 [22]	To identify the **CRC** cases in the causes of death	ICD-10 (C18-C20, D01.0, D01.1, D01.2, K50, K51, M07.4, M07.5), CNAM Algorithm [37]	None
Mellot E, et al. Stud Health Technol Inform 2019 [23]	Definition of a **IBD** as chronic disease	ICD-10 (K501, K508, K509, K51, M0753, M0758)	None

Meyer A, et al. Aliment Pharmacol Ther 2020 [24]	To describe pregnancy in women with **IBD**	ICD-10 (K50, K51)	None
Meyer A, et al. Ann Inter Med 2019 [25]	Evaluate of a treatment in **Crohn's disease**	ICD-10 (K50)	None
Meyer A, et al. Aliment Pharmacol Ther 2019 [26]	Evaluate an **Ulcerative Colitis**'s treatment	ICD-10 (K51)	None
Mitton N, et al. J Cancer Epidemiol 2011 [27]	Estimate **CRC** incidence	ICD-10 (C18-C21	None
Poulalhon C, et al. Rev Epidemiol Sante Publique 2018 [28]	Identification of **CRC** cases in the causes of death	ICD-10 (C18-C20, D01.0, D01.1, D01.2)	None
Poulalhon, C et al. BMC Palliat Care 2018 [29]	Identification of **CRC** cases in the causes of death	ICD-10 (C18-C20, D01.0, D01.1, D01.2)	None
Quantin C, et al. Sante Publique 2014 [30]	Estimate **CRC** incidence	ICD-10 (C18, C19, C20)	2, Comparative data from a Cancer Registry and medical records (SE; PPV)
Quantin C, et al. J Cancer Epidemiol 2012 [31]	Estimate **CRC** incidence	ICD-10 (C18, C19, C20)	2, Comparative data from a Cancer Registry and medical records (SE; PPV)
Tanguy-Melac A, et al. Cancer Med 2019 [32]	Describe medical care and cause of death in **CRC** patients	ICD-10 (See list of codes above (CNAM Algorithm)) [37]	None
Tuppin P, et al. Therap Adv Gastroenterol 2019 [33]	Identification of gastrointestinal therapeutic drug use by chronic disease (**IBD, CRC**) patients	ICD-10 (See list of codes above (CNAM Algorithm)) [37]	None
Tuppin P, et al. Presse Med 2019 [34]	To describe medical car and cause of death in **CRC** patients	ICD-10 (See list of codes above (CNAM Algorithm)) [37]	None
Tuppin P, et al. Bull Cancer 2017 [35]	To estimate **CRC** incidence	ICD-10 (See list of codes above (CNAM Algorithm)) [37]	None

195 **CCAM**: "*classification commune des actes médicaux* "; **CNAM**: "*Caisse Nationale d'Assurance maladie*"; **CRC:** Colorectal cancer; **ICD-10**: Tenth revision of the WHO International
196 Classification of Diseases; **IBD**: inflammatory Bowel Disease; **IR:** Incidence report; **SE**: Sensibility; **SIR:** Standardized incidence ratio; **PPV:** Positive predictive value;

197 ***Table 2: Obvious definitions of morbid situations and frequency of code use in* hospital stays *(homogeneous groups of patients) in 2021.***
198

Morbid situations and corresponding definition codes	Code label	Terminology repository	References of published algorithms	Frequency(n) of use in reports of hospital stays in 2021
CRC				
C18.x	Malignant neoplasm of colon	ICD-10	9, 12-18, 21-23, 27-35	48595
C19.x	Malignant neoplasm of rectosigmoid junction	ICD-10	9, 12-18, 21-23, 27-35	4819
C20.x	Malignant neoplasm of rectum	ICD-10	9, 12-18, 21-23, 27-35	19048
C21.x	Malignant neoplasm of the anus and anal canal	ICD-10	9, 12-18, 21-23, 27-35	2728
C78.5	Secondary malignant neoplasm of the large intestine and rectum	ICD-10	15	422
D01.0	Carcinoma in situ (CIS), colon	ICD-10	22, 28, 29, 32-35	6185
D01.1	CIS, rectosigmoid junction	ICD-10	22, 28, 29, 32-35	298
D01.2	CIS, rectum	ICD-10	22, 28, 29, 32-35	1573
Lynch Syndrome				
C18.9+0	Hereditary Non-Polyposis CRC	ICD-10	9, 12-18, 21-23, 27-35	219
IBD				
K50.x	Crohn's disease	ICD-10	10, 22-26, 32-35	20293
M07.4	Arthropathy in Crohn's disease	ICD-10	10, 22-26, 32-35	40
K51.x	Ulcerative colitis	ICD-10	10, 22-26, 32-35	25,697
M07.5	Arthropathy in Ulcerative colitis	ICD-10	10, 22-26, 32-35	0
Adenoma				
D12.x	*Benign neoplasm of colon, rectum, anus and anal canal (except D12.9: Anus and anal canal)*	ICD-10	12, 16, 17	77579
D12.6+3	Scalloped (colic) polyposis	ICD-10	12, 16, 17	1627

D12.6+4	Juvenile polyposis	ICD-10	12, 16, 17	76
D12.6+5	Polyposis (colic) in Peutz-Jeghers syndrome	ICD-10	12, 16, 17	84
D12.6+6	Polyposis (colic) in Cowden's disease	ICD-10	12, 16, 17	0
D12.6+8	Colon, other and unspecified benign neoplasms	ICD-10	12, 16, 17	7429
FAP				
D12.6+0	Familial adenomatous polyposis (FAP)	ICD-10	12, 16, 17	769
D12.6+1	Polyposis (colonic) linked to the APC gene	ICD-10	12, 16, 17	153
D12.6+2	Polyposis (colic) linked to the MYH gene	ICD-10	12, 16, 17	29
Colonoscopy				
HHQE002	Total colonoscopy, with crossing of the ileocolic orifice	CCAM	11, 19, 20	294624
HHQE003	Complete exploration of the colon after right colectomy, by endoscopy	CCAM	11, 19, 20	4450
HHQE005	Total colonoscopy with visualization of the caecal lower fundus, without crossing the ileocolic orifice	CCAM	11, 19, 20	533002
HHFE001	Removal of 1 to 3 polyps less than	CCAM	11, 19, 20	6236
HHFE002	Removal of 1 to 3 polyps less than 1cm in diameter from the colon and/or rectum, by total colonoscopy	CCAM	11, 19, 20	402784
HHFE004	Removal of a polyp more than 1cm in diameter or 4 or more polyps from the colon and/or rectum, by total colonoscopy	CCAM	11, 19, 20	72300
HHFE005	Removal of a polyp more than 1cm in diameter or 4 or more polyps from the colon and/or rectum, by rectosigmoidoscopy or by partial colonoscopy	CCAM	11, 19, 20	2228
HHNE001	Colon and/or rectum lesion destruction session without laser, by total colonoscopy	CCAM	11, 19, 20	2082
HHNE002	Colon and/or rectum lesion destruction session with laser, by total colonoscopy	CCAM	11, 19, 20	956
HHSE002	Haemostasis of colon lesion without laser, by total colonoscopy	CCAM	11, 19, 20	22118
HHSE003	Haemostasis of colon lesion with laser, by total colonoscopy	CCAM	11, 19, 20	2413

CT Colonography				
HHQH365	CT scan of the colon with insufflation [virtual colonoscopy], and intravenous injection of contrast medium	CCAM	19, 20	0 (Recent code 2022)
HHQK484	CT scan of the colon with insufflation [virtual colonoscopy], without intravenous injection of contrast medium	CCAM	19, 20	0 (Recent code 2022)

***CCAM**: the local classification of medical procedures "Classification commune des actes médicaux"(CCAM) ;* ***CRC:** Colorectal Cancer; FAP: Familial adenomatous polyposis;* ***ICD-10**: tenth revision of the WHO International Classification of Diseases;* ***IBD:** chronic inflammatory bowel disease.*

Table-3: Likely definitions of family and personal history of colorectal cancer

Homogeneous groups of patients (HGP) Code in hospital stays (Description)	*Special examination for bowel neoplasm screening*		*ICD-10 codes of Family history of malignant neoplasm of digestive organs*		*ICD-10 codes of Personal history of malignant neoplasm of digestive organs*		CCAM codes of the most used types of total colonoscopy procedures during hospital stays		
	Z12.1	**Z12.9**	**Z80.0**	**Z80.9**	**Z85.0**	**Z85.9**	**HHQE002**	**HHQE005**	**HHFE002**
06K04J (Diagnostic digestive endoscopy and anaesthesia, outpatient care)	0	0	81 583	0	12 884	749	248877	450511	0
06K02Z (Therapeutic digestive endoscopies and anaesthesia, stays of less than 2 days)	0	0	5 975	0	1 614	83	5815	7583	379702
06M16Z (Explorations and monitoring for digestive tract disorders)	0	0	2 132	0	1 416	87	2604	5715	195
06K05J (Stays including a diagnostic digestive endoscopy without anesthesia, outpatient care)	0	0	688	0	492	33	1547	4458	0
06K06J (Digestive disorders without surgery, with anaesthesia, outpatient care)	0	0	64	0	77	0	0	0	0
06K03J (Stays including therapeutic digestive endoscopy without anaesthesia, outpatient care)	0	0	40	0	52	0	31	76	3082
06C14J (Procedures on the rectum and anus other than rectal resections, outpatient care)	0	0	22	0	0	0	195	325	217
06C19J (Haemorrhoidectomies, outpatient care)	0	0	14	0	0	0	134	313	178
17K07J (Myeloproliferative diseases and neoplasms with imprecise location without surgery, with anaesthesia, outpatient care)	0	0	6	85	0	0	405	688	225
06C151(Other interventions on the digestive tract apart from laparotomies, level 1)	0	0	0	0	11	0	15	16	0
23K03J (Reasons for seeking care without surgery, with anaesthesia, in outpatient care)	704	87	0	0	0	0	2245	4541	488
23M20T (Other symptoms and reasons for seeking care, very short length of hospital stays)	44	54	0	0	0	0	216	702	77
23M20Z (Other symptoms and reasons for seeking care)	36	12	0	0	0	0	327	824	221
Others HGP	1	0	17	15	22	14	32213	57250	18399
Total	**785**	**153**	**90 541**	**100**	**16568**	**966**	**294624**	**533002**	**402784**

***CCAM**: the local classification of medical procedures "Classification commune des actes médicaux"(CCAM) ; **CRC:** Colorectal Cancer; **HGP**:* Homogeneous groups of patients; ***ICD-10**: tenth revision of the WHO International Classification of Diseases.*

Discussion

This study showed that there was no unique algorithm that can identify in the French database, all morbid situations that can justify medical exclusion from CRCSP campaigns. By listing all the codes defining in the SNDS, each of the morbid situations that can justify medical exclusion from the CRCSP, this study documents the constituent elements of a unique algorithm for extracting CRCSP data from the SNDS. Despite the small proportion of studies having benefited from an algorithm validation process, the high frequency of use of definition codes in reports of hospital stays, argue the feasibility of the medical exclusion of people not eligible for CRC screening using the SNDS database. The absence of the use of certain codes which are less frequent elsewhere (Examples: M075 and D12.6+6), raises the question of the quality of the coding in the SNDS.

This study highlights the use of a single phenotype model (list of codes) among the 5 frequently used models [42]. The study also highlights a lack of validation of the chronic colorectal disease's phenotypes in the SNDS. The few studies that adopted a validation approach only targeted the codes defining the CRC but not the other morbid situations defining a medical exclusion from the CRCSP. Although comparison with a standard benchmark is the common means of validating a phenotype, Chapman et al. [42] described 6 other phenotype validation mechanisms (i.e., Comparison with registry data; Comparison with medical record data; Inter-registration concordance analysis; Risk factor analysis; prognosis; genetic association). In this review, the validation mechanism was mainly the comparison with data from a cancer registry. In addition, the high proportions of false positives reported in the study by Quantin et al. [34] plead for the validation of any search algorithm for potential cases of medical exclusion in the SNDS databases.

Despite the alert on the quality of data in the SNDS, the need expressed by the management structures of the CRCSP for a standard method of medical exclusion from the CRCSP is relevant. Indeed, there was an inter-departmental disparity in terms of medical exclusion rates mainly due to differences in means in collecting medical exclusions. Concerning the method of collecting medical exclusion data, France Public Health Agency admitted that the number of people excluded was provided by the management structures from various sources of usable data (responses from people, records of acts of colonoscopies reimbursed by the primary health insurance funds, records of diagnostic and therapeutic acts, standardized reports of colonoscopy results and anatomo-cytopathological examinations, etc.), but access to which varied

234 according to the CRCSP management structures[6]. This study highlights the feasibility of standard access
235 to this medical exclusion data using the SNDS.
236 While initiating the answer to the question of the choice of the most appropriate algorithm in each context,
237 this literature review also emphasizes the need for validation studies because the quality of the algorithms
238 used conditions the quality of the studies carried out in the medico-administrative databases. If it passes the
239 validation stage successfully, the decisional extraction algorithm that results from our analysis of the
240 literature could validly support this need for a standard method of medical exclusion in CRCSP. The
241 relevance of the constituent elements of this decision-making algorithm is eloquently expressed in the
242 literature included in this study. Contrary to existing algorithms, the desired model for the extraction of
243 medical exclusion data from the CRCSP should not introduce any temporal variable, to comply with the
244 specifications of the CRCSP. Indeed, unlike the mapping of health insurance expenditure, the CRCSP
245 specifications make no distinction between cured cancer (or under surveillance) and active cancer because
246 in either case, the person is considered high risk and permanently excluded from CRCSP campaigns.
247 Similarly, since the diagnostic codes are almost systematic in the summaries of hospital stays, we did not
248 find it relevant to search for the codes for colon surgery procedures because not all colorectal lesions have
249 medical exclusion value from the CRCSP. Also, chemotherapy, radiotherapy, treatment with monoclonal
250 antibodies, are not specific to chronic colon-rectum diseases. In Bronson's algorithm[43], the same CPT
251 (Current Procedural Terminology) codes indexed chemotherapy or radiotherapy in the management of three
252 distinct cancers (Colon, Lung, Breast). We deduced that in the decision algorithm, the cases of treatment
253 (chemotherapy or radiotherapy or antibodies) of a chronic colon-rectum disease cannot escape
254 discrimination based on the diagnostic evidence, there is no need to make the algorithm heavier with
255 combinations of treatments and codes of tracer acts as approached in the INCA algorithm[40].
256 The fact of defining the last group of discriminating variables by colonoscopy/CT colonography procedures
257 is justified by the recommendations of the screening specifications[2]. A person can be excluded
258 permanently (high risk and very high risk of CRC) or temporarily (for 5 years if normal colonoscopy, for 2
259 years if normal CT colonography, or a screening test less than 2 years old) from screening campaigns. This
260 algorithm prioritizes permanent exclusion by minimizing the risk of temporarily excluding people who are
261 truly at high risk for CRC.
262 In addition to being widely used in the context of medico-economic and epidemiological studies, the SNDS
263 currently contains almost all the data essential for estimating the rate of medical exclusion during screening
264 campaigns. To achieve this, it is necessary to link the SNDS and the screening databases. This link is

essential because none of the studies included in this analysis was explicit on the existence or not of a linking between the SNDS databases and an external database.

This study has three limitations:

1- The quantification of the use of the codes only concerned hospital stays and main diagnostic codes, which underestimates the CCAM acts carried out in an extra-hospital environment. It also underestimates hospital follow-ups with colorectal pathology only as an associated diagnosis or as a related diagnosis, especially during chemotherapy in a day hospital.

2- There is a difficulty in tracing in the SNDS, people with a personal or family history of CRC who are probably not recorded in the SNDS databases because the Zxx codes are too imprecise, and monitoring is not regular for the most part. Indeed, unlike some European countries, there is no follow-up program for people at high risk of CRC in France, so the follow-up decision is made or not by the attending physician. It is therefore not possible to distinguish in the SNDS, the follow-ups carried out following a CRC diagnosed in a direct ascendant from those carried out following a CRC diagnosed in another member of the family (cousin, brothers …). Similarly, since the SNDS is recent, people with a personal history of CRC prior to 2006 cannot be identified in the SNDS database which would increase the number of false negatives when using the search algorithm in the SNDS database.

3- Our decision-making algorithm compiles many codes whose extraction accuracy has not been validated by studies. Indeed, the fact that the codes are represented in the database and frequently used in the studies included in the review does not guarantee their reliability in capturing the disease situations of interest in the SNDS. This uncertainty is accentuated by the fact that the objectives of the studies included in this review are different from each other, and none of them focuses on medical exclusion from screening campaigns.

These limits will be controlled by a process of monitoring and evaluation of the search algorithm over time. Indeed, the algorithm's rehabilitation needs are justified over time by the evolution of terminologies and vocabularies (i.e., the passage from ICD-10 to ICD-11) have been planned. Similarly, phenotype assessment studies display proportions of false positives (FP) and false negatives (FN) high enough to bias the rates estimated from any data extracted by the search algorithm. Obviously, these proportions (FP and FN) decrease over time with the continuous improvement of the quality of the databases. The algorithm will therefore be regularly updated with a perspective of in-depth evaluation using a benchmark. This methodological perspective will inevitably involve linking the SNDS with a screening database.

Conclusion

In addition to being widely used in the context of medico-economic and epidemiological studies, the French-MADB currently contains almost all the data essential for estimating the rate of medical-exclusion during screening campaigns. By listing all the codes defining in the SNDS, each of the morbid situations that can justify medical exclusion from the CRCSP, this study argues the feasibility of the medical exclusion of people not eligible for CRC screening using the SNDS database despite the rarity of studies having benefited from a phenotype validation process. The study concludes that not all situations that could justify medical exclusion can currently be queried in the SNDS, only recent morbid situations. The personal history may also not be identifiable if the follow-up goes beyond 2006 (year of anteriority of the SNDS data).

Since the screening databases and the evaluation of screening campaigns are departmental, defining a phenotype of the target person to be excluded from the CRCSP based on SNDS data is an essential step. While initiating the answer to the question of the choice of the most appropriate algorithm in each context, this review of the literature also emphasizes the need for validation studies because the quality of the algorithms used conditions the quality of the studies carried out in the medico-administrative databases.

Acknowledgments

The authors are grateful to all contributors who participated in the final revisions.

Authors' contributions

Akoï KOÏVOGUI: *Conceptualization, Methodology, Data curation, Original draft preparation* ***Christian BALAMOU****: Data curation, Writing Original draft preparation* ***Robert BENAMOUZIG****: Visualization, Supervision, Validation* ***Catherine DUCLOS****: Conceptualization, Methodology, Visualization, Supervision, Validation.* All the authors revised the manuscript, read, and approved the final version of this manuscript.

Conflict of Interest:

No conflict of interest.

Funding Sources

The study did not receive any special funding.

Summary points

What was already known on the topic:

- In the French population-based colorectal cancer screening program, the medical exclusion rate was estimated only after a collection of voluntary statements from subjects and GPs, which compromises an exhaustive collection of potential cases of medical exclusion.
- The French medico-administrative database contains medical and healthcare information have to date never been used to refine the target population of the screening program.

What this study added to our knowledge:

- This review showed that there was no unique algorithm that can identify in the French database, all morbid situations that can justify medical exclusion from CRCSP campaigns.
- The use of a single phenotype model and the lack of validation of research phenotypes of the chronic colorectal diseases in the French database.
- The French medico-administrative database currently contains almost all the data essential for estimating the rate of medical exclusion during screening campaigns.

338 **References**

339 1. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised
340 controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet (London, England).
341 1996;348(9040):1472-7. doi: 10.1016/S0140-6736(96)03386-7. PubMed PMID: 8942775.
342 2. JORF. Arrêté du 19 mars 2018 modifiant l'arrêté du 29 septembre 2006 relatif aux programmes de
343 dépistage organisé des cancers et portant modification du cahier des charges du dépistage organisé du
344 cancer colorectal. JORF n° 0068 du 22 mars 2018 texte n° 13. Journal Officiel de la République
345 Française (JORF). 2018;0068:NOR: SSAP1807442A
346 3. Leuraud K, Jezewski-Serra D, Viguier J, Salines E. Colorectal cancer screening by guaiac faecal occult
347 blood test in France: Evaluation of the programme two years after launching. Cancer epidemiology.
348 2013;37(6):959-67. doi: 10.1016/j.canep.2013.07.008. PubMed PMID: 24035240.
349 4. Pornet C, Denis B, Perrin P, Gendre I, Launoy G. Predictors of adherence to repeat fecal occult blood
350 test in a population-based colorectal cancer screening program. British journal of cancer.
351 2014;111(11):2152-5. doi: 10.1038/bjc.2014.507. PubMed PMID: 25314056; PubMed Central PMCID:
352 PMCPMC4260031.
353 5. Pollack CE, Mallya G, Polsky D. The impact of consumer-directed health plans and patient
354 socioeconomic status on physician recommendations for colorectal cancer screening. J Gen Intern Med.
355 2008;23(10):1595-601. doi: 10.1007/s11606-008-0714-x. PubMed PMID: 18629590; PubMed Central
356 PMCID: PMCPMC2533392.
357 6. SPF. Taux de participation au programme de dépistage organisé du cancer colorectal 2018-2019. Saint
358 Maurice, France: Santé Publique France(SPF), 2020.
359 7. Boustière C. Pratique de la coloscopie en France : enquête “une semaine d’endoscopie en France”.
360 Hegel. 2013;3(3):219-22. doi: 10.3917/heg.033.0219.
361 8. Bernardini D, Bulois P, Barthet M, Chaussade S, Gronier O, Lecomte T, et al. « Une semaine de
362 coloscopie en France » : résultats 2017 de l’enquête annuelle de la Société française d’endoscopie
363 digestive. Acta Endoscopica. 2017;47(5):242-51. doi: 10.1007/s10190-017-0613-2.
364 9. Tuppin P, Rudant J, Constantinou P, Gastaldi-Menager C, Rachas A, de Roquefeuil L, et al. Value of a
365 national administrative database to guide public decisions: From the systeme national d'information
366 interregimes de l'Assurance Maladie (SNIIRAM) to the systeme national des donnees de sante (SNDS)
367 in France. Rev Epidemiol Sante Publique. 2017;65 Suppl 4:S149-S67. doi:
368 10.1016/j.respe.2017.05.004. PubMed PMID: 28756037.

10. Scansante. The aggregated data from the SNDS. Website: https://www.scansante.fr/applications/statistiques-activite-MCO-par-diagnostique-et-actes [updated 06/25/2023; cited 2023].
11. CNAM. Open Data platforms. Website: (https://assurance-maladie.ameli.fr/etudes-et-donnees/donnees/liste-bases-de-donnees-open-data#text_193464, [updated 202306/26/2023].
12. CNAM. list of studies carried out using the SNDS database and listed on the health insurance ("Caisse Nationale d'Assurance maladie" website https://assurance-maladie.ameli.fr/etudes-et-donnees/etudes-publications/references-publications-articles-assurance-maladie. [cited 2023 06/26/2023].
13. Ajrouche A, Estellat C, De Rycke Y, Tubach F. Evaluation of algorithms to identify incident cancer cases by using French health administrative databases. Pharmacoepidemiol Drug Saf. 2017;26(8):935-44. doi: 10.1002/pds.4225. PubMed PMID: 28485129.
14. Blotière PO, Rudant J, Barré A, Racine A, Weill A, Peyrin-Biroulet L, et al. Conditions of prescription of anti-TNF agents in newly treated patients with inflammatory bowel disease in France (2011-2013). Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2016;48(6):620-5. Epub 2016/03/28. doi: 10.1016/j.dld.2016.02.022. PubMed PMID: 27017107.
15. Blotière PO, Weill A, Ricordeau P, Alla F, Allemand H. Perforations and haemorrhages after colonoscopy in 2010: a study based on comprehensive French health insurance data (SNIIRAM). Clinics and research in hepatology and gastroenterology. 2014;38(1):112-7. Epub 2013/11/26. doi: 10.1016/j.clinre.2013.10.005. PubMed PMID: 24268997.
16. Bousquet PJ, Lefeuvre D, Tuppin P, BenDiane MK, Rocchi M, Bouée-Benhamiche E, et al. Cancer care and public health policy evaluations in France: Usefulness of the national cancer cohort. PLoS One. 2018;13(10):e0206448. Epub 2018/11/01. doi: 10.1371/journal.pone.0206448. PubMed PMID: 30379946; PubMed Central PMCID: PMCPMC6209307.
17. Doat S, Samson S, Fagot-Campagna A, Tuppin P, Menegaux F. Estimation of breast, prostate, and colorectal cancer incidence using a French administrative database (general sample of health insurance beneficiaries). Rev Epidemiol Sante Publique. 2016;64(3):145-52. doi: 10.1016/j.respe.2015.12.020. PubMed PMID: 27238161.
18. Doat S, Thiebaut A, Samson S, Ricordeau P, Guillemot D, Mitry E. Elderly patients with colorectal cancer: treatment modalities and survival in France. National data from the ThInDiT cohort study. Eur J Cancer. 2014;50(7):1276-83. doi: 10.1016/j.ejca.2013.12.026. PubMed PMID: 24447833.

19. Gouverneur A, Dolatkhani D, Rouyer M, Grelaud A, Francis F, Gilleron V, et al. Agreement between hospital discharge diagnosis codes and medical records to identify metastatic colorectal cancer and associated comorbidities in elderly patients. Rev Epidemiol Sante Publique. 2017;65(4):321-5. Epub 2017/06/04. doi: 10.1016/j.respe.2017.03.132. PubMed PMID: 28576381.
20. Janah A, Le Bihan-Benjamin C, Mancini J, Bouhnik AD, Bousquet PJ, Bendiane MK. Access to inpatient palliative care among cancer patients in France: an analysis based on the national cancer cohort. BMC Health Serv Res. 2020;20(1):798. Epub 2020/08/28. doi: 10.1186/s12913-020-05667-8. PubMed PMID: 32847565; PubMed Central PMCID: PMCPMC7448507.
21. Janah A, Gauthier LR, Morin L, Bousquet PJ, Le Bihan C, Tuppin P, et al. Access to palliative care for cancer patients between diagnosis and death: a national cohort study. Clinical epidemiology. 2019;11:443-55. Epub 2019/06/27. doi: 10.2147/clep.s198499. PubMed PMID: 31239783; PubMed Central PMCID: PMCPMC6559764.
22. Kudjawu Y, de Maria F, Decool E, M. A, J.M. H, Weill A, et al. Croisement de deux bases de données médico administratives : méthodologie et étude descriptive pour une application à la surveillance épidémiologique des cancers en France. BEH. 2013;Hors-Série:49-58.
23. Laanani M, Weill A, Carbonnel F, Pouchot J, Coste J. Incidence of and Risk Factors for Systemic Adverse Events After Screening or Primary Diagnostic Colonoscopy: A Nationwide Cohort Study. The American journal of gastroenterology. 2020;115(4):537-47. doi: 10.14309/ajg.0000000000000574. PubMed PMID: 32149780.
24. Laanani M, Coste J, Blotière PO, Carbonnel F, Weill A. Patient, Procedure, and Endoscopist Risk Factors for Perforation, Bleeding, and Splenic Injury After Colonoscopies. Clin Gastroenterol Hepatol. 2019;17(4):719-27.e13. Epub 2018/08/14. doi: 10.1016/j.cgh.2018.08.005. PubMed PMID: 30099110.
25. Lamarche-Vadel A, Pavillon G, Aouba A, Johansson LA, Meyer L, Jougla E, et al. Automated comparison of last hospital main diagnosis and underlying cause of death ICD10 codes, France, 2008-2009. BMC medical informatics and decision making. 2014;14:44. Epub 2014/06/06. doi: 10.1186/1472-6947-14-44. PubMed PMID: 24898538; PubMed Central PMCID: PMCPMC4057818.
26. Melac AT, Lesuffleur T, Bousquet PJ, Fagot-Campagna A, Gastaldi-Menager C, Tuppin P. Cancer and end of life: the management provided during the year and the month preceding death in 2015 and causes of death in France. Support Care Cancer. 2020;28(8):3877-87. doi: 10.1007/s00520-019-05188-z. PubMed PMID: 31845006.

27. Mellot E, Balcaen T, Calafiore M, Bouzillé G, Beuscart JB, Ficheur G, et al. What Is a Chronic Disease? A Contribution Based on the Secondary Use of 161 Million Discharge Records. Studies in health technology and informatics. 2019;264:263-7. Epub 2019/08/24. doi: 10.3233/shti190224. PubMed PMID: 31437926.
28. Meyer A, Drouin J, Weill A, Carbonnel F, Dray-Spira R. Pregnancy in women with inflammatory bowel disease: a French nationwide study 2010-2018. Aliment Pharmacol Ther. 2020;52(9):1480-90. doi: 10.1111/apt.16074. PubMed PMID: 33095502.
29. Meyer A, Rudant J, Drouin J, Weill A, Carbonnel F, Coste J. Effectiveness and Safety of Reference Infliximab and Biosimilar in Crohn Disease: A French Equivalence Study. Ann Intern Med. 2019;170(2):99-107. Epub 2018/12/12. doi: 10.7326/m18-1512. PubMed PMID: 30534946.
30. Meyer A, Rudant J, Drouin J, Coste J, Carbonnel F, Weill A. The effectiveness and safety of infliximab compared with biosimilar CT-P13, in 3112 patients with ulcerative colitis. Aliment Pharmacol Ther. 2019;50(3):269-77. doi: 10.1111/apt.15323. PubMed PMID: 31115919; PubMed Central PMCID: PMCPMC6767082.
31. Mitton N, Colonna M, Trombert B, Olive F, Gomez F, Iwaz J, et al. A Suitable Approach to Estimate Cancer Incidence in Area without Cancer Registry. Journal of cancer epidemiology. 2011;2011:418968. Epub 2011/04/30. doi: 10.1155/2011/418968. PubMed PMID: 21527984; PubMed Central PMCID: PMCPMC3065037.
32. Poulalhon C, Rotelli-Bihet L, Raso C, Aubry R, Fagot-Campagna A, Tuppin P. Deaths in France: Characteristics, place of death, hospitalisations and use of palliative care during the year before death. Rev Epidemiol Sante Publique. 2018;66(1):33-42. doi: 10.1016/j.respe.2017.06.008. PubMed PMID: 29203132.
33. Poulalhon C, Rotelli-Bihet L, Moine S, Fagot-Campagna A, Aubry R, Tuppin P. Use of hospital palliative care according to the place of death and disease one year before death in 2013: a French national observational study. BMC Palliat Care. 2018;17(1):75. doi: 10.1186/s12904-018-0327-z. PubMed PMID: 29764435; PubMed Central PMCID: PMCPMC5954461.
34. Quantin C, Benzenine E, Hagi M, Auverlot B, Abrahamowicz M, Cottenet J, et al. [Evaluation of medical information systems as a mean of identification of incident cases of colorectal cancer]. Sante publique (Vandoeuvre-les-Nancy, France). 2014;26(1):55-63. PubMed PMID: 24893515.
35. Quantin C, Benzenine E, Hagi M, Auverlot B, Abrahamowicz M, Cottenet J, et al. Estimation of national colorectal-cancer incidence using claims databases. Journal of cancer epidemiology.

461 2012;2012:298369. doi: 10.1155/2012/298369. PubMed PMID: 22792103; PubMed Central PMCID:
462 PMCPMC3390047.
463 36. Tanguy-Melac A, Aguade AS, Fagot-Campagna A, Gastaldi-Menager C, Sabate JM, Tuppin P.
464 Management and intensity of medical end-of-life care in people with colorectal cancer during the year
465 before their death in 2015: A French national observational study. Cancer Med. 2019;8(15):6671-83.
466 doi: 10.1002/cam4.2527. PubMed PMID: 31553130; PubMed Central PMCID: PMCPMC6825985.
467 37. Tuppin P, Riviere S, Deutsch D, Gastaldi-Menager C, Sabate JM. Burden of drug use for
468 gastrointestinal symptoms and functional gastrointestinal disorders in France: a national study using
469 reimbursement data for 57 million inhabitants. Therap Adv Gastroenterol. 2019;12:1756284819853790.
470 doi: 10.1177/1756284819853790. PubMed PMID: 31320929; PubMed Central PMCID:
471 PMCPMC6628544.
472 38. Tuppin P, Tanguy-Melac A, Lesuffleur T, Janah A, Gastaldi-Ménager C, Fagot-Campagna A. Intensity
473 of care for cancer patients treated mainly at home during the month before their death: An observational
474 study. Presse medicale (Paris, France : 1983). 2019;48(11 Pt 1):e293-e306. Epub 2019/11/18. doi:
475 10.1016/j.lpm.2019.09.044. PubMed PMID: 31734050.
476 39. Tuppin P, Pestel L, Samson S, Cuerq A, Riviere S, Tala S, et al. [The human and economic burden of
477 cancer in France in 2014, based on the Sniiram national database]. Bulletin du cancer. 2017;104(6):524-
478 37. doi: 10.1016/j.bulcan.2017.01.010. PubMed PMID: 28285755.
479 40. INCA. Algorithme de sélection des hospitalisations liées à la prise en charge du cancer dans les bases
480 nationales d'activité hospitalière de court séjour «algorithme cancer». Boulogne-Billancourt, France,
481 Institut national du cancer (INCA), 2013. Référence : OUTALGO13 Available from: https://www.e-
482 cancer.fr/Expertises-et-publications/Catalogue-des-publications/Algorithme-de-selection-des-
483 hospitalisations-liees-a-la-prise-en-charge-du-cancer-dans-les-bases-nationales-d-activite-hospitaliere-
484 de-court-sejour.
485 41. CNAM. Méthodologie médicale de la cartographie des pathologies et des dépenses, version G7(années
486 2012 à 2018). Caisse Nationale d'Assurance Maladie(CNAM): Direction de la Stratégie, des Etudes et
487 des Statistiques-Département des Etudes sur les Pathologies et les Patients. 2020 17 Fév 2020. Report
488 No.
489 42. Chapman M, Mumtaz S, Rasmussen LV, Karwath A, Gkoutos GV, Gao C, et al. Desiderata for the
490 development of next-generation electronic health record phenotype libraries. Gigascience. 2021;10(9).

491 doi: 10.1093/gigascience/giab059. PubMed PMID: 34508578; PubMed Central PMCID:
492 PMCPMC8434766.
493 43. Bronson MR, Kapadia NS, Austin AM, Wang Q, Feskanich D, Bynum JPW, et al. Leveraging Linkage
494 of Cohort Studies With Administrative Claims Data to Identify Individuals With Cancer. Med Care.
495 2018;56(12):e83-e9. doi: 10.1097/MLR.0000000000000875. PubMed PMID: 29334524; PubMed
496 Central PMCID: PMCPMC6043405.

4.4 Synthèse de la définition opérationnelle de la personne à exclure du dépistage à partir du SNDS.

Il ressort de l'enquête auprès des SGDO (***Étude 3***) qu'aucun département ou région n'exploite actuellement les données du SNDS pour affiner la population cible des campagnes. La collecte des informations en rapport avec les 11 critères d'exclusion se fait actuellement à l'aide d'un questionnaire auto-administré ou administré par un professionnel de santé selon un schéma décisionnel exposé dans ***l'étude 4*** (***Figure-4.2.1***).

En plus des trois stratégies majeures listées, plusieurs autres stratégies sont déployées par les SGDO en France de manière disparate pour réaliser les exclusions médicales. Il existait un partenariat entre les registres de cancer et seulement 5 SGDO parmi les 59 ayant répondu à l'enquête. Il faut cependant rappeler que les registres de cancer ne couvrent que 20% de la population française[69].

Les résultats de l'enquête montrent que quelques structures de gestion recevaient des données de coloscopies de certaines CPAM sans qu'un partenariat ait été formalisé pour la plupart. L'usage de ces données de coloscopie obtenue de la CPAM sans aucun compte rendu de prise en charge médico-chirurgicale devrait interroger sur la place de la coloscopie comme critère unique d'exclusion médicale, dans la mesure où une coloscopie de moins de 5 ans ne disculpe pas de l'existence d'autres situations morbides (existence d'un diagnostic excluant, ou ATCD familiaux) justifiant d'une exclusion définitive des campagnes.

Certes, les 3 stratégies majeures d'exclusion médicale permettent de collecter des données d'exclusion en rapport avec les 11 critères majeurs mais l'exhaustivité de la collecte n'est pas assurée et, il existe un risque d'inclusion à tort (au mieux) ou d'exclusion à tort (au pire). En effet, le risque que les personnes soient incluses à tort dans l'estimation des taux de participation (ou du taux d'exclusion) est nul lorsque le médecin vérifie l'éligibilité et non nul quand l'éligibilité est vérifiée que par le pharmacien ou par le patient lui-même ou quand le SNDS est utilisé. En effet, les scénarii 2, 3 et 4 ont les proportions de VP et FP respectivement <100% et >0% (***Tableau 4.4.1***).

A propos du SNDS, ***l'étude 5*** a montré qu'il contient à ce jour la quasi-totalité des données indispensables à l'estimation du taux d'exclusion médicale lors des campagnes de dépistage. L'étude 6 a aussi compilé dans un algorithme d'exclusion médicale, l'ensemble des algorithmes permettant l'extraction de ces données. Ce phénotype de la personne à exclure des campagnes de DOCCR en France devra être testé dans le SNDS. Un protocole d'accès aux données a été rédigé (***Annexe 5***), mais l'accès aux données du SNDS n'a pas été obtenu dans le temps imparti de la thèse.

Cependant, des précédentes études d'évaluation d'une des bases du SNDS (Programme de médicalisation des systèmes d'information- PMSI -) ont rapporté une sensibilité atteignant 87.7%[151], une Valeur prédictive positive atteignant 58.9%[151] et un pourcentage de concordance(coefficient kappa) atteignant 95%[101] pour l'identification des phénotypes de CCR. Ces valeurs des mesures antérieures alimentent la réflexion avec une question essentielle : A quel point le filtrage amont de la population éligible à une invitation au dépistage du CCR en par une requête SNDS améliorera la valeur du taux de participation et du taux d'exclusion ?

Dans la perspective de réaliser des exclusions en amont des invitations à l'aide d'un chaînage du SNDS avec les bases de données du dépistage, d'autres nouvelles stratégies doivent être mises en place pour minimiser le nombre de personnes non inclues à tort dans les campagnes car l'usage du SNDS génèrerait aussi des FP et FN à cause de la sensibilité du SNDS <100%. La définition de ces nouvelles stratégies devra inéluctablement passer par une évaluation du SNDS avec pour référentiel la base de données du dépistage.

Tableau 4.4.1 : Analyse d'impact des scénarios d'accès au dépistage du CRC en France.

Scénarios d'accès au dépistage du CRC				
***Scénario S1 :* Consultation médicale avec vérification complète des critères d'admissibilité (utilisation de la base de données de la SGDO ou du DMI)**				
Risque d'erreur de classification				
	Vrai statut d'éligibilité des personnes			
	Éligible	Inéligible		
Décision en faveur de l'éligibilité	VP = 100 %	FP = 0 %		
Décision en faveur de l'inéligibilité	FN = 0 %	VN = 100 %		
Conséquence d'une classification erronée				
	Décision en faveur de l'éligibilité		***Décision en faveur de l'inéligibilité***	
Vrai statut d'éligibilité des personnes	Éligible (VP)	Inéligible (FP)	Inéligible (VN)	Éligible (FN)
Appréciation de la décision d'éligibilité	Correct		Correct	
Si un test est réalisé, est-ce pris en compte dans le calcul du taux de participation?	Oui		Non	
Si aucun test n'est réalisé, est-ce pris en compte dans le calcul du taux de non-participation?	Oui		Non	
Ces personnes sont-elles incluses dans l'estimation du taux d'exclusion?	Non		Oui	
***Scénario S1 : Consultation d'*un pharmacien avec vérification incomplète des critères d'admissibilité en raison du manque d'utilisation de la CMR et de la base de données du PCSCR.**				
Risque d'erreur de classification				
	Vrai statut d'éligibilité des personnes			
	Éligible	Inéligible		
Décision en faveur de l'éligibilité	VP<100%	>0 %		

Décision en faveur de l'inéligibilité	FN = 0 %	VN = 100 %

Conséquence d'une classification erronée

	Décision en faveur de l'éligibilité		***Décision en faveur de l'inéligibilité***	
Vrai statut d'éligibilité des personnes	Éligible (VP)	Inéligible (FP)	Inéligible (VN)	Éligible (FN)
Appréciation de la décision d'éligibilité	Correct	Pas correct	Correct	
Si un test est réalisé, est-ce pris en compte dans le calcul du taux de participation?	Oui	Oui, à tort	Non	
Si aucun test n'est réalisé, est-ce pris en compte dans le calcul du taux de non-participation?	Oui	Oui, à tort	Non	
Ces personnes sont-elles incluses dans l'estimation du taux d'exclusion?	Non	Non, à tort	Oui	

***Scénario S3 : Aucune consultation avec un médecin ou un pharmacien, l'admissibilité est* établie à l'aide d'un questionnaire auto-administré.**

Risque d'erreur de classification

	Vrai statut d'éligibilité des personnes	
	Éligible	Inéligible
Décision en faveur de l'éligibilité	VP<100%	>0 %
Décision en faveur de l'inéligibilité	FN=>0 %	VN<100%

Conséquence d'une classification erronée

	Décision en faveur de l'éligibilité		***Décision en faveur de l'inéligibilité***	
Vrai statut d'éligibilité des personnes	Éligible (VP)	Inéligible (FP)	Inéligible (VN)	Éligible (FN)
Appréciation de la décision d'éligibilité	Correct	Pas de correct	Correct	Pas de correct
Si un test est réalisé, est-ce pris en compte dans le calcul du taux de participation?	Oui	Oui, à tort	Non	Non, à tort
Si aucun test n'est réalisé, est-ce pris en compte dans le calcul du taux de non-participation?	Oui	Oui, à tort	Non	Non, à tort
Ces personnes sont-elles incluses dans l'estimation du taux d'exclusion?	Non	Non, à tort	Oui	Oui, à tort

***Scénario S4 (en perspective)* : Comme le scénario 3 en raison de la sensibilité et de la valeur prédictive positive des bases de données SNDS inférieure à 100%**

DMI: *Dossier médical informatisé;* ***DOCCR :*** *Programme de dépistage organisé du cancer colorectal;* ***SNDS :*** *Système national des données de santé ;* ***FN :*** *Faux négatifs ;* ***FP :*** *Faux positifs ;* ***VN :*** *Vrais négatifs ;* ***VP :*** *Vrais positifs.*

5. AXE 3 : MONTRER QUE LES RESEAUX SOCIAUX PEUVENT ETRE UTILISES COMME SOURCES SUPPLEMENTAIRES DE MOTIVATION DES PERSONNES A PARTICIPER AU DOCCR

5.1 – Faisabilité d'une incitation à la participation aux campagnes de dépistage du cancer colorectal, par la motivation des personnes à travers les réseaux sociaux: Étude 6.

Le domaine de la santé publique n'est pas épargné par la vague de digitalisation et l'usage de l'intelligence artificielle comme outil d'optimisation des services de santé. A propos, une commission d'experts français admettait que s'il y a bien un domaine de la médecine dans lequel l'intelligence artificielle offrira de nombreuses avancées, ce sera celui de la prévention et du dépistage [159]. Cependant, comparativement aux pays anglosaxons, les réseaux sociaux (RS) sont moins utilisés en France, notamment dans le domaine de la prévention. Dans une perspective de recherche de moyens supplémentaires d'optimisation des campagnes de DOCCR, l'usage d'algorithmes d'IA comme outils de mobilisation dans le DOCCR nous a semblé innovant.

5.1.1 – CONTEXTE

En Seine-Saint-Denis, le taux de participation n'a jamais dépassé 30% depuis le début de la campagne en 2003. Du lancement du programme à ce jour, la motivation à réaliser un test se faisait principalement par courrier postal et/ou à la suite d'une consultation médicale spontanée. Pour augmenter le taux de participation, plusieurs stratégies ont été proposées sans succès notable. Dans ce département, la forte précarité sociale, les barrières linguistiques fréquentes, la faible accessibilité de certains blocs résidentiels sont des facteurs qui pourraient significativement influencer le taux de participation. Au cours d'une récente évaluation des campagnes départementales de DOCCR, plusieurs villes s'étaient distinguées avec une proportion élevée de personnes n'habitant pas à l'adresse indiquée (NPAI : > 5% de la population ciblée par la campagne). Pour connaitre les motifs de non-participation de ces personnes NPAI, la mise en place de nouvelle stratégie (sans courrier postal) de contact, nous a semblé évidente.

5.1.2 – METHODES

La stratégie de mobilisation expérimentée dans cette étude consistait en l'envoi à l'aide d'un module internet, des messages électroniques de sensibilisation et de motivation pour le dépistage du cancer colorectal à l'aide d'un algorithme de géolocalisation et d'envoi des messages électroniques via le réseau social Facebook.

L'algorithme a été développé en partenariat avec ***Medorion*** (Start-Up experte dans le domaine du développement de l'intelligence artificielle dans le domaine de la promotion de la santé). A l'image des messages publicitaires qu'elles reçoivent lors des connections aux réseaux sociaux, les personnes âgées de 50 à 74 ans vivant dans l'une des villes ciblées recevaient lors de leur connexion à Facebook, une séquence de messages didacticiels sur les moyens de prévention du CCR, notamment sur l'offre de dépistage en France. Ces messages didacticiels ont été élaborés de sorte qu'ils soient automatiquement livrés dans la langue de connexion de la personne. Une des courtes vidéos institutionnelles de l'Inca était glissée entre deux messages. Au terme de la séquence de messages, un formulaire de contact (questionnaire) était proposé à toute personne cliquant sur l'un des messages. Aucun autre message électronique ne lui était proposé après le refus du contact. Après acceptation, un médecin vérifiait l'éligibilité de la personne sur une plateforme internet régionale (www.neonetidf.org) dédiée. Le médecin de l'association pouvait programmer un entretien téléphonique pour les éventuelles exclusions médicales. Cette vérification de l'éligibilité pouvait se faire aussi par le médecin de la structure de gestion du DOCCR. Dans tous les cas, les informations fournies par les personnes étaient automatiquement transférées par le module sur un espace sécurisé de stockage conforme à la nouvelle réglementation du RGPD (Règlement Général sur la Protection des Données). Si la personne était éligible au dépistage, il était possible de lui proposer à travers le module :

- Le retrait d'un kit de test de dépistage soit à l'adresse de l'association de dépistage du CCR, soit au service de gastroentérologie de l'hôpital Avicenne, soit chez son médecin traitant (s'il en a un), soit à la structure de gestion du DOCCR à Bondy.
- L'envoi d'un kit de test de dépistage seulement si la personne a une adresse postale fiable et que toutes les questions de contre-indication du dépistage lui ont été posées par le médecin en charge de l'envoi postal dans la structure de gestion.

Pour tester la fiabilité du module de contact et faire une estimation des résultats probants d'une telle action, le module de contact a été ouvert pendant deux mois (Décembre 2018 et janvier 2019), dans quatre communes (***Aubervilliers, La Courneuve, Saint-Denis et Saint Ouen***) ayant un taux de NPAI≥10% et deux communes (***Le Bourget, Clichy-sous-Bois***) ayant un taux de NPAI<10% lors de 3 campagnes antérieures à 2017. Le choix de ces communes ne reposait sur aucune procédure statistique.

5.1.3 – SYNTHESE DES RESULTATS

Selon l'estimation faite par l'algorithme en amont, environ 67 000 personnes âgées de 50-74ans (âge déclaré sur le réseau social Facebook) se connectaient régulièrement au réseau Facebook en décembre 2018 dans les

communes ciblées. A l'ouverture du module, l'objectif initial était d'atteindre 50 000 personnes sur les 67 000. Au terme des deux mois de promotion dans ces six communes, 39 900 personnes ont pu être touchées par la campagne et 9,200 (18,4% de la population cible) ont pu regarder au moins une des vidéos diffusées sur Facebook par le module. La plupart des personnes ont regardé les vidéos en dehors des heures de travail, l'heure de pointe étant située entre 18h et 22h. La campagne a pu atteindre des personnes de tous les âges ciblés par le DOCCR, y compris les personnes âgées de plus de 65 ans. La campagne a intéressé les deux sexes dans les proportions similaires quel que soit le groupe d'âge pertinents. La plupart des personnes ont regardé les vidéos via un téléphone mobile(91%) ou un ordinateur de bureau. 4450 (8,9% de la population cible) personnes ont cliqué sur l'un des messages de promotion pour en savoir davantage sur le DOCCR.

Parmi les 4,450 personnes ayant cliqué sur un message, 298 (6,7%) ont finalisé la procédure de contact en réalisant une demande de kit de test de dépistage. Avec une moyenne d'âge estimée à 54 ans, 160 (88 femmes et 72 hommes) des 298 (53,7%) demandeurs de kit de tests étaient éligibles après vérification sur le site d'éligibilité. 38 personnes, bien qu'éligibles, résidaient dans un département autre que la Seine-Saint-Denis. 36 personnes étaient non éligibles soit pour un test de dépistage récent (28), soit une coloscopie récente ou un suivi d'une maladie du colon en cours (6). 16 personnes étaient hors tranche d'âge du dépistage et l'éligibilité de 48 personnes n'a pu être vérifiée probablement à cause de la non-fiabilité des informations fournies. Parmi les 160 personnes éligibles, 9 ont été NPAI au moins une fois et 104 n'avait jamais réalisé un test de dépistage. Chez ces personnes ayant demandé le kit, 84% des demandes ont été effectuées en dehors des heures habituelles (9H-17H) de travail et 100% des 298 personnes ont fourni des adresses électroniques et postales fiables, un numéro de téléphone fiable (portatif dans 100% des cas).

A la date du 30 juin 2019 (soit 4 mois après la fermeture du module), 67 des 160 ayant reçu le kit après la demande sur Facebook, avait réalisé le test dans un délai médian de 14 jours, soit un taux de réalisation de 41.9%. Bien que fermé pour les demandes de kit, le module permettait de faire un rappel par SMS toutes les deux semaines, pour les personnes n'ayant pas encore réalisé un test. De même, un rappel par courriel était réalisé toutes les deux semaines jusqu'au 30 juin 2019. Avec ce rappel régulier par SMS et par courriel, l'algorithme avait estimé le taux de participation à 89,5% (soit 143 kits utilisés sur les 160 envoyés dans un délai de 12 mois après la fermeture du module. Le résultat du test était positif pour 3 tests (4.5%) sur les 67 réalisés. Une seule des trois personnes ayant un résultat positif avait réalisé la coloscopie à la date du 30 juin 2019.

Original Article

doi:10.1111/codi.15121

Feasibility of encouraging participation in colorectal cancer screening campaigns by motivating people through the social network, Facebook

A. Koïvogui*, S. Levi†, M. Finkler†, S. Lewkowicz‡, T. Gombeaud§, J. M. Sabate¶, C. Duclos,†† and R. Benamouzig¶**

*CRCDC-IDF, Site de Seine-Saint-Denis, Bondy, France, †Medorion Technologies Ltd, Tel Aviv, Israel, ‡Zav-Menia NGO, Tel Aviv, Israel, §Association de Dépistage du Cancer Colorectal (ADCCR), Clinique d'Hépato-Gastro-Entérologie, Les Lilas, France, ¶Service d'Hépato-Gastro-Entérologie, Hôpital Avicenne (AP-HP), Bobigny, France, **Laboratoire d'Informatique Médicale et d'Ingénierie des Connaissances pour la e-Santé (LIMICS), INSERM 1142, Univesité Paris 13, Bobigny, France, and ††Unité d'Informatique Hospitalière et de Traitement de l'Information Médicale - Service de de Santé Publique, Hôpital Avicenne (AP-HP), Bobigny, France

Received 3 January 2020; accepted 24 April 2020; Accepted Article online 13 May 2020

Abstract

Aim To describe the results of a feasibility phase and the expected results of a new approach to increase the participation rate in a Colorectal Cancer Organized Screening Program (CRCSP) through Facebook awareness messages.

Method This approach targets people aged 50–74 years, who reside in an urban deprived area and regularly connect to Facebook. The feasibility phase ran over 2 months (December 2018 and January 2019) in six municipalities (Seine-Saint-Denis, France). The full provisional campaign will run over a year. The approach consists of sending electronic awareness messages on the importance of screening for colorectal cancer using a specific Facebook module. Subjects who consent to screening complete a test-kit application form. The eligibility of each subject to participate in screening is determined by a doctor before the kit is sent out.

Results A total of 39 900 people were reached by the feasibility phase campaign, and 9200 were able to watch at least one Facebook message/video. Of those, 4450 people logged to learn more about the CRCSP, 298 applied for a test kit, 160 test kit applicants were eligible to participate and the test completion rate was 41.9%. According to these feasibility results, 366 120 targeted people would connect regularly in the tested area, 141 541 of whom would be interested in a specific promotional message posted on Facebook. Requests could be made for 9770 kits, with 5246 people being eligible to participate in screening. The expected test-completion rate is estimated at 42%–89%. This would represent 5%–11% of the tests carried out in the area during the same period by 'classical' CRCSP.

Conclusion Implementation of the Facebook strategy would significantly improve the rate of participation in the CRCSP by mobilizing people with no previous participation, including younger subjects.

Keywords Colorectal cancer screening, Facebook, awareness campaign, participation rate

What does this paper add to the literature?
In the French Colorectal Cancer Screening Program, several strategies have been tested but the participation rate rarely exceeded 50%. This study presents a new approach to increase the participation rate through the social network Facebook by mobilizing people with no previous participation, including younger subjects.

Introduction

In France, since the launch of the Organized Colorectal Cancer Screening Program (CRCSP) in 2000, participation, to date, has been driven by systematic postal invitation to the target population or self-referring medical consultation [1–3]. With this strategy, the participation rate has rarely exceeded 50% in the French departments between 2009 and 2018 [3]. To optimize the participation rate, several strategies have been tested [4–14]. These strategies were mostly based on motivation, messages and reminders sent by postal mail

Correspondence to: Robert Benamouzig, Hôpital Avicenne (AP-HP), Service d'Hépato-Gastro-Entérologie, 125 Rue de Stalingrad, 93000 Bobigny, France.
E-mail: robert.benamouzig@aphp.fr

[signed, in some cases, by general practitioners (GPs)], by telephone or by email. A postal mailing reminder with the test kit, initially implemented, was stopped because of the high rate of unused test kits but was restarted in 2018 for those who had participated in at least one of the previous three campaigns. Such a strategy is not adapted to people who do not consult a GP (such as young adults in employment who are in good health, people with a social barrier for consultation or people living in an area with no medical doctor) or those with a GP who does not recommend colorectal cancer (CRC) screening [15].

Strategies to implement health prevention through the Facebook network have previously been used [16–19]. There is growing evidence to suggest that Facebook is a useful tool and its use should be considered for public health action [17]. In the study by Harty *et al.*, Facebook advertisement proved to be successful in engaging community members in the outreach campaign for cancer prevention, with over 22 600 exposed to such advertisements [16]. The use of artificial intelligence (AI) in prevention and promotion programmes has also marked the start of an ambitious phase in targeted marketing [20,21]. However, the combination of a geolocation model to target specific populations and sending personalized awareness messages through a social network (Facebook) is still poorly documented in terms of impact on participation rate during a CRC screening campaign.

Concerned by the need for additional actions in promoting CRC screening, several partners (listed in the Acknowledgements) decided to implement a complementary strategy of sending messages of awareness and invitation through the social network Facebook. These messages are aimed at anyone usually targeted by the CRCSP campaign, especially those not reached by classical postal mail. The present paper aims to describe the results of a feasibility phase and the expected results of such an action according to the results of the feasibility phase.

Method

Targeted population

This promotion and awareness-raising action for participation in screening campaigns targets the population aged 50–74 years who are regularly connected to the social network Facebook. Regular connection was defined as the existence of an active Facebook account that had been used within the 48 h preceding the estimation by the geolocation module. The feasibility was tested over 2 months (December 2018 and January 2019) in four Seine-Saint-Denis municipalities with a high rate of undelivered postal mail because of people not living at the indicated address.

Conduct of the motivation action in accordance with organization of the CRCSP campaign

In France, CRCSP campaigns are organized in accordance with national specifications [1,2,22]. In Seine-Saint-Denis (an urban department of 1 623 111 inhabitants in 2017), the 'Comité Départemental des Cancers' (known as 'CRCDC-IDF, site de Seine-Saint-Denis' since 1 May 2019) was the local cancer screening managing centre. According to the CRCSP, people receive the first screening invitation on their 50th birthday and then every 2 years thereafter until they reach 74 years of age. If a positive test result of obtained, the patient is referred for a total colonoscopy.

Each Facebook motivational campaign consists of sending electronic messages of awareness and motivation for CRC screening via the social network Facebook using an internet module. This geolocation and electronic message sending module was developed by a partner, Medorion®, a startup firm expert in the development of AI in the field of health promotion. Medorion® has proprietary AI algorithms that use demographic data, clinical data and online engagement to match the optimal content per person, greatly increasing the chances of engagement and thus the chances for kit return. Medorion® has developed an online platform that facilitates the management team engagement process, which includes: segmenting the audience (in terms of Facebook login language); building engagement journeys that include Facebook advertisements, emails and text messages; tracking campaign progress, such as the number of requests for testing kits received via Facebook and email, and the number of tests completed and returned, alongside analytics tools to manage long-term goals.

People aged 50–74 years and living in one of the targeted municipalities receive a sequence of tutorial messages on CRC prevention, particularly on the free CRC screening offered in France (Table 1). These tutorial messages come in several languages. Between two sequences of messages broadcast in the person's login language, short institutional videos of the National Cancer Institute (INCa) are started. At the end of the message, a contact form (questionnaire) is offered. No other messages are offered if the contact form is refused (Fig. 1). If the person agrees to be contacted, a medical doctor from the CRCSP's managing centre verifies their eligibility. The medical doctor may schedule a telephone interview to check eligibility for screening. Indeed, in France, people at high risk (e.g., those who have a

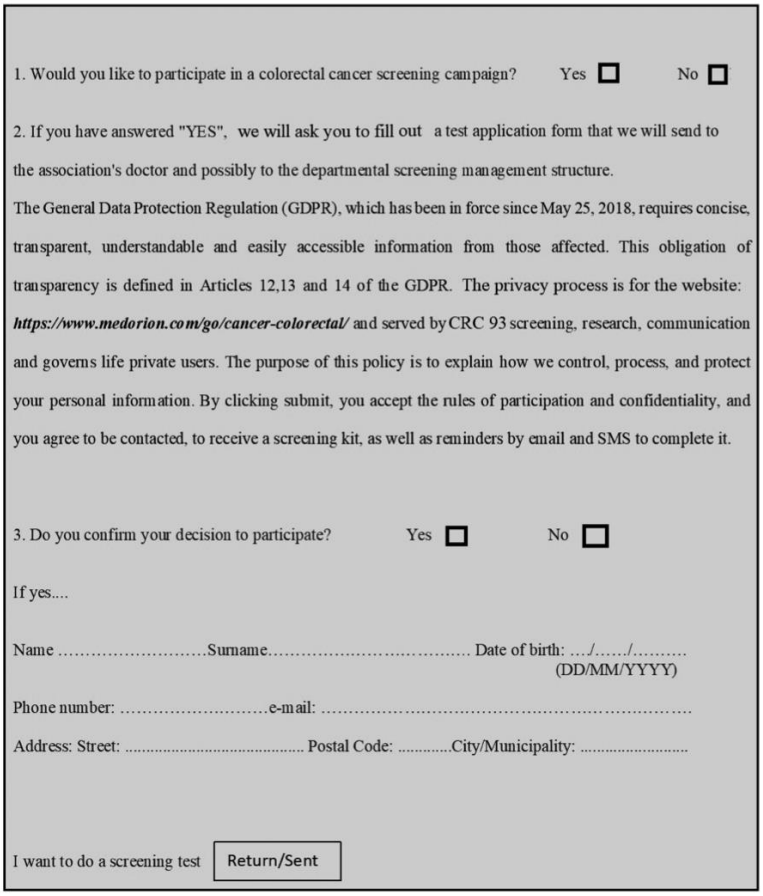
personal or family history of colorectal adenomas, cancer or inflammatory diseases such as ulcerative colitis and Crohn's disease) or at very high risk (e.g., those with familial adenomatous polyposis or Lynch syndrome) of developing CRC and people with severe extra-intestinal pathology are excluded from faecal test screening campaigns. If the person is eligible for screening, it would be possible to offer to them, through the module, a screening test kit either sent by the CRCSP's management structure or by other project partners (GP, gastroenterology department). Sending a screening test kit is only possible if a reliable postal address is provided. In all cases, the information provided by the individuals was automatically transferred by the module to a secure storage space in accordance with the new regulations of the General Data Protection Regulation (GDPR).

Feasibility phase

To test the reliability of the module and to estimate the feasibility of such an action, the contact module was opened for 2 months (December 2018 and January 2019) in six municipalities in the department of Seine-Saint-Denis. This socially deprived department was chosen because of the low participation rate (which has never exceeded 30% since the official start of the CRCSP's campaign in 2009) [1] of its residents, with only 13.9% of the test kits used in the department sent directly by post in 2017 [1]. Similarly, during a recent evaluation of CRCSP's departmental campaigns, several municipalities had a high proportion of not living at the indicated address (NPAI), greater than 5% of the population targeted by a campaign (Activity Report 2017, Comité Départemental des Cancers-CDC93-, 2018) (Fig. 2). The choice of the six municipalities was not based on any procedure except that they were the first in alphabetical order on two separate lists. The first list included all municipalities with an NPAI rate of ≥ 10% and a participation rate of < 15% in the 2017 campaign. Four municipalities (Aubervilliers, La Courneuve, Saint-Denis and Saint-Ouen) were selected from this list. The second list included all municipalities with an NPAI rate of ≥ 10% and a participation rate of ≥ 15% or an NPAI rate of < 10% and a participation rate of < 15% in the

Table 1 Tutorial message sequences: How to prevent colorectal cancer.

Sequence-1 (duration: 10 s): What is colorectal cancer?
Colon cancer develops from the cells that line the inner lining of the colon. In more than 80% of cases, it comes from a benign tumour, called an adenomatous polyp, which progresses slowly and eventually becomes cancerous. In France, colorectal cancer is the third most common cancer in both sexes: it is the second most common cancer in women and the third most common cancer in men. The vast majority of cases of colorectal cancer occur in people aged 50 years of age and over. In 2015, this cancer killed 17 833 (9337 men and 8496 women) people in France
Sequence-2 (duration: 10 s): Am I at risk for colorectal cancer?
If you are over 50, then yes, you are at risk and this risk will only increase with age. Several lifestyle factors (diet, overweight, physical inactivity, etc.) are also risk factors for colorectal cancer
Sequence-3 (duration: 10 s): But, what if I don't have any symptoms?
Colorectal cancer is often called the 'silent killer' because it can develop without symptoms. Unfortunately, by the time symptoms begin to appear, it is likely that the cancer is at a stage where it is more difficult to treat. Therefore, it is important to have colorectal cancer screening even if you are feeling well. Periodic screening for colorectal cancer can help you prevent or detect the disease quickly
Sequence-4 (duration: 10 s): So, what should I do?
We recommend that you get tested every 2 years if you are 50 years old. The fecal immunochemical test (FIT) is very user-friendly and is done at home. If you have recently had a colonoscopy, please check with your doctor before doing a FIT test
Sequence 5 (duration: 10 s): What is the FIT test?
The fecal immunochemical test (FIT) is an examination that can detect small amounts of occult blood (hidden) in the stool. The test is quick and easy to perform and requires no special preparation. In France, it is the test used in the colorectal cancer screening programme. However, the presence of occult blood in the stool can be caused by other diseases of the digestive tract, such as Crohn's disease, which is why a colonoscopy is performed after a positive test result is obtained
Sequence-6 (duration: 10 s): Where can I get a screening test kit?
You can request that the FIT kit be posted to you by submitting the form below. Alternatively, you can purchase the kit from your GP or other GP, or the screening management structure in your home department
Sequence-7 (duration: 10 s): How do I do FIT?
The test is done at home, on the stool as such, it has no side effects. It involves placing a sample of your stool (feces) in a tube and sending it to a laboratory in the postage-paid reply envelope supplied with the kit. You don't need to make special preparations for this test or change your diet the day before. After doing the test, fill out the fact sheet, place it in the envelope with the test sample, seal the envelope and post. You will have the results in about 15 days


1. Would you like to participate in a colorectal cancer screening campaign? Yes ☐ No ☐

2. If you have answered "YES", we will ask you to fill out a test application form that we will send to the association's doctor and possibly to the departmental screening management structure.
The General Data Protection Regulation (GDPR), which has been in force since May 25, 2018, requires concise, transparent, understandable and easily accessible information from those affected. This obligation of transparency is defined in Articles 12,13 and 14 of the GDPR. The privacy process is for the website: ***https://www.medorion.com/go/cancer-colorectal/*** and served by CRC 93 screening, research, communication and governs life private users. The purpose of this policy is to explain how we control, process, and protect your personal information. By clicking submit, you accept the rules of participation and confidentiality, and you agree to be contacted, to receive a screening kit, as well as reminders by email and SMS to complete it.

3. Do you confirm your decision to participate? Yes ☐ No ☐

If yes....

Name ……Surname…… Date of birth: …/……/…… (DD/MM/YYYY)

Phone number: ……e-mail: ……

Address: Street: ..... Postal Code: .....City/Municipality: .....

I want to do a screening test Return/Sent

Figure 1 Module contact form.

2017 campaign. Two municipalities (Clichy-sous-bois and Le Bourget) were selected from this list.

In addition to the telephone interview to identify possible causes for medical exclusion, eligibility was checked either on the dedicated regional internet platform (www.neonetidf.org) or by contacting the CRCSP's management structure in the department. During the feasibility phase, the test kit had been posted to all eligible applicants by the CRCSP's management structure or by the gastroenterology department of the Avicenne University Hospital.

Legal provision

The CRCSP database is supported by the National Commission for Information Technology and Freedoms (CNIL) [23]. According to the current French legislation, a study that does not change the care of patients does not require permission from the Clinical Research Centers Ethics Committee. This study does not change either the type of screening test or the approach to taking care of people participating in screening campaigns. The module developed by Medorion® is consistent with the new regulation of the GDPR.

Results

Feasibility phase results

Before starting the feasibility phase (November 2018), the number of people aged 50–74 years (age declared) regularly connected to the Facebook network in the six municipalities was estimated by the module (Table 2). About 67 000 people were regularly connected in the six municipalities (15 000 in Aubervilliers, 6900 in Le Bourget, 5700 in Clichy-sous-Bois, 7600 in La Courneuve, 8600 in Saint-Denis and 23 000 in Saint-Ouen).

At the beginning of the feasibility phase, the initial goal was to reach 50 000 people out of these 67 000 identified as regularly connected. At the end of the feasibility phase, 39 900 people from these six municipalities were reached and 9200 (18.4% of the target

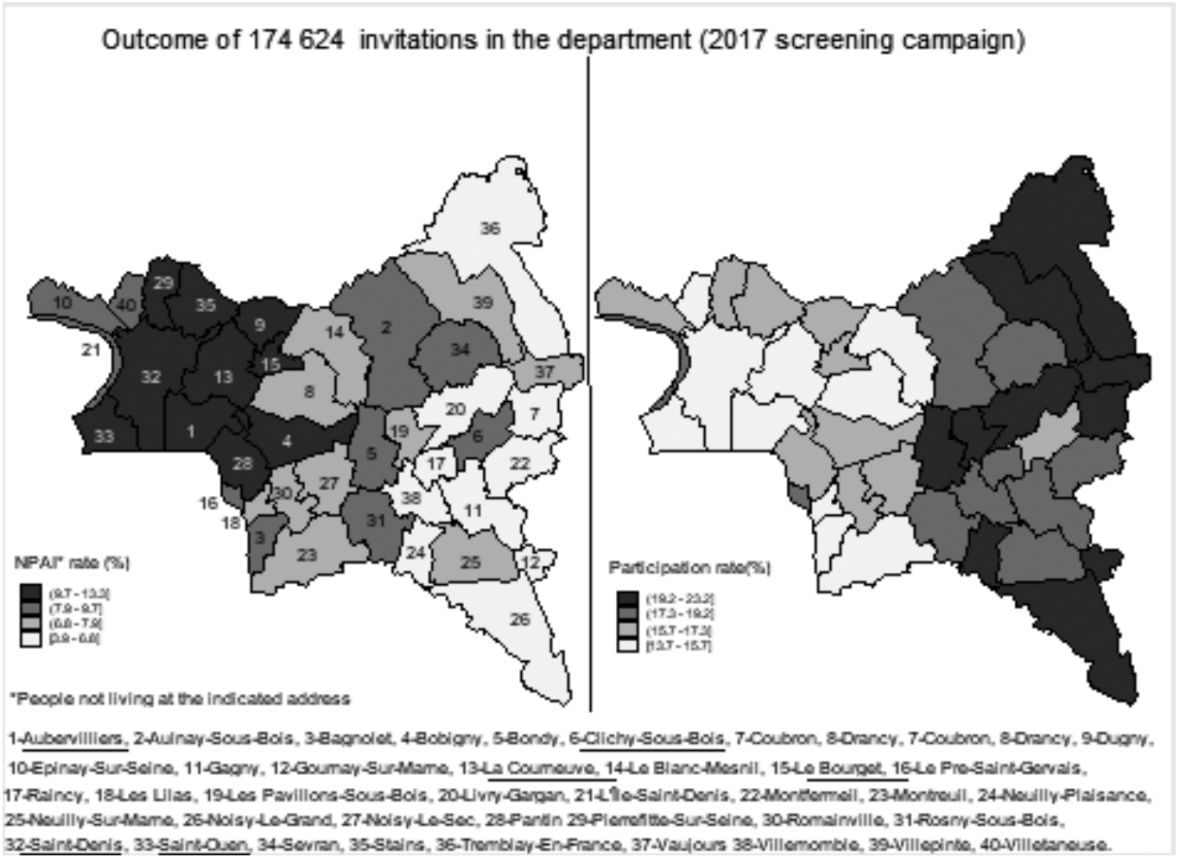


Figure 2 Campaign NPAI rate and Campaign participation rate, 2017. The municipalities chosen for the pilot phase are underlined.

population of 50 000) watched at least one of the videos posted on Facebook by the module. The campaign attracted both sexes in similar proportions, regardless of age. Most (91%) people watched the videos via a mobile phone (Fig. 3). Most people watched the videos outside working hours, with rush hour between 18:00 h and 22:00 h. In total, 4450 (8.9% of the target population) people clicked on at least one of the promotional video/messages to learn more about the CRCSP (Table 2).

Among these 4450 people, 298 (6.7%) requested a test kit. 84% of the test kit requests were made outside normal working hours (09:00–17:00 h) and 100% provided reliable email and postal addresses and telephone numbers (for a mobile in 100% of cases). After verification, 160 (mean age: 54 years; 88 women) of the 298 (53.7%) applicants for the test kit were eligible for CRC screening. Thirty-eight people although eligible, resided outside the targeted department. Thirty-six people were ineligible as a result of either a recent screening test ($n = 30$) or a recent colonoscopy ($n = 6$). Sixteen people were outside the age range (50–74 years), and the eligibility of 48 people could not be verified because of the unreliability of the information provided. Among the 160 eligible persons, 47 had previously carried out a screening test at least once, nine had not received any previous invitation because of their NPAI status and 104 had received an invitation but had never been screened.

Five months after the module closed (30 June 2019), 67 of the 160 people who received the kit after a request through Facebook had completed the test within an average of 23 days (median = 14 days) (i.e., a completion rate of 41.9%). After the end of the 2-month feasibility period, the module was still used to send SMS and email reminders every 2 weeks, until 30 June 2019, to all people provided with a test kit who had not yet completed it. The content of these messages (SMS/email) is summarized in Table 3. According to the module estimation, the test completion rate could reach 89.5% (143 kits used out of 160 sent) with this regular SMS/email reminder, 12 months after closure of the module.

Table 2 Results observed during the 2-month campaign in six municipalities, and the results expected during a 12-month campaign in the department.

	Results in the six municipalities			Results in the department	
	Observed in 2-month campaign	Expected in 12-month campaign		Expected in 12-month campaign	
Type of expenditure	No. of people	No. of people	Estimation method	No. of people	Estimation method
Total population aged 50–74 years in target area (2017–2018 campaign)	75 380	75 380	None	411 914	None
Number of people invited (2018 campaign)	38 758	38 758	None	211 793	None
Weight (%) of the six municipalities in the number of departmental invitations during the 2018 campaign*	–	–	–	18.3[A]	None
Number of people aged 50–74 years, regularly connected to the Facebook network	67 000[B]	67 000[B]	By the module*	366 120[C]	(B × 100)/A
Campaign target population among these people aged 50–74 years, regularly connected to the Facebook network	50 000	67 000[B]	By the module*	366 120[C]	–
Number of Facebook users 'reached' (viewed a promotional message)	39 000	67 000[B]	By the module*	366 120[C]	–
Number of people who watched at least one video/message posted	9200	55 200[D]	9200 × 6	301 639	(D × C)/B
Estimated click-through rate (%) over 12-month campaign	–	38.6[E]	By the module*	–	–
Number of users who clicked on the promotional video/message posted	4450	25 902	(B × E)/100	141 322	(C × E)/100
Number of test kits requested	298	1788[F]	298 × 6	9770	(F × 100)/A
Number of people eligible/sent test kit	160	960[G]	160 × 6	5246	(G × 100)/A
Number of test kits returned (5 months after the module closed)	67	402[H]	67 × 6	2197	(H × 100)/A

*The artificial intelligence module developed makes it possible to make estimates of parameters of an action at a time T.
The capital letters index the numbers on which they are placed as a superscript in columns 2, 3 and 5. For the estimates made in the last column, each capital letter must be replaced by the number indexed by the superscript letter in the previous columns.

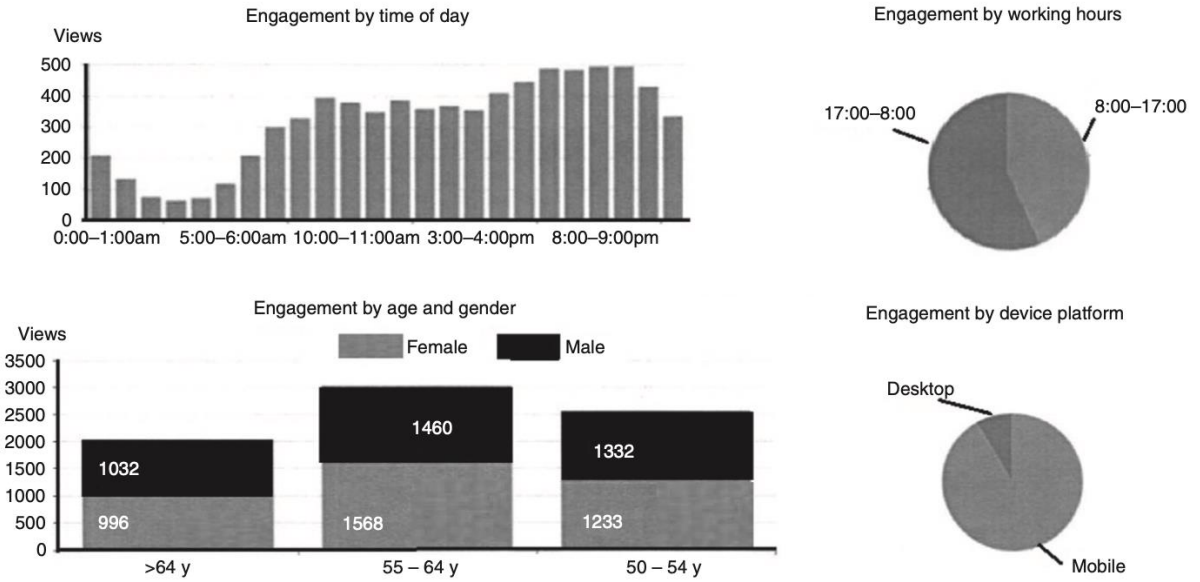


Figure 3 Funnel analytics. y, years.

Table 3 Content of tracking messages.

First SMS after test kit request through Facebook
Thank you for finalizing the test kit application process. Your application will be processed by the staff doctor who will contact you to establish your eligibility. Kind regards
SMS notification of sending the test kit
Your decision to perform a colorectal cancer screening test was an excellent choice. Your eligibility has been established after a telephone interview with the staff doctor. A test kit was sent to you today by postal mail. This test is useless if not used! It will only take a few minutes to complete and requires no preparation
Recall SMS after sending the Test Kit
Your decision to perform a colorectal cancer screening test was an excellent first step. But the test that was sent to you is useless if it is not used! The test takes only a few minutes and requires no preparation. This is a kind reminder to carry out the test and thus consolidate your noble decision to protect yourself from colorectal cancer. If you have already completed and returned the test, thank you and please ignore this message
For more information, call: 00 00 00 00 00.
Wishing you health
Recall email after sending the Test Kit
Subject: a kind reminder to complete the cancer screening test
Dear . . .
Your decision to perform, at home, the screening test for colorectal cancer was an excellent first step. But the test that was sent to you is useless if it is not used! The test takes only a few minutes and requires no preparation
This is a kind reminder to carry out the test and thus consolidate your noble decision to protect yourself from colorectal cancer. If you have already completed and returned the test, be sure to receive our thanks and ignore this message
For more information, call: 00 00 00 00 00
Wishing you health

To carry out the feasibility phase, a fixed cost of €15 000 was discussed with the partner MEDORION (Table 4). The package included: (i) unlimited dissemination, during the feasibility phase, of messages through Facebook to 50 000 people in the six chosen municipalities; (ii) sending email/SMS messages to a maximum of 5000 people who had agreed to learn more about the CRCSP; and (iii) the cost of the MEDORION licenses/software. The fixed cost does not include the cost of the test kit and the cost of database management.

Estimated departmental number

People aged 50–74 years who regularly connected to the Facebook network

According to the CRCSP's management structure, 211 793 people aged 50–74 years were targeted by the 2018 departmental campaign. Of those, 18.3% resided in the six municipalities included in the feasibility phase, and 67 000 people regularly connected to Facebook in these six municipalities; thus, we deduce that 366 120 (67 000 × 100/18.3) connect regularly to Facebook throughout the department. By extrapolating the results observed in the feasibility population to a 1-year campaign, with a similar click-through rate (38.6%), the number of people clicking on a message/video could reach 25 902 subjects in these six municipalities and 141 322 subjects in the department (Table 2).

People who can request a test kit through Facebook

In 12 months, 1788 people in the six municipalities and 9770 people in the department could finalize the contact procedure and request a test kit. The number of eligible persons could reach 5246 in the department, including 295 people with NPAI people and 3410 people without previous history of participation in any CRCSP campaign.

People aged 50–74 years performing the test after a kit request through Facebook

Among 211 793 people invited to the 2018 annual CRCSP campaign (Seine-Saint-Denis), 42 663 (20.1%) performed a test before 30 June 2019 (6 months after the last invitation). The proportion of positive test results was 4.1%, the colonoscopy completion rate was 77.4% and 1106 neoplasms (98 cancers, 1008 polyps) were diagnosed in this target population.

The number of tests performed in the first 5 months through a Facebook campaign was estimated at 2197 out of 5246 likely eligible people. This would represent 5.1% (2197/42 663) of all tests performed in the CRCSP campaign 6 months after the last invitation. If the SMS/email tracking process that the module allows

Table 4 Estimated cost (€) of carrying out a Facebook campaign (Reference: year 2018).

	Cost of a 2-month campaign in six municipalities				Estimated cost of a 1-year campaign in the department			
Type of expenditure	Unit cost (€)	Number of people	Repetition	Total cost (€)	Unit cost (€)	Number of people	Repetition	Total cost (€)
Package (Medorion service and Facebook) cost	15 000	–	–	15 000	150 000	–	–	150 000
1-Package Contents								
Targeted Facebook members		50 000	–			≤ 367 000	–	
Licenses (Includes all Medorion and 3rd party licenses)		Single user account	Unlimited			Single user account	Unlimited	
Professional services (all professional services in the Scope of Work section are included)		Unlimited	Unlimited			Unlimited	Unlimited	
Customization (unlimited reports, journeys, content customization, user access)		Unlimited	Unlimited			Unlimited	Unlimited	
Media expenses (Facebook, email and SMS)		≤ 5000	Unlimited			≤ 50 000	Unlimited	
2-Package consumption after module closing								
Video/commercial broadcast by the module		39 900	Unlimited			366 120	Unlimited	
Click on the contact form		4450	Unlimited			141 541	Unlimited	
First contact SMS-1 after kit request		298	Unlimited			9770	Unlimited	
SMS-2 information after sending the test kit		160	Unlimited			5246	Unlimited	
SMS-3 recall if test kit not used		160	Unlimited			5246	Unlimited	
Recall email if test kit not used		160	Unlimited			5246	Unlimited	
Subtotal (awareness message and contact process)	15 000			15 000	150 000	–	–	150 000
Test kit mailing								
Salary cost (secretary)	2.10	160	1	336	2.10	5246	1	11 016.60
Telephone call (toll-free number)	1.49	160	1	238.4	1.49	5246	1	7816.54
Material cost								
Envelope A4	0.202	160	1	32.32	0.202	5246	1	1059.69
Invitation Form	0.024	160	1	3.84	0.024	5246	1	125.90
Immunological test kit	0.00	160	1	0	0.00	5246	1	0†
Flyer reformatting	0.03	160	1	4.8	0.03	5246	1	157.38
Printing	0.041	160	1	6.56	0.041	5246	1	215.09
Post stamp	1.38	160	1	220.8	1.38	5246	1	7239.48
Subtotal (test kit mailing)	5.267	160	1	842.72	5.267	5246	–	27 630.69
Total				15 842.72€				177 630.69€

*Unit cost decreasing with number of targeted populations.
†Estimate impossible because fully supported by social security.

is maintained 12 months after the module closes, the number of tests that can be achieved through a Facebook campaign could reach 4695 (0.895 × 5246), and this would represent a potential increase of 11% in the CRCSP departmental participation rate.

Cost of an annual Facebook campaign

According to the current CRCSP organization, the purchase price of the test kits is entirely covered by health insurance (not estimable in this approach). The cost of mailing the test kit is limited to salary costs (secretary), logistics and equipment (Table 4). The cost of the campaign was estimated at €99.02 (€15 842.72/160) per unit of kit sent in this pilot phase. This campaign cost is estimated at €33.86 (€177 630.69/5246) per unit of test kit (departmental campaign).

Discussion

Social network use is spreading in all areas of medicine. Prevention and screening are potential fields for this new approach. Seine-Saint-Denis displays the characteristics of a priority area in the field of prevention and public health. This urban department is the most cosmopolitan in France and the need for health intervention is largely supported by the characteristics of its population [1]. As a reminder, the relationship between the participation rate in a screening campaign and the socio–economic level has been well established [24,25]. In order to optimize the participation in CRCSP in this department, new strategies were needed by the regional health authority [26]. Lack of reliable postal addresses and language barriers are obstacles identified for CRCSP.

Electronic messages to patients have been shown to increase the participation rate in CRC screening [13,14]. Social media marketing may be a new way to enrol hard-to-reach populations into cancer screening educational initiatives [16]. The scheme proposed in this methodological approach makes it possible to perpetuate the effect of electronic messages by combining two proven communication methodologies: the Persuasive Health Message (PHM) framework developed by Witte [27] and the relevance of the source of the information argued by Smith *et al.* [28]. Witte identifies *the constant* and *transient* factors in the development of a compelling health message. Like the study of Cutrona *et al.* [29] this study also proposes components adapted to the French context, including: (i) *threat* messages in which people connected to Facebook feel likely to be a healthy carrier of CRC (serious threat); (ii) *efficacy* messages about the screening programme, in which people are convinced that they are able to perform a screening test (self-efficacy) and that this action effectively distances the threat; (iii) various *cues* indirectly influencing the persuasive process (the credibility of project partners, including the Avicenna Hospital Gastroenterology Department); and (iv) *targeting a specific audience* (50–74 years). The messages are delivered in several languages, which is a guarantee of achieving the objectives of the message and respect for beliefs, culture and preferences in France's most cosmopolitan department. Similarly, the repetition of the electronic messages and the display of the coordinates of the various project partners (known to all) will lead to memorization of the objectives by the people, because according to Smith [28], memorized messages have the ability to influence behaviour when they are retained in the memory long after the initial exposure.

The cost per unit of the test kit sent in the pilot phase remains consistent with the usual cost of posting the test kit, estimated at around €33.59 per person [30]. Since the public health research mission is now assigned to a screening programme-organizing body [22], the cost of the action could be limited to the cost of the Facebook campaign and the software's license purchase price for large-scale use. This dematerialization action could optimize the participation rate by 5.1%–11% and is ideal for mobilizing a young population (aged 50–54 years), unmotivated by CRCSP campaigns. It could allow a follow-up, at lower cost (email/SMS), of all test kits sent to people. Similarly, for all positive tests, follow-up to colonoscopy could be considered at a lower cost (SMS/email).

To sustain such a complementary strategy in the CRCSP, an evaluation (epidemiological/medico–economic) of its impact should use one of two schemes. The first scheme involves promotion for 1 year, in the population of 366 120 people regularly connected to Facebook throughout the department. In this case, the sample will be compared with itself in a 'before/after' scenario (Scheme-1). The second scheme involves promotion for 2 years, in half of the population of 366 120 people regularly connected to Facebook. In this case, the other half would serve as a control group in a cluster randomized trial approach (Scheme-2). In the French context, the second scheme should be preferred.

Limits of the study

In such a mobilization strategy through social networks, the risk of a false projection of results is not zero. Indeed, a technological failure in the use of AI could underestimate (at best) or overestimate (at worst) the number of elderly people who regularly connect to

Facebook. However, the results obtained after a 2-month campaign allow us to conclude that the effectiveness of such an initiative measured in a long-term campaign could be optimal.

The cost of the campaign also seems to be very high and therefore very demotivating for the main funders of screening campaigns. However, this cost could be lower than the cost of treating cancers that were not diagnosed early in this population because of the absence of a complementary mobilization strategy.

Conclusion

The strategy to mobilize people through the social network Facebook is feasible. Implementation of this strategy will improve the participation rate through the mobilization of young people with no participation history. However, in order to sustain such a complementary strategy in the CRCSP, it seems obvious to assess (in a before/after or here/elsewhere evaluation process) its impact in terms of: optimizing the participation rate; the time it takes to complete the test; mobilizing people usually NPAI; and colonoscopy completion rates after a positive test.

Acknowledgements

The cost of the 2-month campaign was fully financed by MAST Group (France). The authors are grateful to all partners (CRCDC-IDF, Site de Seine-Saint-Denis; Association de Dépistage du cancer colorectal (ADCCR) en Seine-Saint-Denis; Service de Gastroenterologie de l'Hôpital Avicenne, APHP; Collège des gastroentérologues de Seine-Saint-Denis; Medorion® team). We thank very especially Guillaume LE MAG and Aude TRIEUL (daily monitoring of the CRCSP database).

Conflicts of interest

No conflict of interest (KA, LS[3], GT, SJM and BR).

Conflict of interest: LS[2] and FM declare a conflict of interest because they are employees of the start-up Medorion®.

Accepted Article online 13 May 2020

References

1 Koivogui A, Ecochard R, Le Mab G, Benamouzig R. Impact of stopping sending colorectal cancer screening tests kits by regular mail. *Public Health* 2019; **12:** 10–2.
2 Koivogui A, Le Mab G, Benamouzig R. Detection of colorectal neoplasia in a cohort before and after the change of fecal occult blood test in a french colorectal cancer screening program. *Am J Gastroenterol* 2018; **113:** 1891–9.
3 Lcuraud K, Jczewski-Scrra D, Viguier J, Salincs E. Colorectal cancer screening by guaiac faecal occult blood test in France: evaluation of the programme two years after launching. *Cancer Epidemiol* 2013; **37:** 959–67.
4 Giorgi Rossi P, Grazzini G, Anti M *et al.* Direct mailing of faecal occult blood tests for colorectal cancer screening: a randomized population study from Central Italy. *J Med Screen* 2011; **18:** 121–7.
5 Camilloni L, Ferroni E, Cendales BJ *et al.* Methods to increase participation in organised screening programs: a systematic review. *BMC Public Health* 2013; **13:** 464.
6 Cole SR, Smith A, Wilson C, Turnbull D, Esterman A, Young GP. An advance notification letter increases participation in colorectal cancer screening. *J Med Screen* 2007; **14:** 73–5.
7 Tinmouth J, Ritvo P, McGregor SE *et al.* A qualitative evaluation of strategies to increase colorectal cancer screening uptake. *Can Fam Physician* 2011; **57:** e7–e15.
8 van Roon AH, Hol L, Wilschut JA *et al.* Advance notification letters increase adherence in colorectal cancer screening: a population-based randomized trial. *Prev Med* 2011; **52:** 448–51.
9 Van Roosbroeck S, Hoeck S, Van Hal G. Population-based screening for colorectal cancer using an immunochemical faecal occult blood test: a comparison of two invitation strategies. *Cancer Epidemiol* 2012; **36:** e317–e324.
10 Barthe J, Perrodeau E, Gilberg S *et al.* Impact of a Doctor's invitation on participation in colorectal cancer screening: a cluster randomized trial. *Am J Med* 2015; **128:**1024.e1–7.
11 Courtier R, Casamitjana M, Macia F *et al.* Participation in a colorectal cancer screening programme: influence of the method of contacting the target population. *Eur J Cancer Prev* 2002; **11:** 209–13.
12 Altobelli E, Lattanzi A, Paduano R, Varassi G, di Orio F. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. *Prev Med* 2014; **62:** 132–41.
13 Muller D, Logan J, Dorr D, Mosen D. The effectiveness of a secure email reminder system for colorectal cancer screening. *AMIA Annu Symp Proc* 2009; **2009:** 457–61.
14 Sequist TD, Zaslavsky AM, Colditz GA, Ayanian JZ. Electronic patient messages to promote colorectal cancer screening: a randomized controlled trial. *Arch Intern Med* 2011; **171:** 636–41.
15 Denis B, Gendre I, Perrin P. [Colorectal cancer screening: a disappointing performance needs improvement]. *Rev Prat* 2018; **68:** 839–46.
16 Harty NM, Le Grice K, Cahill C, Bull S, Dwyer A. End-Cancer: development and pilot testing of multimedia recruitment for a text message campaign to increase cancer screening. *mHealth* 2018; **4:** 50.
17 Whitaker C, Stevelink S, Fear N. The use of Facebook in recruiting participants for health research purposes: a systematic review. *J Med Internet Res* 2017; **19:** e290.

18 Brittain K, Pennings Kamp KJ, Salaysay Z. Colorectal cancer awareness for women via Facebook: a pilot study. *Gastroenterol Nurs* 2018; **41:** 14–8.
19 Pellino G, Simillis C, Qiu S *et al.* Social media and colorectal cancer: a systematic review of available resources. *PLoS ONE* 2017; **12:** e0183031.
20 Mohamad Marzuki MF, Yaacob NA, Bin Yaacob NM, Abu Hassan MR, Ahmad SB. Usable mobile app for community education on colorectal cancer: development process and usability study. *JMIR Hum Factors* 2019; **6:** e12103.
21 Shimoda A, Ichikawa D, Oyama H. Using machine-learning approaches to predict non-participation in a nationwide general health check-up scheme. *Comput Methods Programs Biomed* 2018; **163:** 39–46.
22 Journal Officiel de la République Française (JORF). Texte n° 13: Arrêté du 19 mars 2018 modifiant l'arrêté du 29 septembre 2006 relatif aux programmes de dépistage organisé des cancers et portant modification du cahier des charges du dépistage organisé du cancer colorectal. *JORF* 2018; **0068**: NOR: SSAP1807442A.
23 Journal Officiel de la République Française (JORF). Texte N° 67: Délibération n° 2017-215 du 13 juillet 2017 portant adoption d'une norme destinée à simplifier l'obligation de déclaration des traitements de données à caractère personnel ayant pour finalité le dépistage organisé du cancer du sein, du cancer colorectal et du cancer du col de l'utérus mis en œuvre par les structures de gestion conventionnées, et abrogeant la délibération n° 2015-175 du 11 juin 2015 (décision d'autorisation unique n° AU-043) (NS-059). *JORF* 2017; **0204**: NOR: CNIL1724568X.
24 Debordc T, Chatignoux E, Quintin C, Beltzer N, Hamers FF, Rogel A. Breast cancer screening programme participation and socioeconomic deprivation in France. *Prev Med* 2018; **115:** 53–60.
25 Rogers CR, Robinson CD, Arroyo C, Obidike OJ, Sewali B, Okuyemi KS. Colorectal cancer screening uptake's association with psychosocial and sociodemographic factors among homeless blacks and whites. *Health Educ Behav* 2017; **44:** 928–36.
26 ARS. Décision N°6-2016 portant approbation de mise en place d'une stratégie complémentaire de remise de kits de dépistage du cancer colorectal. Paris: Agence Regionale de Santé Ile-De-France (ARS-IDF), 2016.
27 Witte K. The role of threat and efficacy in AIDS prevention. *Int Q Community Health Educ* 1991; **12:** 225–49.
28 Smith SW, Nazione S, Laplante C *et al.* Topics and sources of memorable breast cancer messages and their impact on prevention and detection behaviors. *J Health Commun* 2009; **14:** 293–307.
29 Cutrona SL, Wagner J, Roblin DW *et al.* E-mail to promote colorectal cancer screening within social networks: acceptability and content. *J Health Commun* 2015; **20:** 589–98.
30 Tifratene K, Eisinger F, Rinaldi Y, Didelot R, Seitz JF. Colorectal cancer screening program: cost effectiveness of systematic recall letters. *Gastroenterol Clin Biol* 2007; **31:** 929–33.

5.2. – Synthèses de l'axe 3 et perspectives

Cette étude est la première en France, à avoir expérimenté dans le DOCCR, l'usage d'un algorithme d'intelligence artificielle pour la promotion de la santé à travers le réseau social Facebook. En démontrant qu'il est possible de rehausser la participation aux campagnes (réalisation d'un test) l'étude a aussi montré qu'il est possible d'établir le profil d'éligibilité des personnes à partir d'un contact établi sur un réseau social (optimisation de l'exclusion médicale). L'étude a ainsi ouvert la voie sur l'expérimentation de nouvelles méthodes d'optimisation du DOCCR. En plus de la précision dans le ciblage, l'usage d'un module d'intelligence artificielle permet de faire des projections sur les comportements des personnes ayant demandé un kit de dépistage (***Annexe-1***).

A la date du 1er mars 2022, soit environ 16 mois après la publication de l'étude résumant cet axe 3, la CNAM a mis à la disposition de la population cible, un site internet dédié à la commande des kits de test. A travers ce site de commande en ligne, le profil d'éligibilité des personnes est automatiquement transmis aux SGDO, tel qu'expérimenté dans cet axe. Toutefois, le site de commande mis à la place par la CNAM ne répond que partiellement à la demande de dématérialisation de la procédure d'invitation et de suivi des résultats du dépistage exprimée dans cette étude. En effet, le numéro d'invitation de la SGDO est requis pour passer une commande de kit sur le site de la CNAM. Le site de la CNAM facilite ainsi l'accès au dépistage seulement à ceux qui sont informés de l'existence du programme par le biais des courriers d'invitations reçus des SGDO. Pour les autres personnes qui ne reçoivent aucun courrier, notamment les personnes NPAI, la dématérialisation tel qu'argumentée dans cette étude, permettrait d'établir leur éligibilité mais aussi de les mobiliser et fidéliser aux campagnes de dépistage.

6. DISCUSSION GENERALE

6.1. Principaux résultats

Pour le choix de la population éligible aux campagnes, l'exclusion médicale est couramment pratiquée dans les programmes DOCCR à travers le monde, notamment dans l'espace UE où le DOCCR est complètement déployé. Bien qu'il existe une disparité en termes de type de critère et de nombre de critères, il se dégage 5 critères consensuels (antécédents personnels de cancer colorectal, de maladie inflammatoire chronique, d'adénome, de haut risque génétique et de coloscopie ou sigmoïdoscopie récente) moins subtiles qui définissent parfaitement la personne à exclure des campagnes de dépistage.

Il faut cependant noter que dans la plupart des programmes, il existait une différence importante entre les recommandations de leurs cahiers des charges et ce qui est réellement appliqué. La pratique de l'exclusion est disparate et varie en fonction des moyens d'exclusion à la disposition des programmes, ce qui fait qu'aucun programme européen ou américain n'observe la totalité des facteurs morbides justifiant d'une exclusion médicale. Ceci sous-entend que la notion de risque moyen est partiellement prise en compte dans les programmes avec pour conséquences majeures :

(1)La sous-estimation des taux d'exclusion dans les programmes, ce qui induit un énorme biais d'information dans l'estimation du taux de participation ;

(2) L'impossibilité de comparer les taux entre programmes, notamment dans l'espace UE. Pour espérer arriver un jour à comparer, nos analyses sur la nécessité de standardiser (dans l'espace UE) les définitions des indicateurs mesurés, convergent avec d'autres analyses antérieures [16, 17, 160]. Une de nos recommandations a d'ailleurs été de n'appliquer que les critères consensuels révélés dans l'axe 1 de cette thèse et surtout de faciliter la connexion des bases du dépistage avec les autres bases de données de santé, notamment les BDMA. Toutefois, pour un bon usage des BDMA, une standardisation des algorithmes d'extraction est nécessaire.

Si la France partage avec ses voisins de l'UE plusieurs difficultés susmentionnées, cette thèse met en évidence d'autres difficultés inhérentes au DOCCR français. En effet, le cahier des charges du DOCCR n'ayant pas été explicite sur les moyens à mettre à la disposition des SGDO, celles-ci utilisent divers moyens et stratégies pour optimiser les campagnes de DOCCR. Il existe une disparité assez évidente entre départements (ou régions) en termes d'accès aux moyens d'optimisation du DOCCR. A l'inverse des pays voisins qui ont des registres

nationaux de cancers, la France a des région entière (IDF en exemple) non couverte par un registre de cancers[43, 46, 49, 62, 64-77]. De même, à l'inverse de quelques pays voisins qui utilisent les BDMA pour affiner la population cible du DOCCR, les SGDO en France n'ont à ce jour aucun accès au SNDS. En effet, à l'exception des données de coloscopies transmises par quelques CPAM, le SNDS n'était pas utilisé par les SGDO qui ont par ailleurs exprimé lors de l'enquête transversale (axe 2), leur intérêt de l'avoir comme moyen d'exclusion médicale.

Cette thèse avait été initiée dans le but de faciliter cet accès du SNDS par les SGDO. A l'aide d'une revue de la littérature nationale et internationale, elle a permis l'élaboration d'une définition opérationnelle du phénotype de la personne à exclure des campagnes de DOCCR à partir des données du SNDS. Cette définition a été faite par la synthèse des codes définissant chaque situation morbide qui sera transcrite conformément aux vocabulaire et terminologie en vigueur dans le SNDS (CCAM, CIM-10, ATC, GHM...) en vue d'un test d'efficacité et de validation sur les bases de données du SNDS.

Comme moyen supplémentaire, cette thèse a argumenté une nouvelle stratégie utilisant un module d'intelligence artificielle pour optimiser une campagne de DOCCR. Cette stratégie qui consiste en l'envoi de message de promotion du DOCCR par le biais des réseaux sociaux, notamment Facebook a déjà fait ses preuves dans d'autres programmes de promotion de la santé [161, 162].

6.2. Apport des sources de données pour une meilleure définition de la population cible du dépistage

En France, la sous-estimation du taux d'exclusion médicale a déjà été évaluée. Seitz et collaborateur avaient démontré que le taux d'exclusion de 13% observé en 2016 et 2017 était sous-estimé car dans les Bouches du Rhône, environ 20% de la population cible avaient une coloscopie récente [163]. Ceci nous laisse espérer que l'usage du SNDS permettrait d'attendre un taux minimum d'exclusion de 25%. Ce taux minimum escompté est largement justifié par les 25% observé dans le département Indre-et-Loire où la SGDO dans le cadre d'un partenariat avec la CPAM, avait régulièrement accès aux données de coloscopies entre 2016 et 2020. Ce taux d'exclusion de 25% est aussi celui observé en Flandre (Belgique) où l'usage des BDMA et des registres de cancer est en vigueur (étude 1 de l'axe 1).

En outre, l'usage du SNDS permettra d'avoir une méthodologie d'exclusion reproductible et consensuelle en cette période de discussion sur la mise en place d'un espace européen de donnée de santé (EEDS). S'il est validé, l'algorithme d'exclusion médicale développé dans cette thèse pourrait être standardisé en perspective d'une

éventuelle intégration des données des DOCCR dans ce futur EEDS. En effet, les MADB révélées dans l'axe 1 de cette thèse utilisent des terminologies diverses et non interopérables pour certaines, la définition des normes et la standardisation d'un algorithme sont des préalables pour la migration des données de dépistage dans l'EEDS. L'usage d'algorithme d'IA permettrait d'optimiser la participation aux campagnes de DOCCR à hauteur de 5.1 à 11%. Elle pourrait permettre un suivi à moindre coût (courriel et SMS), des kits de test envoyés aux personnes. De même pour les personnes ayant un test positif, le suivi jusqu'à la réalisation de la coloscopie pourrait être envisagé à moindre coût (SMS ou courriel).

7. CONCLUSION

Cette thèse met en évidence des outils et moyens d'exploitation de plusieurs sources de données pour une optimisation des campagnes de DOCCR en France. La thèse montre aussi que ces sources de données n'étaient pas encore exploitées par les SGDO en France.

Ce travail montre qu'en plus d'être largement utilisé dans le cadre d'études médico-économiques et épidémiologiques, le SNDS contiendrait la quasi-totalité des données indispensables à l'estimation du taux d'exclusion médicale lors des campagnes de DOCCR. En répertoriant dans un algorithme d'exclusion médicale, l'ensemble des codes définissant dans le SNDS, chacune des situations morbides pouvant justifier d'une exclusion médicale du DOCCR, cette thèse argumente la faisabilité de l'exclusion médicale des personnes non éligibles au dépistage à l'aide de la base de données SNDS. La thèse démontre que l'usage du SNDS comme moyen permettrait d'attendre un taux minimum d'exclusion de 25% dans le programme français.

A propos de la faisabilité, la thèse a argumenté que la plupart des facteurs morbides cibles de l'algorithme d'exclusion médicale sont requêtables dans la plupart des bases de données médico-administratives référencées dans les pays ayant un DOCCR bien que l'usage de ces types de bases de données de santé ne soit pas en vigueur dans tous les programmes de DOCCR. Cependant, pour son usage, l'algorithme d'exclusion médicale proposé dans cette thèse devra inéluctablement passer par une évaluation en deux phase : une première phase de test de son efficacité en termes d'optimisation des taux d'exclusion et de participation aux campagnes de DOCCR et une seconde phase de chainage en vue d'une évaluation des données d'invitation extraites du SNDS. Eventuellement, d'autres nouvelles stratégies devront être mises en place pour corriger les imprécisions du SNDS dans le classement des personnes en éligibles ou Non-éligibles aux campagnes de DOCCR.

En expérimentant dans le DOCCR, l'usage d'un algorithme d'intelligence artificielle, cette thèse a démontré que les RS sont de sources supplémentaires exploitables pour l'optimisation des campagnes de DOCCR. La mise en œuvre de cette nouvelle stratégie permettra une amélioration du taux de participation par la mobilisation des personnes relativement jeunes et sans antécédent de participation. Toutefois, pour pérenniser une telle stratégie complémentaire dans le DOCCR, il semble évident d'établir son impact dans un schéma d'étude épidémiologique à grande échelle.

.

8. PERSPECTIVES DE LA THESE

8.1 – Validation de l'algorithme d'extraction des données du SNDS

Pour consolider la définition de la population éligible au DOCCR, l'intérêt du développement d'une méthode d'identification à partir des données de consommation de soins du SNDS a été largement argumenté dans cette thèse. En effet, une meilleure connaissance de la couverture endoscopique et des facteurs de risque menant à une exclusion du DOCCR permettrait d'optimiser le calcul du taux d'exclusion médicale à une campagne biannuelle de DOCCR. La validation de l'algorithme d'exclusion médicale passera par deux niveaux : (1) Dans un premier temps, l'algorithme d'exclusion médicale élaboré dans cette thèse sera testé sur les bases du SNDS ; (2) Puis, une étude de validation de l'algorithme d'exclusion médicale sera faite à l'aide d'un chainage du SNDS et de la base de données du dépistage.

8.1.1 – TEST DE L'EFFICACITE DE L'ALOGRITHME D'EXCLUSION MEDICALE EN TERMES D'EXTRACTION DE DONNEES DES BASES DU SNDS

Dans le cadre de cette thèse, le protocole de ce test d'efficacité de l'algorithme a été écrit conformément aux recommandations de l'INSERM et son exécution est en cours (***Annexe 2***). L'algorithme sera testé en termes d'efficacité à identifier les caractéristiques des personnes participantes ou non participantes aux campagnes de dépistage à partir des données du SNDS, en utilisant la variable de remboursement des kits de test par l'assurance maladie. Deux objectifs sont visés dans ce projet de test :

8.1.1.A. – PHENOTYPER LA PERSONNE CIBLE A EXCLURE OU PAS DES CAMPAGNES DE DOCCR A L'AIDE DES DONNEES DU SYSTEME NATIONAL DES DONNEES DE SANTE (SNDS)

Il s'agira d'une approche descriptive qui ciblera une cohorte de 3.258.975 personnes âgées de 50 à 74 ans, domiciliées en Ile-de-France et cibles de la campagne biannuelle (DOCCR) de 2021-2022 (Cohorte_21_22). Dans un premier temps, il sera fait une description du nombre de personnes âgées de 50 à 74 ans domiciliées en Ile-de-France et ayant au moins un enregistrement dans une base du SNDS, en rapport avec une situation morbide justifiant d'une exclusion médicale du DOCCR. Puis, avec l'hypothèse que le taux d'exclusion médicale est optimal s'il est estimé à partir des données du SNDS, l'étude descriptive comparera le taux d'exclusion médicale publié par SPF à celui estimé à partir des données du SNDS dans cette cohorte de 3,5 millions. Dans cette Cohorte_21_22, SPF avait estimé le taux d'exclusion à 10,9% et le taux de participation à la

campagne à 31,3%. Cette évaluation comparative sera faite avec l'hypothèse de non-infériorité du taux estimé à partir du nombre de cas d'exclusion médicale extraits du SNDS par l'algorithme. En plus d'une déclinaison géographique par département/région ou par zone (ultra-rural, ultra-urbain), la comparaison tiendra compte des stratégies actuelles mises en place par les structures de gestion pour optimiser leur collecte des données d'exclusion médicale (*Réf. Les profils des structures de gestion selon les stratégies sont ceux décrits dans l'axe 2 de la thèse : Enquête d'exclusion médicale*).

8.1.1.B. – ANALYSER L'APPORT DES DONNEES DU SNDS POUR UNE ESTIMATION DU TAUX DE PARTICIPATION A UNE CAMPAGNE DE DOCCR EN FRANCE

Il s'agira d'une étude cas-témoins visant à analyser les facteurs de risque de non-réalisation d'un test de dépistage lors de de la campagne biannuelle (DOCCR) de 2020-2021. L'étude ciblera la cohorte de 3.258.975 personnes âgées de 50 à 74 ans, domiciliées en Ile-de-France et cibles de la campagne biannuelle (DOCCR) de 2021-2022. Les Cas seront toutes les personnes de la cohorte n'ayant eu aucun remboursement de test de dépistage récent (2021-2022) dans les bases du SNDS. Les Témoins seront toutes les personnes de la cohorte ayant un remboursement de test de dépistage sur la période 2021-2022 dans les bases du SNDS. Les fréquences de situations morbides (maladies chroniques, exploration diagnostique, traitement médicamenteux au long cours) antérieures à 2021 seront comparées entre Cas et Témoins, avec l'hypothèse que les cas n'ont pas réalisé de test sur la campagne biannuelle 2021-2022 à cause d'une situation morbide enregistrée en amont de cette campagne biannuelle.

8.1.2. – PROJET DE CHAINAGE DES DONNEES DU PROGRAMME DE DEPISTAGE DU CANCER COLORECTAL ET DES DONNEES DU SYSTEME NATIONAL DES DONNEES DE SANTE (SNDS).

Le chainage de la base de données du dépistage et celle du Système National des Données de Santé (SNDS) permettrait d'améliorer le programme de dépistage. Dans le cadre de cette thèse, le protocole de chainage a été écrit conformément aux recommandations du Health Data Hub (HDH). Deux objectifs majeurs sont visés dans ce projet de chainage.

Le premier objectif visé sera la validation de l'algorithme d'exclusion médicale testé. En effet, dans l'axe 2 de cette thèse, il a été argumenté que l'usage du SNDS génèrerait aussi des FP et FN à cause de la sensibilité du SNDS <100% et la nécessité d'une évaluation du SNDS avec pour référentiel la base de données du dépistage. En fonction de la précision des données extraites du SNDS, des recommandations pourraient être faite pour la

mise en place d'autres nouvelles stratégies visant à minimiser les imprécisions des exclusions à faire à partir du SNDS.

Le second objectif visé sera l'analyse du risque de cancer colorectal (ainsi que le risque de gravité des lésions cancéreuses) en fonction des facteurs prédictifs identifiés, notamment ceux en rapport avec les habitudes de consommation médicamenteuse ou la prise en charge d'autres affections chroniques (consommation régulière d'anti-inflammatoire, traitement d'une obésité morbide…). L'hypothèse est que les habitudes de consommation médicamenteuse en rapport avec certaines affections chroniques sont prédictives du risque de CCR. En effet, différents facteurs de risque du CCR ont été décrits. Il a été montré que la consommation régulière d'aspirine était associée à une diminution du risque de CCR de l'ordre de 30% [164]. Des études suspectent un effet protecteur dans l'analyse du lien entre un traitement hormonal substitutif de la ménopause et la survenue d'un CCR[165]. L'impact de la consommation de certains médicaments sur le taux d'Hémoglobine fécale (mesurée lors du test de DOCCR) reste non évalué à ce jour dans une large population. Les résultats quantifiés de l'hémoglobine fécale par le FIT ne sont d'ailleurs pas pris en compte et seul le caractère positif (au-delà du seuil de positivité fixé à 30µgHb/g de selle) est pris en compte dans le DOCCR alors qu'il existe dans la base des SGDO les valeurs de mesure de l'hémoglobine fécale. La probabilité qu'un test « négatif » avec un taux d'hémoglobine fécale « élevé » puisse être un élément prédictif de futur néoplasie est de ce fait non nulle. De même, le lien entre la consommation régulière de certains médicaments et la proportion de « faux positifs» pourrait être établie. Cette analyse étiologique des faux positifs trouverait sa principale justification dans la forte proportion (34.3% entre 2015 et 2020) de coloscopies normales après un FIT positif dans la région IDF.

Cette nouvelle approche pourrait permettre d'identifier des personnes ayant un risque plus élevé d'avoir un test positif deux ans plus tard [166]. A terme, elle pourrait émettre des hypothèses sur la stratégie de dépistage en fonction du niveau antérieur du taux d'hémoglobine dans les selles et en fonction de tel ou tel traitement médicamenteux ou en fonction de telle ou telle comorbidité. Ceci pourrait enfin alimenter la réflexion sur l'adaptation du calendrier des invitations. Tel que recommandé par certains auteurs[167], la mise en œuvre d'un scores de risque de CCR tenant compte des facteurs de risque identifiés, pourrait permettre d'améliorer les pratiques de prévention par l'adaptation des modalités du dépistage selon le niveau de risque.

8.2 Projet de mise en œuvre d'une stratégie complémentaire basée sur la géolocalisation et la motivation des personnes par le biais de leur connexion sur les réseaux sociaux

La principale recommandation faite au terme de l'étude 6 (axe 3), était la mise en place et l'évaluation à grande échelle d'une stratégie complémentaire d'optimisation des campagnes de DOCCR à travers les réseaux sociaux notamment Facebook. Ce projet de recherche et d'action mettra à disposition après validation, une stratégie complémentaire de mobilisation des personnes cibles, habituellement non participantes aux campagnes de DOCCR. La validation à travers une expérimentation sur la population de personnes domiciliées dans la région Ile-de-France nécessite une étude de suivi et d'évaluation en guise de base scientifique. Le protocole a été finalisé et budgétisé, en attente d'un accord de principe pour financement.

ANNEXES

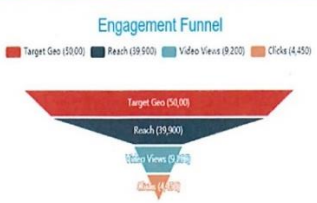
Liste des annexes

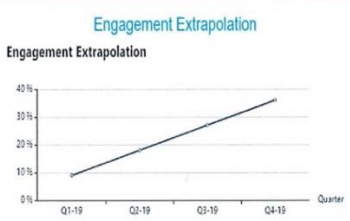
ANNEXE 1 : ANALYSE DU FLUX ET PROJECTIONS FAITES PAR LE MODULE

ANNEXE 2 : PROJET DE PHENOTYPAGE DE LA PERSONNE CIBLE A EXCLURE OU PAS DES CAMPAGNES DE DEPISTAGE ORGANISE DU CANCER COLORECTAL A L'AIDE DES DONNEES DU SYSTEME NATIONAL DES DONNEES DE SANTE (SNDS).

Annexe 1 : Analyse du flux et projections faites par le module.

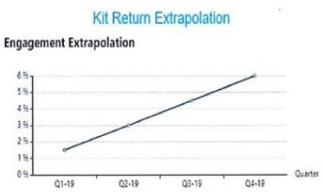
Funnel analytics





The engagement funnel demonstrates the number of people reached by the campaign and their online engagement. In only 2 months, 18% of the targeted population watched a video and 9% clicked the link. Extrapolating this data to a year long campaign, we estimate a click through engagement rate of 36%.





The kit return funnel demonstrates the number of people who were actively engaged in the "real world", beyond the online engagement boundaries. Out of 298 people who signed up and requested a FIT kit, 160 people (54%) were eligible for it. Out of the 160 FIT kits that were sent, 64 (40%) were returned for testing.

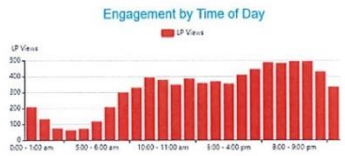


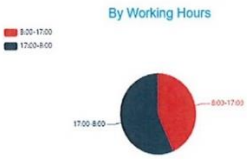
4450 people engaged with the campaign and clicked on the link. Out of them, 6.7% requested a kit, but only 3.6% of them were eligible to get a FIT kit. The return rate compared to the clicks rate was 1.5%. Extrapolating the 2 months results of the return rate compared the clicking rate (1.5%), to a year-long campaign, yields a FIT kit return rate expectancy of 6%.

While previous efforts reached a maximum of 25% FIT kit return rate from the number of kits sent out, this campaign achieved a 40% return rate, which shows 60% higher engagement rates.

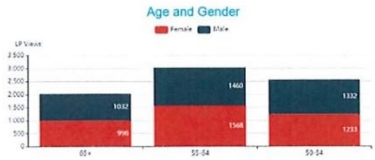
Demographic analytics

Most people engaged with the campaign outside of working hours. The peak time was between 18:00-22:00.

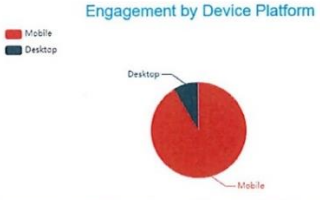




The campaign engaged both genders, almost equally, across all relevant age groups. Using Facebook, we were able to reach people of all relevant ages, including the elderly population, aged over 65 years old.



Most people were engaged with the campaigns via mobile compared to desktop.



Annexe 2 : Projet de phénotypage de la personne cible à exclure ou pas des campagnes de dépistage organisé du cancer colorectal à l'aide des données du Système National des Données de Santé (SNDS).

Inserm
La science pour la santé
From science to health

[Phénotypage de la personne cible à exclure]

Ce protocole d'étude sur les données du SNDS est à destination des agents officiellement rattachés à une équipe labellisée Inserm et souhaitant bénéficier de l'accès permanent de l'Inserm aux données du SNDS pour un projet d'étude sous responsabilité de traitement de l'Inserm.

****

Seul ce document doit être utilisé
par le porteur de projet pour établir son protocole d'étudeLa structure et le format doivent être conservés.

****

Le présent document doit être rédigé en Français

Inserm
La science pour la santé
From science to health

1. INTRODUCTION

1.1. Identification du projet et de son porteur

Phénotypage de la personne cible à exclure ou pas des campagnes de dépistage organisé du cancer colorectal à l'aide des données du Système National des Données de Santé (SNDS).

Intitulé du projet	Phénotypage de la personne cible à exclure ou pas des campagnes de dépistage organisé du cancer colorectal à l'aide des données du Système National des Données de	
Acronyme du projet	Etude Pheno-DOCCR	
Calendrier prévisionnel du projet (et non des droits d'accès demandés)	Date de début : 01/10/2021	
	Date de fin : 31/12/2023	
Porteur de projet / Demandeur des droits d'accès EGB/SNDS	Nom : KOIVOGUI	Nom de l'unité Inserm : LIMICS
	Prénom : AKOI	Numéro de l'unité : 1142
	E-mail Inserm : akoi.koivogui@inserm.fr	Nom de l'équipe : LIMICS
	Tél : 01558xxxxxx	Nom du directeur JAULENT Prénom du directeur MARIE-CHRISTINE

2. OBJECTIFS ET FINALITES

2.1 Contexte et justification de l'étude

Toutes les conditions justifiant la mise en place d'une politique de dépistage sont remplies par le cancer colorectal (CCR) [1-3]. En France, une évaluation récente du dépistage organisé du CCR (DOCCR) montrait que la réalisation d'un test fécal immunologique (FIT) tous les deux ans dans une population à risque moyen était la stratégie la plus rentable parmi les stratégies évaluées[4]. Le test utilisé actuellement a fait la preuve de son efficience en termes de diminution de la mortalité par cancer colorectal [2, 3, 5-8]. Malgré l'intérêt démontré du programme de dépistage, le taux de participation aux campagnes biannuelles dépasse rarement 50% dans la plupart des départements français [9-11]. Plusieurs raisons ont déjà été évoquées pour expliquer la faible participation à une campagne de dépistage [10, 12-15]. Dans le contexte français, il semble pertinent d'investiguer la fiabilité des données d'invitation fournies par la CNAM et la méthode d'estimation du taux de participation.

A ce jour, la plupart des politiques de DOCCR mises en place ciblent une population à risque moyen de CCR, définie sur le seul critère âge [6]. L'efficacité du dépistage dépendant de son application dans un cadre approprié et à une fréquence appropriée[16]. La plupart des programmes européens incluent les personnes de tout sexe et âgées de 50 à 74 ans et observent un intervalle de 1 ou 2 ans entre deux dépistages [17]. En France, en prélude de chaque campagne, une mise à jour des fichiers de personnes éligibles est faite après la transmission des données individuelles (Nom et Prénom, date de naissance, sexe, Numéro de sécurité sociale, adresse de résidence) par les différents Régimes d'Assurance Maladie (RAM). Les personnes sont ainsi invitées à leur 50ème anniversaire (primo-invitation) et tous les deux ans après un test de dépistage (Invitation-subséquente) jusqu'à l'âge de 74 ans révolus [18].

Dans son protocole d'évaluation du DOCCR, Santé Publique France (SPF) définit le taux de participation par le rapport entre le nombre de personnes dépistées et la population éligible, soit la population cible du programme (hommes et femmes de 50 à 74 ans) dont sont extraites les personnes exclues du programme pour raisons médicales[19]. La définition de cette population éligible est basée sur la notion de risque moyen caractérisé par l'âge (50-74 ans) et l'absence de risque personnel ou familial connu [20]. Ainsi des personnes à haut risque (antécédents personnels ou familiaux d'adénomes colorectaux, de cancers ou de maladies inflammatoires : colite ulcéreuse et maladie de Crohn) ou à très haut risque (polypose adénomateuse familiale ou syndrome de Lynch) de développer un cancer colorectal et les personnes atteintes de pathologies extra-intestinales sévères sont exclues de cette population éligible. De même, les personnes ayant réalisé une coloscopie dans les 5 ans sont temporairement exclues si le résultat de cette coloscopie est normal [9]. Lors de la première évaluation du programme, ce taux d'exclusion était estimé à 10.5% dans 46 départements de France [9]. En IDF, le taux d'exclusion standardisé était de 9,9% dans la cohorte de personnes invitées à la campagne biannuelle de 2015-2016 comparativement au taux national qui était de 12,3% [21].
Une enquête réalisée en Seine-Saint-Denis avait montré que cette approche d'estimation du taux d'exclusion était sous-optimale et que de nombreux sujets étaient considérés comme à risque moyen alors qu'ils étaient à risque élevé et que d'autres étaient considérés comme non participants alors qu'ils étaient en fait « couverts » par une coloscopie récente non déclarée. Dans ce département de Seine-Saint-Denis, le taux de participation avait été estimé à 22,6% dans la cohorte des 386541 personnes âgées de 50-74 ans, invitées à la campagne biannuelle de 2018-2019. Dans cette cohorte, la proportion des exclusions n'était que de 10,4%. Ces chiffres pourraient être sous-estimés car en 2018 dans le département, le taux standardisé des personnes prises en charge pour maladies inflammatoires chroniques intestinales était de 4‰ ; le taux standardisé des personnes prises en charge pour un cancer colorectal actif était de 2,2‰ et pour un cancer colorectal sous surveillance de 3,2‰ [22]. En rappel, le nombre de coloscopies réalisées en France, en 2016, était estimé à 1.402.047 dont 76,9% chez les personnes âgées de plus de 50 ans [23].
Le fait que le taux d'exclusion soit estimé seulement après un recueil de déclarations volontaires des sujets (questionnaire joint au courrier d'invitation) et des médecins traitants pourrait compromettre un recueil exhaustif de ces cas potentiels cas d'exclusion médicale. A propos de la méthode de recueil, SPF admettait que le nombre de personnes exclues était fourni par les structures de gestion à partir de diverses sources de données exploitables (réponses des personnes invitées, fichiers des actes de coloscopies remboursés par les CPAM, fichiers des actes diagnostiques et thérapeutiques…), mais dont l'accès variait selon les structures de gestion [19].

2.2 Objectif(s) de l'étude

Objectif général :

Ce travail vise à définir le phénotype d'une personne à exclure des campagnes de DOCCR en France et à tester son application dans le SNDS. Le phénotype sera construit en identifiant les critères décisionnels définissant les personnes à risque moyen et élevé de CCR et en les transformant pour qu'ils soient transposables dans une requête sur des données de consommation de soins (SNDS). En effet, une meilleure connaissance de la couverture endoscopique et des facteurs de risque menant à une exclusion du DOCCR permettrait d'optimiser le calcul du taux d'exclusion médicale à une campagne biannuelle de DOCCR.

Objectifs secondaires :

- Une revue non systématique de la littérature incluant la liste exhaustive des travaux réalisées sur les bases du SNDS a permis d'élaborer un algorithme composé de l'ensemble des codes définissant une situation morbide pouvant justifier une exclusion des campagnes de dépistages. Cet objectif spécifique vise à évaluer la performance de l' algorithme par le décompte du nombre de potentiels cas d'exclusion médicale qu'il peut extraire des bases du SNDS;
- Décrire le profil clinique et épidémiologique des personnes ayant réalisé un test de dépistage comparativement à celui des personnes n'ayant pas réalisé un test lors de la même campagne de dépistage ;

2.3 Justification de l'intérêt public

Cette étude a vocation à améliorer non seulement, la qualité du programme de DOCCR mais aussi les connaissances scientifiques dans le domaine de l'exploitation des bases du SNDS, ce qui justifie son intérêt public. En termes de qualité, l'optimisation du taux d'exclusion attendue pourrait réduire significativement le coût actuel du programme. A terme, le projet pourrait aider les structures de gestion du dépistage à conduire leurs missions de gestion du système d'information et d'évaluation des campagnes. En effet, l'arrêté du 4 avril 2018 relatif aux programmes de dépistages des cancers qui attribue au centre régional de coordination des dépistages des cancers une mission de gestion du système d'information du dépistage (intégration et mise à jour des fichiers, conservation des données, interfaces avec les autres systèmes d'informations, enregistrement des refus, respect des obligations liées à la loi n° 78-17 relative à l'information, aux fichiers et aux libertés du 6 janvier 1978), une mission d'évaluation des programmes de dépistages organisés, une mission de valorisation des données issues des programmes et une mission de vigilance sanitaire du dispositif [24].

2.4 Publication des résultats et valorisation

Les résultats de ce projet seront diffusés à travers :

- Des journées de présentation des résultats aux structures ressources et aux professionnels de santé ;
  - Rapport public à destination de l'INSERM
  - Rapport public à destination des structures partenaires (Inca, Hôpital Avicenne, LIMICS)
- Des publications et communications scientifiques : l'équipe projet envisage une valorisation du projet au moyen de publications scientifiques qui pourraient présenter, entre autres :
  - Apport des données du SNDS pour une estimation du taux de participation à une campagne de dépistage organisé du cancer colorectal en fonction du type de population cible
  - Exploration de nouvelles méthodes d'exclusion médicale aux campagnes de DOCCR en France : une comparaison des données d'exclusion publiées par SPF et des données du SNDS.
  - Faisabilité de la validation du phénotypage de la personne à exclure des campagnes de DOCCR à partir des données des dossiers du SNDS.
  - Estimation du nombre de personnes qui pourraient être exclues à tort du DOCCR à partir du SNDS ;
  - Profil clinique et épidémiologique des personnes ne réalisant jamais un test de dépistage du cancer colorectal en France
- Toute publication ou communication sur ce projet, sera disponible en accès libre HAL

3. METHODOLOGIE

3.1 Données requises

3.1.1. Description de la cohorte étudiée

Définition de la population d'étude :

Le projet cible la cohorte de 3 258 975 personnes âgées de 50 à 74 ans, domiciliées en IDF et cibles de la campagne biannuelle de 2021-2022.

Critères d'inclusion :

Personnes âgées de 50 à 74 ans domiciliées en Ile-de-France et ayant au moins un enregistrement dans une base du SNDS

Critères d'exclusion :

Aucun

3.1.2. Sources de données et ciblage des données pertinentes

Définition et mesure des événements étudiés :

Pour cette étude, l'algorithme collectera les codes des diagnostics ou actes de soins en rapports avec :

- ***Le diagnostic et l'année de prise en charge (ou de déclaration en ALD) d'une affection chronique sur la période antérieure à 2020:*** La recherche de diagnostic d'affection chronique se fera par le repérage (Oui/Non) des codes CIM-10 dans les bases du PMSI (DP, DA ou DR) des ALD et du CEPIDC. Si le diagnostic est présent dans le PMSI, l'année de prise en charge sera la date de sortie du dernier RSA. L'année de prise en charge sera l'année de la dernière déclaration ALD ou de décès(CEPIDC) si le diagnostic n'est présent que dans la base des ALD ou du CEPIDC. Les affections chroniques recherchées seront : *i)Le cancer colorectal actif ou sous surveillance ; ii)Les maladies inflammatoires chroniques intestinales ; iii)Les 3 cancers les plus morbides (Prostate, Poumon, Sein) actifs ou sous surveillance, iv) les autres cancers, v)L'obésité morbide, vi)Le diabète.*
- ***La réalisation et l'année de réalisation d'une coloscopie ou d'un coloscanner, ou d'une polypectomie sur la période antérieure à 2020:*** La recherche se fera dans les bases du SNIIRAM et du PMSI par le repérage(Oui/Non) d'un code CCAM en rapport avec une coloscopie ou une polypectomie. Sur la période précédant la campagne ciblée par l'étude (2010-2017), l'année de réalisation sera la date de réalisation de l'acte le plus récent. Sur les périodes per et post campagne (2018-2020), l'année de réalisation sera la date de réalisation de l'acte le plus ancien. Dans la base du PMSI, il s'agira de la date de sortie du dernier RSA contenant le code CCAM de l'acte.

L'existence, la fréquence et l'année de prescription d'un traitement à long terme, polypectomie sur la période antérieure à 2020 : La recherche se fera dans la base des remboursements, par le repérage (Oui/Non) d'un code ATC (année et nombre de délivrances) en rapport avec : *i)Une prise en charge du diabète, ii)Un traitement à long terme par aspirine et autres antiagrégants plaquettaires ; iii) Une hormonothérapie substitutive de la ménopause.*

Variables utilisées :

Pour cette analyse, nous utiliserons :

1 - Les variables sociodémographiques :

- La variable définissant l'âge en années révolues à défaut les variables année et mois de naissance ;
- La variable définissant le genre ;
- La variable définissant le code de département de résidence ;

2 - Les variables pouvant définir l'existence d'une maladie chronique ou d'une exploration colorectale :

- La variable cause médicale de décès ;
- La variable année de décès ;
- Toute variable pouvant contenir un code diagnostic (CIM-1O) ou traitement (CCAM, ATC) d'une maladie chronique ;
- Toute variable pouvant contenir un code de procédure d'exploration colorectal (CCAM).

3 - Les variables permettant d'approfondir l'analyse étiologique sur les facteurs d'adhésion ou de rejet du dépistage:

- Toute variable pouvant contenir un code médicament ;
- Toute variable de prestation (médicaments, consultation…) ;
- Tout variable caractérisant la spécialité des professionnels de santé;
- La variable définissant la réalisation d'un test de dépistage du cancer colorectal ;
- La variable définissant la réalisation d'un test de dépistage du cancer du sein pour les femmes
- Toute variable sur le type de remboursement (70%, 30%) du test de dépistage
- La variable mois et année COVID (si existent)

Si demande d'accès aux données exhaustives du SNDS, justifier en quoi l'EGB n'est pas en mesure de répondre aux besoins de l'étude :

Il s'agit d'une étude visant à décrire un effectif, le plus maximal possible. L'EGB ne nous permettra pas d'avoir cet effectif maximal du nombre de personnes ayant une situation morbide justifiant d'une exclusion médicale du dépistage

Si demande d'accès EGB/SNDS avec croisement de variables sensibles (identifiants potentiels), justifiez ce besoin :

[Les « identifiants potentiels » : Date de naissance (mois et année) ; Dates de soins (jour, mois et année) ; Code de la commune de résidence ; Date de décès (jour, mois, année) ; Code de la commune de décès]

Nous souhaitons avoir la variable année (pas le mois) de décès et la variable cause de décès pour l'analyse la plus exhaustive possible (dans le SNDS) des causes de non-participation aux campagnes de dépistage et éventuellement évaluer dans le cadre de l'objectif spécifique 2, le nombre de tests de dépistage réalisés par une personne qui décède d'un cancer colorectal.
Pour les dates, nous souhaitons n'avoir que les mois et années, ce qui nous permettra d'établir une séquence temporelle entre le test de dépistage, le diagnostic ou le soin d'un cancer.

3.2 Méthodes, traitements et analyses des données

3.2.1 Type d'étude

Avec l'hypothèse que le taux d'exclusion médicale est optimal s'il est estimé à partir des données du SNDS, le premier objectif spécifique sera exploré dans une étude descriptive et évaluative. L'étude descriptive comparera le taux d'exclusion médicale publié par SPF à celui estimé à partir des données du SNDS dans la cohorte de 20,5 millions de personnes âgées de 50 à 74ans et cibles de la campagne biannuelle de 2020-2021 en France (Cohorte_20_21). Dans cette Cohorte_20_21, SPF avait estimé le taux d'exclusion à 13,2% et le taux de participation à la campagne à 35,7%[27].
Dans cette Cohorte_20_21, l'évaluation comparative sera faite avec l'hypothèse de non-infériorité du taux estimé à partir du nombre de cas d'exclusion médicale extraits du SNDS par l'algorithme. En plus d'une déclinaison géographique par département/région ou par zone (ultra-rural, ultra-urbain), cette comparaison tiendra compte des stratégies actuelles mises en place par les structures de gestion pour optimiser leur collecte des données d'exclusion médicale (Une enquête en cours définira les profils des structures de gestion selon les stratégies).
Le second objectif spécifique sera exploré dans une étude Cas-Témoins. Les Cas seront toutes les personnes de la Cohorte_20_21 n'ayant aucun remboursement de test de dépistage récent (2020-2022). Les Témoins seront toutes les personnes de la Cohorte_20_21 ayant un remboursement de test de dépistage sur la période 2020-2022. Les fréquences de situations morbides (maladies chroniques, exploration diagnostique, traitement médicamenteux au long cours, antérieures à 2020 seront comparées en Cas et Témoins, avec l'hypothèse que les cas n'ont pas réalisé de test sur la campagne 2020-2021 à cause d'une situation morbide enregistrée en amont de la campagne.

2.1
2.2

3.2.2 Choix du schéma de circulation du NIR

Aucun chainage ou appariement ne sera fait dans cette étude. Pas besoin de choix d'un schéma.

3.2.3 Infrastructure informatique utilisable

Nous souhaiterions recourir aux plateformes sécurisées de l'INSERM et du portail de la CNAM pour le traitement des données.
En rapport avec le premier objectif spécifique, la première partie de cette analyse statistique sera essentiellement descriptive. Elle consistera à comparer à l'aide d'un test du chi-2, le taux d'exclusion théorique (issu des données actuelles du DOCCR publiées par SPF) au taux d'exclusion réel obtenu après l'exploitation des données du SNDS. Le taux de participation estimé avant et après cette amélioration du taux d'exclusion sera décrit en fonction des stratégies d'exclusion dans les départements qui répondront à l'enquête en cours.
En rapport avec le second objectif spécifique, la seconde partie des analyses statistiques sera à la fois descriptive et étiologique (étude cas-temoin). Les modèles prédictifs seront construits pour évaluer les sous objectifs énumérés au point «Publication et valorisation des résultats ».

3.2.4 Circulation synthétique des données et appariement

Les données issues de l'extraction du SNDS seront stockées dans l'espace de traitement dédié à l'INSERM et sur le portail de données de la CNAM. Toutes les analyses statistiques seront réalisées dans cet espace dédié. Seuls les résultats agrégés issus des analyses statistiques seront exportés sur le site du LIMICS pour interprétation et valorisation.

3.2.1 Information individuelle des patients

L'article 21 (point 6) du RGPD recommande : « Lorsque des données à caractère personnel sont traitées à des fins de recherche scientifique ou historique ou à des fins statistiques en application de l'article 89, paragraphe 1, la personne concernée a le droit de s'opposer, pour des raisons tenant à sa situation particulière, au traitement de données à caractère personnel la concernant, à moins que le traitement ne soit nécessaire à l'exécution d'une mission d'intérêt public ». Le traitement envisagé dans ce projet sera mené à des fins de recherche scientifique visant à améliorer la qualité d'une mission d'intérêt public (Le programme national du dépistage des cancers), les droits prévus aux articles 15 à 20 du RGPD (accès, rectification, effacement, limitation et portabilité) ne sont pas applicables.

3.2.2 Respect du droit des personnes

L'extraction est anonymisée et aucun NIR ne sera utilisé.

3.2.3 Support des données et sécurité

A l'exception de la plateforme INSERM, l'usage d'un autre support de données n'est pas envisagé. Toutefois, les résultats des analyses statistiques réalisées sur l'espace dédié l'INSERM, pourraient être sauvegardés sur le serveur local du LIMICS. Seuls des résultats agrégés seront exportés de l'environnement de travail vers le serveur local du LIMICS.

3.2.4 Archivage des documents et Rapport final de l'étude

Le rapport d'étude ainsi que l'ensemble des exportations (données agrégées et résultats d'analyses) seront archivés par le LIMICS.

3.3 Calendrier prévisionnel et faisabilité du projet

Calendrier : Date de début du Projet: 01/10/2021 ; Date de Fin du projet: 30/12/2023, soit une durée de 26 mois
Grandes lignes du calendrier:
1 - Elaboration du protocole : Initié en octobre 2021 et finalisé en décembre 2022.
2 - Revue de la littérature et élaboration de l'algorithme: Initié en octobre 2021 et finalisé en septembre 2022 sous la coordination de Professeur Catherine DUCLOS (LIMICS).
3 - Formation à l'exploitation des données du SNDS: Professeur Catherine DUCLOS (LIMICS) et Docteur Akoï KOÏVOGU (LIMICS) ont à ce jour finalisé les trois formations obligatoires (SNIIRAM - Données Individuelles Bénéficiaires ; SNIIRAM SAS Entreprise Guide - Prise en main du logiciel et utilisation dans le SNIIRAM; Architecture et données du SNDS).
4 - Dépôt de la demande d'accès au SNDS, en cours.
5 - Analyse de données et Interprétation de résultats: Initié en octobre 2021 (revue de la littérature) et date prévisionnelle de fin décembre 2023. Lieu d'analyse: LIMICS, site Université Paris 13.
6 - Publications et diffusion des résultats: Mars - Juin 2024.

2.1
Faisabilité : Le LIMICS est spécialisée dans le traitement informatique des données de santé. Ce projet de phénotypage qui s'inscrit dans son domaine d'expertise est très compatible avec leurs projets en cours. Le Docteur Akoï KOIVOGUI ayant une longue expérience en gestion des bases de données mais aussi en analyse statistique, consacrera 5 vacations par semaine. Ayant plusieurs années d'expérience dans le traitement des données du PMSI, le professeur Catherine DUCLOS supervisera l'élaboration et la validation des requêtes informatiques et participera à l'analyse des données. Centre d'expertise reconnue dans la recherche en oncologie digestive, le Service de Gastroentérologie de l'Hôpital Avicenne a toujours été très impliqué dans les projets de recherche conduit par le centre de coordination du dépistage. Le Professeur Robert BENAMOUZIG (supervision) et le docteur David DEUTSCH (aide à l'élaboration des requêtes) consacreront chacun une vacation par mois.

4. BIBLIOGRAPHIE

1.Binder-Foucard F, Belot A, Delafosse P, et al. Estimation nationale de l'incidence et de la mortalité par cancer en France entre 1980 et 2012. Partie 1 - Tumeurs solides. Saint-Maurice (Fra): Institut de veille sanitaire, 2013.
2.Faivre J, Dancourt V, Lejeune C, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. Gastroenterology. 2004;126(6):1674 - 80.
3.Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet (London, England). 1996;348(9040):1472-7.
4.Barre S, Leleu H, Benamouzig R, et al. Cost-effectiveness analysis of alternative colon cancer screening strategies in the context of the French national screening program. Therap Adv Gastroenterol. 2020;13:1756284820953364.
5.Faivre J, Lepage C, Viguier J. [Colorectal cancer: from diagnosis to screening]. Gastroenterol Clin Biol. 2009;33(8-9):660-71.
6.Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA: a cancer journal for clinicians. 2009;59(6):366-78. Epub 2009/11/10.
7.Welch HG, Robertson DJ. Colorectal Cancer on the Decline--Why Screening Can't Explain It All. N Engl J Med. 2016;374(17):1605-7.
8.Lejeune C, Arveux P, Dancourt V, et al. Cost-effectiveness analysis of fecal occult blood screening for colorectal cancer. Int J Technol Assess Health Care. 2004;20(4):434-9.
9.Leuraud K, Jezewski-Serra D, Viguier J, Salines E. Colorectal cancer screening by guaiac faecal occult blood test in France: Evaluation of the programme two years after launching. Cancer epidemiology. 2013;37(6):959-67.
10.Tazi MA, Faivre J, Dassonville F, Lamour J, Milan C, Durand G. Participation in faecal occult blood screening for colorectal cancer in a well defined French population: results of five screening rounds from 1988 to 1996. Journal of medical screening. 1997;4(3):147-51.
11.Moss SM, Campbell C, Melia J, et al. Performance measures in three rounds of the English bowel cancer screening pilot. Gut. 2012;61(1):101-7.
12.Pornet C, Denis B, Perrin P, et al. Predictors of adherence to repeat fecal occult blood test in a population-based colorectal cancer screening program. British journal of cancer. 2014;111(11):2152-5.
13.Pollack CE, Mallya G, Polsky D. The impact of consumer-directed health plans and patient socioeconomic status on physician recommendations for colorectal cancer screening. J Gen Intern Med. 2008;23(10):1595-601.
14.Guignot C. Cancer colorectal. Vers un dépistage mieux accepté. Science et Santé. 2015;26:36-7.
15.Le Pimpec F, Moutel G, Piette C, et al. Fecal immunological blood test is more appealing than the guaiac-based test for colorectal cancer screening. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2017;49 (11):1267-72.
16.Lieberman D. Screening/early detection model for colorectal cancer. Why screen? Cancer. 1994;74(7 Suppl):2023-7.
17.Altobelli E, Lattanzi A, Paduano R, et al. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. Preventive medicine. 2014;62:132-41.
18.JORF. Arrêté du 19 mars 2018 modifiant l'arrêté du 29 septembre 2006 relatif aux programmes de dépistage organisé des cancers et portant modification du cahier des charges du dépistage organisé du cancer colorectal. JORF n° 0068 du 22 mars 2018 texte n° 13. Journal Officiel de la République Française (JORF). 2018;0068:NOR: SSAP1807442A
19.SPF. Taux de participation au programme de dépistage organisé du cancer colorectal 2018-2019. Saint Maurice, France: Santé Publique France(SPF), 2020.
20.JORF. Arrêté du 29 septembre 2006 relatif aux programmes du dépistage des cancers. Journal Officiel de la République Française. Annexe au n° 295.32003-48. 2006.
21.Sannino N. Tableau de bord du dépistage organise du cancer colorectal en Ile-de-France, donnees 2015-2016. Paris, France: 2018 19-Mai-2018. Report No.
22.CNAM. Cartographie des pathologies et des dépenses: Prévalence départementale par pathologie (2018). Caisse Nationale d'Assurance Maladie (CNAM): Direction de la Stratégie, des Etudes et des Statistiques-Département des Etudes sur les Pathologies et les Patients. 2020.
23.Bernardini D, Bulois P, Barthet M, et al. « Une semaine de coloscopie en France » : résultats 2017 de l'enquête annuelle de la Société française d'endoscopie digestive. Acta Endoscopica. 2017;47(5):242-51.
24.JORF. Arrêté du 23 mars 2018 portant modification de l'arrêté du 29 septembre 2006 relatif aux programmes de dépistage des cancers. NOR: SSAP1805226A. JORF. 2018;0078(9).
25.Tuppin P, Rudant J, Constantinou P, et al. Value of a national administrative database to guide public decisions: From the systeme national d'information interregimes de l'Assurance Maladie (SNIIRAM) to the systeme national des donnees de sante (SNDS) in France. Rev Epidemiol Sante Publique. 2017;65 Suppl 4:S149-S67.
26.Aguade AS, Gastaldi-Menager C, Denis P, et al. How to Estimate The Cost Of Diabetes Based on Information From The French Health Insurance Database (Sniiram)? Value Health. 2014;17(7):A341.
27.SPF. Programme de dépistage organisé du cancer colorectal. Un Taux de participation en deça des objectifs attendus. Saint Maurice, France: Santé Publique France(SPF), 2022.

Inserm
La science pour la santé
From science to health

5. ENGAGEMENTS

Le demandeur habilité* :

Le Président Directeur-Général de l'Inserm a délégué son pouvoir d'Autorité d'Enregistrement (AE) à des Autorités d'Enregistrement Déléguées (AED).

Le demandeur habilité s'engage à :

1 / Signaler aux AED de l'Inserm, tout changement administratif dans ses fonctions et projets (départ de l'unité au plus tard une semaine avant celui-ci, changement d'affectation dans l'unité concernée, changement de projet, fin de projet, fin d'utilisation des données du SNDS, etc…) pouvant impacter les droits d'accès aux données EGB/SNDS.

2 / Signaler aux AED de l'Inserm, tout nouveau projet d'exploitation des données de l'EGB ou du SNDS via une nouvelle demande sur le site internet du Centre de services SNDS de l'Inserm (y compris si les droits d'accès précédemment accordés couvrent les besoins du nouveau projet).

3 / Se connecter au portail de données de la Cnam uniquement avec le bureau déporté sécurisé (PAS) du DSI Inserm (soumis aux exigences de la PSSI-MCAS).

4 / Respecter les Conditions Générales d'Utilisation (CGU) du SNDS qui seront signées lors de la première connexion au portail de données de la Cnam.

Fait à BOBIGNY **le** 18-Jan-2023

Nom : KOIVOGUI

Prénom : AKOI

Signature :



Le directeur d'unité Inserm du demandeur habilité précité :

Le directeur d'unité Inserm du demandeur atteste que :

1 / Les informations renseignées dans le présent document concernant le demandeur habilité et son projet d'étude sont correctes et que le demandeur habilité est bien un personnel emmargeant au profil de son unité/équipe Inserm (employé officiellement rattaché à son unité/équipe Inserm, étudiant ayant signé une convention de stage ou de doctorat avec son unité/équipe Inserm ou prestataire sous contrat et agissant pour le compte de son unité/équipe Inserm).

2 / Le projet d'étude susmentionné soit bien sous responsabilité de traitement** ou coresponsabilité de traitement Inserm.

Fait à BOBIGNY **le** 18/01/2023

Nom : JAULENT

Prénom : MARIE-CHRISTINE

Signature :



*Ayant suivi les formations SNDS obligatoires en fonction de l'accès demandé dispensées par l'Institut 4.10)

**Un projet est considéré sous responsabilité de traitement Inserm s'il est commandité par l'Inserm. Les moyens du traitement ainsi que ses finalités doivent être déterminées par l'Inserm (« *Le responsable d'un traitement de données à caractère personnel est (…) la personne, l'autorité publique, le service ou l'organisme qui détermine ses finalités et ses moyens. En pratique et en général, il s'agit de la personne morale incarnée par son représentant légal.* » Source : Cnil).

9. BIBLIOGRAPHIE

1. Binder-Foucard F, Belot A, Delafosse P, Remontet L, Woronoff A-S, Bossard N. Estimation nationale de l'incidence et de la mortalité par cancer en France entre 1980 et 2012. Partie 1 - Tumeurs solides. Saint-Maurice (Fra): Institut de veille sanitaire, 2013.
2. Faivre J, Dancourt V, Lejeune C, Tazi MA, Lamour J, Gerard D, et al. Reduction in colorectal cancer mortality by fecal occult blood screening in a French controlled study. Gastroenterology. 2004;126(6):1674 - 80. doi: 10.1053/j.gastro.2004.02.018. PubMed PMID: 15609792.
3. Hardcastle JD, Chamberlain JO, Robinson MH, Moss SM, Amar SS, Balfour TW, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. Lancet (London, England). 1996;348(9040):1472-7. doi: 10.1016/S0140-6736(96)03386-7. PubMed PMID: 8942775.
4. Kronborg O, Fenger C, Olsen J, Jorgensen OD, Sondergaard O. Randomised study of screening for colorectal cancer with faecal-occult-blood test. Lancet (London, England). 1996;348(9040):1467-71. doi: 10.1016/S0140-6736(96)03430-7. PubMed PMID: 8942774.
5. Recommendations on cancer screening in the European union. Advisory Committee on Cancer Prevention. Eur J Cancer. 2000;36(12):1473-8. doi: 10.1016/s0959-8049(00)00122-2. PubMed PMID: 10930794.
6. Faivre J. Colorectal cancer screening in Europe: an overview. Cancero dig. 2010;2(2):101 - 3. doi: 10.4267/2042/32415.
7. Leuraud K, Jezewski-Serra D, Viguier J, Salines E. Colorectal cancer screening by guaiac faecal occult blood test in France: Evaluation of the programme two years after launching. Cancer epidemiology. 2013;37(6):959-67. doi: 10.1016/j.canep.2013.07.008. PubMed PMID: 24035240.
8. JORF. Arrêté du 29 septembre 2006 relatif aux programmes du dépistage des cancers. Journal Officiel de la République Française. Annexe au n° 295.32003-48. 2006.
9. Barre S, Leleu H, Benamouzig R, Saurin JC, Vimont A, Taleb S, et al. Cost-effectiveness analysis of alternative colon cancer screening strategies in the context of the French national screening program. Therap Adv Gastroenterol. 2020;13:1756284820953364. doi: 10.1177/1756284820953364. PubMed PMID: 33014138; PubMed Central PMCID: PMCPMC7509710.
10. Faivre J, Lepage C, Viguier J. [Colorectal cancer: from diagnosis to screening]. Gastroenterol Clin Biol. 2009;33(8-9):660-71. doi: 10.1016/j.gcb.2009.07.008. PubMed PMID: 19709833.

11. Center MM, Jemal A, Smith RA, Ward E. Worldwide variations in colorectal cancer. CA: a cancer journal for clinicians. 2009;59(6):366-78. Epub 2009/11/10. doi: 10.3322/caac.20038. PubMed PMID: 19897840.
12. Welch HG, Robertson DJ. Colorectal Cancer on the Decline--Why Screening Can't Explain It All. N Engl J Med. 2016;374(17):1605-7. doi: 10.1056/NEJMp1600448. PubMed PMID: 27119236.
13. Lejeune C, Arveux P, Dancourt V, Bejean S, Bonithon-Kopp C, Faivre J. Cost-effectiveness analysis of fecal occult blood screening for colorectal cancer. Int J Technol Assess Health Care. 2004;20(4):434-9. PubMed PMID: 15609792.
14. JORF. Arrêté du 23 mars 2018 portant modification de l'arrêté du 29 septembre 2006 relatif aux programmes de dépistage des cancers. NOR: SSAP1805226A. JORF. 2018;0078(9).
15. European Colorectal Cancer Screening Guidelines Working G, von Karsa L, Patnick J, Segnan N, Atkin W, Halloran S, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis: overview and introduction to the full supplement publication. Endoscopy. 2013;45(1):51-9. doi: 10.1055/s-0032-1325997. PubMed PMID: 23212726; PubMed Central PMCID: PMCPMC4482205.
16. Bulliard JL, Garcia M, Blom J, Senore C, Mai V, Klabunde C. Sorting out measures and definitions of screening participation to improve comparability: the example of colorectal cancer. Eur J Cancer. 2014;50(2):434-46. doi: 10.1016/j.ejca.2013.09.015. PubMed PMID: 24144735.
17. Klabunde C, Blom J, Bulliard JL, Garcia M, Hagoel L, Mai V, et al. Participation rates for organized colorectal cancer screening programmes: an international comparison. Journal of medical screening. 2015;22(3):119-26. doi: 10.1177/0969141315584694. PubMed PMID: 25967088.
18. Tazi MA, Faivre J, Dassonville F, Lamour J, Milan C, Durand G. Participation in faecal occult blood screening for colorectal cancer in a well defined French population: results of five screening rounds from 1988 to 1996. Journal of medical screening. 1997;4(3):147-51.
19. Moss SM, Campbell C, Melia J, Coleman D, Smith S, Parker R, et al. Performance measures in three rounds of the English bowel cancer screening pilot. Gut. 2012;61(1):101-7. doi: 10.1136/gut.2010.236430. PubMed PMID: 21561880; PubMed Central PMCID: PMCPMC3230830.
20. Koïvogui A, Ecochardb R, Le Mab G, Benamouzigd R. Impact of stopping sending colorectal cancer screening test kits by regular mail. Public Health. 2019;173:33-41. doi: 10.1016/j.puhe.2019.05.020.
21. Barre S, Leleu H, Vimont A, Kaufmanis A, Gendre I, Taleb S, et al. [Estimated impact of the current colorectal screening program in France]. Rev Epidemiol Sante Publique. 2020;68(3):171-7. doi: 10.1016/j.respe.2020.04.053. PubMed PMID: 32417153.

22. Pornet C, Denis B, Perrin P, Gendre I, Launoy G. Predictors of adherence to repeat fecal occult blood test in a population-based colorectal cancer screening program. British journal of cancer. 2014;111(11):2152-5. doi: 10.1038/bjc.2014.507. PubMed PMID: 25314056; PubMed Central PMCID: PMCPMC4260031.
23. Pollack CE, Mallya G, Polsky D. The impact of consumer-directed health plans and patient socioeconomic status on physician recommendations for colorectal cancer screening. J Gen Intern Med. 2008;23(10):1595-601. doi: 10.1007/s11606-008-0714-x. PubMed PMID: 18629590; PubMed Central PMCID: PMCPMC2533392.
24. Guignot C. Cancer colorectal. Vers un dépistage mieux accepté. Science et Santé. 2015;26:36-7.
25. Le Pimpec F, Moutel G, Piette C, Lievre A, Bretagne JF. Fecal immunological blood test is more appealing than the guaiac-based test for colorectal cancer screening. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2017;49(11):1267-72. doi: 10.1016/j.dld.2017.08.018. PubMed PMID: 28867474.
26. Giorgi Rossi P, Grazzini G, Anti M, Baiocchi D, Barca A, Bellardini P, et al. Direct mailing of faecal occult blood tests for colorectal cancer screening: a randomized population study from Central Italy. Journal of medical screening. 2011;18(3):121-7. Epub 2011/11/03. doi: 10.1258/jms.2011.011009. PubMed PMID: 22045820.
27. Camilloni L, Ferroni E, Cendales BJ, Pezzarossi A, Furnari G, Borgia P, et al. Methods to increase participation in organised screening programs: a systematic review. BMC Public Health. 2013;13:464. Epub 2013/05/15. doi: 10.1186/1471-2458-13-464. PubMed PMID: 23663511; PubMed Central PMCID: PMCPmc3686655.
28. Cole SR, Smith A, Wilson C, Turnbull D, Esterman A, Young GP. An advance notification letter increases participation in colorectal cancer screening. Journal of medical screening. 2007;14(2):73-5. Epub 2007/07/14. doi: 10.1258/096914107781261927. PubMed PMID: 17626705.
29. Tinmouth J, Ritvo P, McGregor SE, Claus D, Pasut G, Myers RE, et al. A qualitative evaluation of strategies to increase colorectal cancer screening uptake. Canadian family physician Medecin de famille canadien. 2011;57(1):e7-15. Epub 2011/02/17. PubMed PMID: 21322288; PubMed Central PMCID: PMCPmc3024174.
30. van Roon AH, Hol L, Wilschut JA, Reijerink JC, van Vuuren AJ, van Ballegooijen M, et al. Advance notification letters increase adherence in colorectal cancer screening: a population-based randomized trial. Preventive medicine. 2011;52(6):448-51. Epub 2011/04/05. doi: 10.1016/j.ypmed.2011.01.032. PubMed PMID: 21457725.

31. Van Roosbroeck S, Hoeck S, Van Hal G. Population-based screening for colorectal cancer using an immunochemical faecal occult blood test: a comparison of two invitation strategies. Cancer epidemiology. 2012;36(5):e317-24. Epub 2012/05/09. doi: 10.1016/j.canep.2012.04.003. PubMed PMID: 22560885.
32. Barthe J, Perrodeau E, Gilberg S, Ravaud P, Ghasarossian C, Marchand-Buttin F, et al. Impact of a Doctor's Invitation on Participation in Colorectal Cancer Screening: A Cluster Randomized Trial. The American journal of medicine. 2015;128(9):1024.e1-7. Epub 2015/04/25. doi: 10.1016/j.amjmed.2015.03.026. PubMed PMID: 25908395.
33. Courtier R, Casamitjana M, Macia F, Panades A, Castells X, Gil MJ, et al. Participation in a colorectal cancer screening programme: influence of the method of contacting the target population. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2002;11(3):209-13. Epub 2002/07/20. PubMed PMID: 12131653.
34. Altobelli E, Lattanzi A, Paduano R, Varassi G, di Orio F. Colorectal cancer prevention in Europe: burden of disease and status of screening programs. Preventive medicine. 2014;62:132-41. doi: 10.1016/j.ypmed.2014.02.010. PubMed PMID: 24530610.
35. Goulard H, Boussac-Zarebska M, Bloch J. Epidemiological assessment of the pilot programme for organized colorectal cancer screening, France,2007. BEH. 2009;2-3:22-5.
36. Tifratene K, Eisinger F, Rinaldi Y, Didelot R, Seitz JF. Colorectal cancer screening program: cost effectiveness of systematic recall letters. Gastroenterol Clin Biol. 2007;31(11):929-33. PubMed PMID: 18166880.
37. Faivre J, Manfredi S. Le dépistage et la prévention du cancer colorectal : il n'y a plus de temps à perdre. Hépato Gastro. 2015;22(4):273-7. doi: 10.1684/hpg.2015.1134.
38. JORF. Arrêté du 19 mars 2018 modifiant l'arrêté du 29 septembre 2006 relatif aux programmes de dépistage organisé des cancers et portant modification du cahier des charges du dépistage organisé du cancer colorectal. JORF n° 0068 du 22 mars 2018 texte n° 13. Journal Officiel de la République Française (JORF). 2018;0068:NOR: SSAP1807442A
39. Sannino N. Tableau de bord du dépistage organise du cancer colorectal en Ile-de-France, donnees 2015-2016. Paris, France: 2018 19-Mai-2018. Report No.
40. SPF. Taux de participation au programme de dépistage organisé du cancer colorectal 2018-2019. Saint Maurice, France: Santé Publique France(SPF), 2020.

41. CNAM. Cartographie des pathologies et des dépenses: Prévalence départementale par pathologie (2018). Caisse Nationale d'Assurance Maladie(CNAM): Direction de la Stratégie, des Etudes et des Statistiques-Département des Etudes sur les Pathologies et les Patients.
. 2020.
42. Bernardini D, Bulois P, Barthet M, Chaussade S, Gronier O, Lecomte T, et al. « Une semaine de coloscopie en France » : résultats 2017 de l'enquête annuelle de la Société française d'endoscopie digestive. Acta Endoscopica. 2017;47(5):242-51. doi: 10.1007/s10190-017-0613-2.
43. Goldsbury D, Weber M, Yap S, Banks E, O'Connell DL, Canfell K. Identifying incident colorectal and lung cancer cases in health service utilisation databases in Australia: a validation study. BMC medical informatics and decision making. 2017;17(1):23. doi: 10.1186/s12911-017-0417-5. PubMed PMID: 28241763; PubMed Central PMCID: PMCPMC5327557.
44. Worthington J, He E, Lew JB, St John J, Horn C, Grogan P, et al. Colonoscopies in Australia - how much does the National Bowel Cancer Screening Program contribute to colonoscopy use? Public Health Res Pract. 2022. doi: 10.17061/phrp32232216. PubMed PMID: 36477980.
45. Kelty E, Ward SV, Cadby G, McCarthy NS, O'Leary P, Moses EK, et al. Familial and non-familial risk factors associated with incidence of colorectal cancer in young and middle-aged persons in Western Australia. Cancer epidemiology. 2019;62:101591. doi: 10.1016/j.canep.2019.101591. PubMed PMID: 31494463.
46. Macken E, Van Dongen S, De Brabander I, Francque S, Driessen A, Van Hal G. Post-colonoscopy colorectal cancer in Belgium: characteristics and influencing factors. Endosc Int Open. 2019;7(5):E717-E27. doi: 10.1055/a-0751-2660. PubMed PMID: 31073539; PubMed Central PMCID: PMCPMC6506335.
47. Tinmouth J, Sutradhar R, Liu N, Baxter NN, Paszat L, Rabeneck L. Validation of 5 key colonoscopy-related data elements from Ontario health administrative databases compared to the clinical record: a cross-sectional study. CMAJ Open. 2018;6(3):E330-E8. doi: 10.9778/cmajo.20180013. PubMed PMID: 30104417; PubMed Central PMCID: PMCPMC6182115.
48. Castelo M, Paszat L, Hansen BE, Scheer AS, Faught N, Nguyen L, et al. Measurement of clinical delay intervals among younger adults with colorectal cancer using health administrative data: a population-based analysis. BMJ Open Gastroenterol. 2022;9(1). doi: 10.1136/bmjgast-2022-001022. PubMed PMID: 36410773; PubMed Central PMCID: PMCPMC9680148.
49. Hol L, Sutradhar R, Gu S, Baxter NN, Rabeneck L, Tinmouth JM, et al. Repeat colonoscopy after a colonoscopy with a negative result in Ontario: a population-based cohort study. CMAJ Open.

2015;3(2):E244-50. doi: 10.9778/cmajo.20140063. PubMed PMID: 26389103; PubMed Central PMCID: PMCPMC4565168.

50. Benchimol EI, Kuenzig ME, Bernstein CN, Nguyen GC, Guttmann A, Jones JL, et al. Rural and urban disparities in the care of Canadian patients with inflammatory bowel disease: a population-based study. Clinical epidemiology. 2018;10:1613-26. doi: 10.2147/CLEP.S178056. PubMed PMID: 30519110; PubMed Central PMCID: PMCPMC6233859.
51. Paszat LF, Sutradhar R, Luo J, Baxter NN, Tinmouth J, Rabeneck L. Morbidity and mortality after major large bowel resection of non-malignant polyp among participants in a population-based screening program. Journal of medical screening. 2021;28(3):261-7. doi: 10.1177/0969141320967960. PubMed PMID: 33153368; PubMed Central PMCID: PMCPMC8366188.
52. Larsen PT, Jorgensen SF, Njor SH. Colorectal cancer screening participation among citizens not recommended to be screened: a cohort study. BMC Gastroenterol. 2022;22(1):256. doi: 10.1186/s12876-022-02331-9. PubMed PMID: 35596148; PubMed Central PMCID: PMCPMC9123659.
53. Schmidt M, Schmidt SA, Sandegaard JL, Ehrenstein V, Pedersen L, Sorensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. Clinical epidemiology. 2015;7:449-90. doi: 10.2147/CLEP.S91125. PubMed PMID: 26604824; PubMed Central PMCID: PMCPMC4655913.
54. Lash TL, Riis AH, Ostenfeld EB, Erichsen R, Vyberg M, Thorlacius-Ussing O. A validated algorithm to ascertain colorectal cancer recurrence using registry resources in Denmark. Int J Cancer. 2015;136(9):2210-5. doi: 10.1002/ijc.29267. PubMed PMID: 25307704.
55. Albaek Jacobsen H, Jess T, Larsen L. Validity of Inflammatory Bowel Disease Diagnoses in the Danish National Patient Registry: A Population-Based Study from the North Denmark Region. Clinical epidemiology. 2022;14:1099-109. doi: 10.2147/CLEP.S378003. PubMed PMID: 36226162; PubMed Central PMCID: PMCPMC9550174.
56. Kirkegard J, Farkas DK, Jorgensen JOL, Cronin-Fenton D. Hyper- and hypothyroidism and gastrointestinal cancer risk: a Danish nationwide cohort study. Endocr Connect. 2018;7(11):1129-35. doi: 10.1530/EC-18-0258. PubMed PMID: 30352404; PubMed Central PMCID: PMCPMC6215792.
57. Czech Hospitalization yearbook, 2017. Institute of Health Information and Statistics of the Czech Republic, Prague. ISSN 1210- 8731, (0862-5654) Available at: https://www.uzis.cz/index-en.php?pg=publications--library&id=238.
58. Friedman MY, Leventer-Roberts M, Rosenblum J, Zigman N, Goren I, Mourad V, et al. Development and validation of novel algorithms to identify patients with inflammatory bowel diseases in Israel: an epi-IIRN

group study. Clinical epidemiology. 2018;10:671-81. doi: 10.2147/CLEP.S151339. PubMed PMID: 29922093; PubMed Central PMCID: PMCPMC5995295.

59. Zorzi M, Gennaro N, Capodaglio G, Urso EDL, Pucciarelli S, Memo L, et al. Colorectal cancer screening: The surgery rates they are a-changing. A nationwide study on surgical resections in Italy. Digestive and liver disease : official journal of the Italian Society of Gastroenterology and the Italian Association for the Study of the Liver. 2019;51(2):304-9. doi: 10.1016/j.dld.2018.10.011. PubMed PMID: 30449608.
60. Meregaglia M, Banks H, Fattore G. Hospital Burden and Gastrointestinal Surgery in Inflammatory Bowel Disease Patients in Italy: A Retrospective Observational Study. J Crohns Colitis. 2015;9(10):853-62. doi: 10.1093/ecco-jcc/jjv104. PubMed PMID: 26069194.
61. Santiago M, Magro F, Correia L, Portela F, Ministro P, Lago P, et al. Inflammatory Bowel Disease Reoperation Rate Has Decreased Over Time If Corrected by Prevalence. Clin Transl Gastroenterol. 2020;11(9):e00227. doi: 10.14309/ctg.0000000000000227. PubMed PMID: 33094955.
62. Recalde M, Manzano-Salgado CB, Diaz Y, Puente D, Garcia-Gil MDM, Marcos-Gragera R, et al. Validation Of Cancer Diagnoses In Electronic Health Records: Results From The Information System For Research In Primary Care (SIDIAP) In Northeast Spain. Clinical epidemiology. 2019;11:1015-24. doi: 10.2147/CLEP.S225568. PubMed PMID: 31819655; PubMed Central PMCID: PMCPMC6899079.
63. Mendez-Bailon M, Jimenez-Garcia R, Munoz-Rivas N, Hernandez-Barrera V, de Miguel-Yanes JM, de Miguel-Diez J, et al. Trends and Clinical Impact of Gastrointestinal Endoscopic Procedures on Acute Heart Failure in Spain (2002-2017). J Clin Med. 2021;10(3). doi: 10.3390/jcm10030546. PubMed PMID: 33540753; PubMed Central PMCID: PMCPMC7867243.
64. Yan MK, Adler NR, Heriot N, Shang C, Zalcberg JR, Evans S, et al. Opportunities and barriers for the use of Australian cancer registries as platforms for randomized clinical trials. Asia Pac J Clin Oncol. 2022;18(4):344-52. doi: 10.1111/ajco.13670. PubMed PMID: 34811922.
65. Decker KM, Lambert P, Nugent Z, Biswanger N, Samadder J, Singh H. Time Trends in the Diagnosis of Colorectal Cancer With Obstruction, Perforation, and Emergency Admission After the Introduction of Population-Based Organized Screening. JAMA Netw Open. 2020;3(5):e205741. doi: 10.1001/jamanetworkopen.2020.5741. PubMed PMID: 32453385; PubMed Central PMCID: PMCPMC7251446.
66. Kirac I, Sekerija M, Simunovic I, Zgaga L, Velimir Vrdoljak D, Kovacevic D, et al. Incidence and mortality trends of gastric and colorectal cancers in Croatia, 1988-2008. Croat Med J. 2012;53(2):124-34. doi: 10.3325/cmj.2012.53.124. PubMed PMID: 22522990; PubMed Central PMCID: PMCPMC3342651.

67. Christensen J, Hojsgaard Schmidt LK, Kejs AMT, Sogaard J, Rasted MC, Andersen O, et al. Agreement between the Danish Cancer Registry and the Danish Colorectal Cancer Group Database. Acta Oncol. 2020;59(1):116-23. doi: 10.1080/0284186X.2019.1669815. PubMed PMID: 31559881.
68. Pehalova L, Krejci D, Snajdrova L, Dusek L. Cancer incidence trends in the Czech Republic. Cancer epidemiology. 2021;74:101975. doi: 10.1016/j.canep.2021.101975. PubMed PMID: 34247066.
69. Tron L, Belot A, Fauvernier M, Remontet L, Bossard N, Launay L, et al. Socioeconomic environment and disparities in cancer survival for 19 solid tumor sites: An analysis of the French Network of Cancer Registries (FRANCIM) data. Int J Cancer. 2019;144(6):1262-74. doi: 10.1002/ijc.31951. PubMed PMID: 30367459.
70. Magid A, Robinson E, Keinan-Boker L. Preventable Cancer in Israel. Int J Environ Res Public Health. 2022;19(17). doi: 10.3390/ijerph191710521. PubMed PMID: 36078236; PubMed Central PMCID: PMCPMC9518384.
71. Giorgi Rossi P, Carretta E, Mangone L, Baracco S, Serraino D, Zorzi M, et al. Incidence of interval cancers in faecal immunochemical test colorectal screening programmes in Italy. Journal of medical screening. 2018;25(1):32-9. doi: 10.1177/0969141316686391. PubMed PMID: 28195508.
72. Gauci D, Allemani C, Woods L. Population-level cure of colorectal cancer in Malta: An analysis of patients diagnosed between 1995 and 2004. Cancer epidemiology. 2016;42:32-8. doi: 10.1016/j.canep.2016.03.001. PubMed PMID: 26999401.
73. Morais S, Antunes L, Rodrigues J, Fontes F, Bento MJ, Lunet N. The impact of the COVID-19 pandemic on the short-term survival of patients with cancer in Northern Portugal. Int J Cancer. 2021;149(2):287-96. doi: 10.1002/ijc.33532. PubMed PMID: 33634852; PubMed Central PMCID: PMCPMC8014057.
74. Boltezar L, Novakovic BJ, Moltara ME. Trends in specialized palliative care referrals at an oncology center from 2007 to 2019. BMC Palliat Care. 2021;20(1):135. doi: 10.1186/s12904-021-00828-w. PubMed PMID: 34479512; PubMed Central PMCID: PMCPMC8418009.
75. Feller A, Matthes KL, Bordoni A, Bouchardy C, Bulliard JL, Herrmann C, et al. The relative risk of second primary cancers in Switzerland: a population-based retrospective cohort study. BMC cancer. 2020;20(1):51. doi: 10.1186/s12885-019-6452-0. PubMed PMID: 31964352; PubMed Central PMCID: PMCPMC6974968.
76. Schulpen M, van den Brandt PA. Mediterranean diet adherence and risk of colorectal cancer: the prospective Netherlands Cohort Study. European journal of epidemiology. 2020;35(1):25-35. doi: 10.1007/s10654-019-00549-8. PubMed PMID: 31494792; PubMed Central PMCID: PMCPMC7058569.

77. Orlich MJ, Singh PN, Sabate J, Fan J, Sveen L, Bennett H, et al. Vegetarian dietary patterns and the risk of colorectal cancers. JAMA Intern Med. 2015;175(5):767-76. doi: 10.1001/jamainternmed.2015.59. PubMed PMID: 25751512; PubMed Central PMCID: PMCPMC4420687.
78. Verma M, Sarfaty M, Brooks D, Wender RC. Population-based programs for increasing colorectal cancer screening in the United States. CA: a cancer journal for clinicians. 2015;65(6):497-510. doi: 10.3322/caac.21295. PubMed PMID: 26331705.
79. Coronado GD, Leo MC, Ramsey K, Coury J, Petrik AF, Patzel M, et al. Mailed fecal testing and patient navigation versus usual care to improve rates of colorectal cancer screening and follow-up colonoscopy in rural Medicaid enrollees: a cluster-randomized controlled trial. Implement Sci Commun. 2022;3(1):42. doi: 10.1186/s43058-022-00285-3. PubMed PMID: 35418107; PubMed Central PMCID: PMCPMC9006522.
80. Monteiro H, Tavares F, Reis J, Ferreira G, Campos MJ, Costa S, et al. Colorectal Screening Program in Northern Portugal: First Findings. Acta Med Port. 2022;35(3):164-9. doi: 10.20344/amp.15904. PubMed PMID: 34283017.
81. Benchimol EI, Guttmann A, Mack DR, Nguyen GC, Marshall JK, Gregor JC, et al. Validation of international algorithms to identify adults with inflammatory bowel disease in health administrative data from Ontario, Canada. J Clin Epidemiol. 2014;67(8):887-96. doi: 10.1016/j.jclinepi.2014.02.019. PubMed PMID: 24774473.
82. Broe MO, Jensen PB, Mattsson TO, Pottegard A. Validity of Antineoplastic Procedure Codes in the Danish National Patient Registry: The Case of Colorectal Cancer. Epidemiology. 2020;31(4):599-603. doi: 10.1097/EDE.0000000000001208. PubMed PMID: 32483069.
83. Bronson MR, Kapadia NS, Austin AM, Wang Q, Feskanich D, Bynum JPW, et al. Leveraging Linkage of Cohort Studies With Administrative Claims Data to Identify Individuals With Cancer. Med Care. 2018;56(12):e83-e9. doi: 10.1097/MLR.0000000000000875. PubMed PMID: 29334524; PubMed Central PMCID: PMCPMC6043405.
84. Carrell D, Grafton J. Group Health/UW. Colorectal Cancer (CRC). PheKB; 2016 Available from: https://phekb.org/phenotype/514.
85. Cea Soriano L, Soriano-Gabarro M, Garcia Rodriguez LA. Validity and completeness of colorectal cancer diagnoses in a primary care database in the United Kingdom. Pharmacoepidemiol Drug Saf. 2016;25(4):385-91. doi: 10.1002/pds.3877. PubMed PMID: 26436320; PubMed Central PMCID: PMCPMC5054928.

86. Clarke CL, Feigelson HS. Developing an Algorithm to Identify History of Cancer Using Electronic Medical Records. EGEMS (Wash DC). 2016;4(1):1209. doi: 10.13063/2327-9214.1209. PubMed PMID: 27195308; PubMed Central PMCID: PMCPMC4862761.
87. Colov EP, Fransgaard T, Klein M, Gogenur I. Validation of a register-based algorithm for recurrence in rectal cancer. Dan Med J. 2018;65(10). PubMed PMID: 30269749.
88. Cozzolino F, Bidoli E, Abraha I, Fusco M, Giovannini G, Casucci P, et al. Accuracy of colorectal cancer ICD-9-CM codes in Italian administrative healthcare databases: a cross-sectional diagnostic study. BMJ Open. 2018;8(7):e020630. doi: 10.1136/bmjopen-2017-020630. PubMed PMID: 29980543; PubMed Central PMCID: PMCPMC6042611.
89. Dacosta Byfield S, Yu E, Morlock R, Evans D, Teitelbaum A. Corroboration of claims algorithm for second-line costs of metastatic colorectal cancer treatment with targeted agents. J Med Econ. 2013;16(8):1071-81. doi: 10.3111/13696998.2013.813513. PubMed PMID: 23777222.
90. Daniels B, Tervonen HE, Pearson SA. Identifying incident cancer cases in dispensing claims: A validation study using Australia's Repatriation Pharmaceutical Benefits Scheme (PBS) data. Int J Popul Data Sci. 2019;5(1):1152. doi: 10.23889/ijpds.v5i1.1152. PubMed PMID: 32935055.
91. de Luise C, Sugiyama N, Morishima T, Higuchi T, Katayama K, Nakamura S, et al. Validity of claims-based algorithms for selected cancers in Japan: Results from the VALIDATE-J study. Pharmacoepidemiol Drug Saf. 2021;30(9):1153-61. doi: 10.1002/pds.5263. PubMed PMID: 33960542; PubMed Central PMCID: PMCPMC8453514.
92. Denny J, Basford M. Crohn's Disease - Demonstration Project. Vanderbilt University. PheKB; 2012 Available from: https://phekb.org/phenotype/77.
93. Diop M, Strumpf EC, Datta GD. Measuring colorectal cancer incidence: the performance of an algorithm using administrative health data. BMC Med Res Methodol. 2018;18(1):38. doi: 10.1186/s12874-018-0494-x. PubMed PMID: 29739338; PubMed Central PMCID: PMCPMC5941340.
94. Dregan A, Moller H, Murray-Thomas T, Gulliford MC. Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study. Cancer epidemiology. 2012;36(5):425-9. doi: 10.1016/j.canep.2012.05.013. PubMed PMID: 22727737.
95. Earles A, Liu L, Bustamante R, Coke P, Lynch J, Messer K, et al. Structured Approach for Evaluating Strategies for Cancer Ascertainment Using Large-Scale Electronic Health Record Data. JCO Clin Cancer Inform. 2018;2:1-12. doi: 10.1200/CCI.17.00072. PubMed PMID: 30652546; PubMed Central PMCID: PMCPMC6874038.

96. Escriba JM, Banque M, Macia F, Galvez J, Esteban L, Pareja L, et al. Detection of incident breast and colorectal cancer cases from an administrative healthcare database in Catalonia, Spain. Clin Transl Oncol. 2020;22(6):943-52. doi: 10.1007/s12094-019-02219-3. PubMed PMID: 31586294.

97. Fishman E, Barron J, Liu Y, Gautam S, Bekelman JE, Navathe AS, et al. Using claims data to attribute patients with breast, lung, or colorectal cancer to prescribing oncologists. Pragmat Obs Res. 2019;10:15-22. doi: 10.2147/POR.S197252. PubMed PMID: 31015772; PubMed Central PMCID: PMCPMC6446985.

98. Gil M, Rodriguez-Miguel A, Montoya-Catala H, Gonzalez-Gonzalez R, Alvarez-Gutierrez A, Rodriguez-Martin S, et al. Validation study of colorectal cancer diagnosis in the Spanish primary care database, BIFAP. Pharmacoepidemiol Drug Saf. 2019;28(2):209-16. doi: 10.1002/pds.4686. PubMed PMID: 30548462.

99. Goldsbury DE, Armstrong K, Simonella L, Armstrong BK, O'Connell DL. Using administrative health data to describe colorectal and lung cancer care in New South Wales, Australia: a validation study. BMC Health Serv Res. 2012;12:387. doi: 10.1186/1472-6963-12-387. PubMed PMID: 23140341; PubMed Central PMCID: PMCPMC3512511.

100. Goldsbury DE, Weber MF, Canfell K, O'Connell DL. Identifying incident cancer cases in routinely collected hospital data: a retrospective validation study. BMC Res Notes. 2019;12(1):674. doi: 10.1186/s13104-019-4726-x. PubMed PMID: 31639061; PubMed Central PMCID: PMCPMC6805366.

101. Gouverneur A, Dolatkhani D, Rouyer M, Grelaud A, Francis F, Gilleron V, et al. Agreement between hospital discharge diagnosis codes and medical records to identify metastatic colorectal cancer and associated comorbidities in elderly patients. Rev Epidemiol Sante Publique. 2017;65(4):321-5. Epub 2017/06/04. doi: 10.1016/j.respe.2017.03.132. PubMed PMID: 28576381.

102. Greco MT, Russo A. [Comparison of different methods for the definition of indicators to assess the diagnostic and therapeutic paths of colorectal cancer]. Epidemiologia e prevenzione. 2014;38(6):364-72. PubMed PMID: 25651768.

103. Hashibe M, Ou JY, Herget K, Bolton D, McPherson J, Hawley C, et al. Feasibility of Capturing Cancer Treatment Data in the Utah All-Payer Claims Database. JCO Clin Cancer Inform. 2019;3:1-10. doi: 10.1200/CCI.19.00027. PubMed PMID: 31756128; PubMed Central PMCID: PMCPMC6882521.

104. Hassett MJ, Uno H, Cronin AM, Carroll NM, Hornbrook MC, Ritzwoller D. Detecting Lung and Colorectal Cancer Recurrence Using Structured Clinical/Administrative Data to Enable Outcomes Research and Population Health Management. Med Care. 2017;55(12):e88-e98. doi: 10.1097/MLR.0000000000000404. PubMed PMID: 29135771.

105. Hess LM, Li X, Wu Y, Goodloe RJ, Cui ZL. Defining treatment regimens and lines of therapy using real-world data in oncology. Future Oncol. 2021;17(15):1865-77. doi: 10.2217/fon-2020-1041. PubMed PMID: 33629590.

106. Jeon H, You SC, Kang SY, Seo SI, Warner JL, Belenkaya R, et al. Characterizing the Anticancer Treatment Trajectory and Pattern in Patients Receiving Chemotherapy for Cancer Using Harmonized Observational Databases: Retrospective Study. JMIR medical informatics. 2021;9(4):e25035. doi: 10.2196/25035. PubMed PMID: 33720842; PubMed Central PMCID: PMCPMC8058693.

107. Ko CW, Dominitz JA, Neradilek M, Polissar N, Green P, Kreuter W, et al. Determination of colonoscopy indication from administrative claims data. Med Care. 2014;52(4):e21-9. doi: 10.1097/MLR.0b013e31824ebdf5. PubMed PMID: 22437619; PubMed Central PMCID: PMCPMC3387505.

108. Lamarche-Vadel A, Pavillon G, Aouba A, Johansson LA, Meyer L, Jougla E, et al. Automated comparison of last hospital main diagnosis and underlying cause of death ICD10 codes, France, 2008-2009. BMC medical informatics and decision making. 2014;14:44. Epub 2014/06/06. doi: 10.1186/1472-6947-14-44. PubMed PMID: 24898538; PubMed Central PMCID: PMCPMC4057818.

109. Lambert P, Pitz M, Singh H, Decker K. Evaluation of algorithms using administrative health and structured electronic medical record data to determine breast and colorectal cancer recurrence in a Canadian province : Using algorithms to determine breast and colorectal cancer recurrence. BMC cancer. 2021;21(1):763. doi: 10.1186/s12885-021-08526-9. PubMed PMID: 34210266.

110. Li X, Hilsden R, Hossain S, Fleming J, Winget M. Validation of administrative data sources for endoscopy utilization in colorectal cancer diagnosis. BMC Health Serv Res. 2012;12:358. doi: 10.1186/1472-6963-12-358. PubMed PMID: 23062117; PubMed Central PMCID: PMCPMC3508878.

111. Luhn P, Kuk D, Carrigan G, Nussbaum N, Sorg R, Rohrer R, et al. Validation of diagnosis codes to identify side of colon in an electronic health record registry. BMC Med Res Methodol. 2019;19(1):177. doi: 10.1186/s12874-019-0824-7. PubMed PMID: 31426736; PubMed Central PMCID: PMCPMC6700780.

112. Margulis AV, Fortuny J, Kaye JA, Calingaert B, Reynolds M, Plana E, et al. Validation of Cancer Cases Using Primary Care, Cancer Registry, and Hospitalization Data in the United Kingdom. Epidemiology. 2018;29(2):308-13. doi: 10.1097/EDE.0000000000000786. PubMed PMID: 29135571.

113. Meng W, Ou W, Chandwani S, Chen X, Black W, Cai Z. Temporal phenotyping by mining healthcare data to derive lines of therapy for cancer. J Biomed Inform. 2019;100:103335. doi: 10.1016/j.jbi.2019.103335. PubMed PMID: 31689549.

114. Morikubo H, Kobayashi T, Fukuda T, Nagahama T, Hisamatsu T, Hibi T. Development of algorithms for identifying patients with Crohn's disease in the Japanese health insurance claims database. PLoS One. 2021;16(10):e0258537. doi: 10.1371/journal.pone.0258537. PubMed PMID: 34644342.

115. Musselman RP, Rothwell D, Auer RC, Moloo H, Boushey RP, van Walraven C. Can Text-Search Methods of Pathology Reports Accurately Identify Patients with Rectal Cancer in Large Administrative Databases? J Pathol Inform. 2018;9:18. doi: 10.4103/jpi.jpi_71_17. PubMed PMID: 29862128; PubMed Central PMCID: PMCPMC5952547.

116. Nguyen AN, Moore J, O'Dwyer J, Philpot S. Automated Cancer Registry Notifications: Validation of a Medical Text Analytics System for Identifying Patients with Cancer from a State-Wide Pathology Repository. AMIA Annu Symp Proc. 2016;2016:964-73. PubMed PMID: 28269893; PubMed Central PMCID: PMCPMC5333242.

117. Nordstrom BL, Simeone JC, Malley KG, Fraeman KH, Klippel Z, Durst M, et al. Validation of Claims Algorithms for Progression to Metastatic Cancer in Patients with Breast, Non-small Cell Lung, and Colorectal Cancer. Front Oncol. 2016;6:18. doi: 10.3389/fonc.2016.00018. PubMed PMID: 26870695; PubMed Central PMCID: PMCPMC4733924.

118. Nordstrom BL, Whyte JL, Stolar M, Mercaldi C, Kallich JD. Identification of metastatic cancer in claims data. Pharmacoepidemiol Drug Saf. 2012;21 Suppl 2:21-8. doi: 10.1002/pds.3247. PubMed PMID: 22552976.

119. Ritchie MD, Denny JC, Crawford DC, Ramirez AH, Weiner JB, Pulley JM, et al. Robust replication of genotype-phenotype associations across multiple diseases in an electronic medical record. Am J Hum Genet. 2010;86(4):560-72. doi: 10.1016/j.ajhg.2010.03.003. PubMed PMID: 20362271; PubMed Central PMCID: PMCPMC2850440.

120. Ritzwoller DP, Carroll N, Delate T, O'Keeffe-Rossetti M, Fishman PA, Loggers ET, et al. Validation of electronic data on chemotherapy and hormone therapy use in HMOs. Med Care. 2013;51(10):e67-73. doi: 10.1097/MLR.0b013e31824def85. PubMed PMID: 22531648; PubMed Central PMCID: PMCPMC3406224.

121. Song LD, Newhouse JP, Garcia-De-Albeniz X, Hsu J. Changes in screening colonoscopy following Medicare reimbursement and cost-sharing changes. Health Serv Res. 2019;54(4):839-50. doi: 10.1111/1475-6773.13150. PubMed PMID: 30941767; PubMed Central PMCID: PMCPMC6606542.

122. Stavrou E, Pesa N, Pearson SA. Hospital discharge diagnostic and procedure codes for upper gastro-intestinal cancer: how accurate are they? BMC Health Serv Res. 2012;12:331. doi: 10.1186/1472-6963-12-331. PubMed PMID: 22995224; PubMed Central PMCID: PMCPMC3506480.

123. Strongman H, Williams R, Bhaskaran K. What are the implications of using individual and combined sources of routinely collected data to identify and characterise incident site-specific cancers? a concordance and validation study using linked English electronic health records data. BMJ Open. 2020;10(8):e037719. doi: 10.1136/bmjopen-2020-037719. PubMed PMID: 32819994.

124. Thompson CA, Jin A, Luft HS, Lichtensztajn DY, Allen L, Liang SY, et al. Population-Based Registry Linkages to Improve Validity of Electronic Health Record-Based Cancer Research. Cancer Epidemiol Biomarkers Prev. 2020;29(4):796-806. doi: 10.1158/1055-9965.EPI-19-0882. PubMed PMID: 32066621.

125. Vachani A, Wong YN, Israelite J, Mitra N, Hin S, Yang L, et al. Validation of Molecular Pathology Codes for the Identification of Mutational Testing in Lung and Colon Cancer. Med Care. 2017;55(12):e131-e6. doi: 10.1097/MLR.0000000000000484. PubMed PMID: 29135776.

126. Vesteghem C, Brondum RF, Falkmer UG, Pottegard A, Poulsen LO, Bogsted M. High Validity of the Danish National Patient Registry for Systemic Anticancer Treatment Registration from 2009 to 2019. Clinical epidemiology. 2021;13:1085-94. doi: 10.2147/CLEP.S332776. PubMed PMID: 34853537.

127. Whyte JL, Engel-Nitz NM, Teitelbaum A, Gomez Rey G, Kallich JD. An Evaluation of Algorithms for Identifying Metastatic Breast, Lung, or Colorectal Cancer in Administrative Claims Data. Med Care. 2015;53(7):e49-57. doi: 10.1097/MLR.0b013e318289c3fb. PubMed PMID: 23524464.

128. Ye Y, Manne S, Bennett D. Identifying Patients With Inflammatory Bowel Diseases in an Administrative Health Claims Database: Do Algorithms Generate Similar Findings? Inquiry. 2019;56:46958019887816. doi: 10.1177/0046958019887816. PubMed PMID: 31744340.

129. HDH. Documentation du SNDS & SNDS OMOP. Healt-Data-Hub (HDH). https://documentation-snds.health-data-hub.fr/snds/formation_snds/initiation/schema_relationnel_snds.html#_3-3-dcir consulté le 27/10/2023.

130. Tuppin P, Rudant J, Constantinou P, Gastaldi-Menager C, Rachas A, de Roquefeuil L, et al. Value of a national administrative database to guide public decisions: From the systeme national d'information interregimes de l'Assurance Maladie (SNIIRAM) to the systeme national des donnees de sante (SNDS) in France. Rev Epidemiol Sante Publique. 2017;65 Suppl 4:S149-S67. doi: 10.1016/j.respe.2017.05.004. PubMed PMID: 28756037.

131. Aguade AS, Gastaldi-Menager C, Denis P, Fagot-Campagna A, Gissot C, Polton D. How to Estimate The Cost Of Diabetes Based on Information From The French Health Insurance Database (Sniiram)? Value Health. 2014;17(7):A341. doi: 10.1016/j.jval.2014.08.675. PubMed PMID: 27200625.

132. Gardy J, Wilson S, Guizard AV, Bouvier V, Tron L, Launay L, et al. Cancer Survival and Travel Time to Nearest Reference Care Center for 10 Cancer Sites: An Analysis of 21 French Cancer Registries. Cancers (Basel). 2023;15(5). doi: 10.3390/cancers15051516. PubMed PMID: 36900308.

133. Jooste V, Remontet L, Colonna M, Belot A, Launoy G, Binder F, et al. Trends in the incidence of digestive cancers in France between 1980 and 2005 and projections for the year 2010. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2011;20(5):375-80. doi: 10.1097/CEJ.0b013e3283465426. PubMed PMID: 21478752.

134. Lepage C, Remontet L, Launoy G, Tretarre B, Grosclaude P, Colonna M, et al. Trends in incidence of digestive cancers in France. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2008;17(1):13-7. doi: 10.1097/CEJ.0b013e32809b4cba. PubMed PMID: 18090905.

135. Valori R, Rey JF, Atkin WS, Bretthauer M, Senore C, Hoff G, et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition--Quality assurance in endoscopy in colorectal cancer screening and diagnosis. Endoscopy. 2012;44 Suppl 3:SE88-105. doi: 10.1055/s-0032-1309795. PubMed PMID: 23012124.

136. Force USPST, Davidson KW, Barry MJ, Mangione CM, Cabana M, Caughey AB, et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA. 2021;325(19):1965-77. doi: 10.1001/jama.2021.6238. PubMed PMID: 34003218.

137. Schreuders EH, Ruco A, Rabeneck L, Schoen RE, Sung JJ, Young GP, et al. Colorectal cancer screening: a global overview of existing programmes. Gut. 2015;64(10):1637-49. doi: 10.1136/gutjnl-2014-309086. PubMed PMID: 26041752.

138. Vale DB, Anttila A, Ponti A, Senore C, Sankaranaryanan R, Ronco G, et al. Invitation strategies and coverage in the population-based cancer screening programmes in the European Union. European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP). 2019;28(2):131-40. doi: 10.1097/CEJ.0000000000000426. PubMed PMID: 29570103.

139. Chapman M, Mumtaz S, Rasmussen LV, Karwath A, Gkoutos GV, Gao C, et al. Desiderata for the development of next-generation electronic health record phenotype libraries. Gigascience. 2021;10(9). doi: 10.1093/gigascience/giab059. PubMed PMID: 34508578.

140. Ananthakrishnan AN, Cai T, Savova G, Cheng SC, Chen P, Perez RG, et al. Improving case definition of Crohn's disease and ulcerative colitis in electronic medical records using natural language processing: a

novel informatics approach. Inflamm Bowel Dis. 2013;19(7):1411-20. doi: 10.1097/MIB.0b013e31828133fd. PubMed PMID: 23567779.

141. Chawla N, Yabroff KR, Mariotto A, McNeel TS, Schrag D, Warren JL. Limited validity of diagnosis codes in Medicare claims for identifying cancer metastases and inferring stage. Ann Epidemiol. 2014;24(9):666-72, 72 e1-2. doi: 10.1016/j.annepidem.2014.06.099. PubMed PMID: 25066409; PubMed Central PMCID: PMCPMC4135357.

142. Fujiwara T, Kanemitsu T, Tajima K, Yuri A, Iwasaku M, Okumura Y, et al. Accuracy of algorithms to identify patients with a diagnosis of major cancers and cancer-related adverse events in an administrative database: a validation study in an acute care hospital in Japan. BMJ Open. 2022;12(7):e055459. doi: 10.1136/bmjopen-2021-055459. PubMed PMID: 35831049.

143. Hou JK, Tan M, Stidham RW, Colozzi J, Adams D, El-Serag H, et al. Accuracy of diagnostic codes for identifying patients with ulcerative colitis and Crohn's disease in the Veterans Affairs Health Care System. Dig Dis Sci. 2014;59(10):2406-10. doi: 10.1007/s10620-014-3174-7. PubMed PMID: 24817338.

144. Hwang YJ, Kim N, Yun CY, Yoon H, Shin CM, Park YS, et al. Validation of Administrative Big Database for Colorectal Cancer Searched by International Classification of Disease 10th Codes in Korean: A Retrospective Big-cohort Study. J Cancer Prev. 2018;23(4):183-90. doi: 10.15430/JCP.2018.23.4.183. PubMed PMID: 30671401.

145. Ihira H, Goto A, Yamagishi K, Iso H, Iwasaki M, Sawada N, et al. Validity of claims data for identifying cancer incidence in the Japan public health center-based prospective study for the next generation. Pharmacoepidemiol Drug Saf. 2022;31(9):972-82. doi: 10.1002/pds.5494. PMID: 35726806.

146. Lamont EB, Lan L. Sensitivity of Medicare claims data for measuring use of standard multiagent chemotherapy regimens. Med Care. 2014;52(3):e15-20. doi: 10.1097/MLR.0b013e31824e342f. PMID: 22410411.

147. Lee CK, Ha HJ, Oh SJ, Kim JW, Lee JK, Kim HS, et al. Nationwide validation study of diagnostic algorithms for inflammatory bowel disease in Korean National Health Insurance Service database. J Gastroenterol Hepatol. 2020;35(5):760-8. doi: 10.1111/jgh.14855. PubMed PMID: 31498502.

148. Lo B, Vind I, Vester-Andersen MK, Burisch J. Validation of ulcerative colitis and Crohn's disease and their phenotypes in the Danish National Patient Registry using a population-based cohort. Scand J Gastroenterol. 2020;55(10):1171-5. doi: 10.1080/00365521.2020.1807598. PubMed PMID: 32838593.

149. Ogino H, Morikubo H, Fukaura K, Okui T, Gardiner S, Sugiyama N, et al. Validation of a claims-based algorithm to identify cases of ulcerative colitis in Japan. J Gastroenterol Hepatol. 2022;37(3):499-506. doi: 10.1111/jgh.15732. PubMed PMID: 34738649; PubMed Central PMCID: PMCPMC9298722.

150. Quantin C, Benzenine E, Hagi M, Auverlot B, Abrahamowicz M, Cottenet J, et al. Estimation of national colorectal-cancer incidence using claims databases. Journal of cancer epidemiology. 2012;2012:298369. doi: 10.1155/2012/298369. PubMed PMID: 22792103.

151. Quantin C, Benzenine E, Hagi M, Auverlot B, Abrahamowicz M, Cottenet J, et al. [Evaluation of medical information systems as a mean of identification of incident cases of colorectal cancer]. Sante publique (Vandoeuvre-les-Nancy, France). 2014;26(1):55-63. PubMed PMID: 24893515.

152. Rezaie A, Quan H, Fedorak RN, Panaccione R, Hilsden RJ. Development and validation of an administrative case definition for inflammatory bowel diseases. Can J Gastroenterol. 2012;26(10):711-7. doi: 10.1155/2012/278495. PubMed PMID: 23061064; PubMed Central PMCID: PMCPMC3472911.

153. Rye C, Rubin KH, Moller FT, Julsgaard M, Jess T, Andersen V. Positive Predictive Value of Diagnostic Codes for Inflammatory Bowel Disease in the Danish National Patient Registry Among Individuals 50+ Years, Using Patient Records as Reference Standard. Clinical epidemiology. 2021;13:335-44. doi: 10.2147/CLEP.S298770. PubMed PMID: 34079376.

154. Takahashi S, Obara T, Kakuta Y, Shimoyama Y, Naito T, Moroi R, et al. Validity of Diagnostic Algorithms for Inflammatory Bowel Disease in Japanese Hospital Claims Data. Int J Environ Res Public Health. 2022;19(13). doi: 10.3390/ijerph19137933. PubMed PMID: 35805591; PubMed Central PMCID: PMCPMC9266263.

155. Thirumurthi S, Chowdhury R, Richardson P, Abraham NS. Validation of ICD-9-CM diagnostic codes for inflammatory bowel disease among veterans. Dig Dis Sci. 2010;55(9):2592-8. doi: 10.1007/s10620-009-1074-z. PubMed PMID: 20033847.

156. Melac AT, Lesuffleur T, Bousquet PJ, Fagot-Campagna A, Gastaldi-Menager C, Tuppin P. Cancer and end of life: the management provided during the year and the month preceding death in 2015 and causes of death in France. Support Care Cancer. 2020;28(8):3877-87. doi: 10.1007/s00520-019-05188-z. PMID: 31845006.

157. Bousquet PJ, Lefeuvre D, Tuppin P, BenDiane MK, Rocchi M, Bouée-Benhamiche E, et al. Cancer care and public health policy evaluations in France: Usefulness of the national cancer cohort. PLoS One. 2018;13(10):e0206448. Epub 2018/11/01. doi: 10.1371/journal.pone.0206448. PubMed PMID: 30379946.

158. Laanani M, Coste J, Blotière PO, Carbonnel F, Weill A. Patient, Procedure, and Endoscopist Risk Factors for Perforation, Bleeding, and Splenic Injury After Colonoscopies. Clin Gastroenterol Hepatol. 2019;17(4):719-27.e13. Epub 2018/08/14. doi: 10.1016/j.cgh.2018.08.005. PubMed PMID: 30099110.

159. Thomassin-Naggara I, Balleyguier C, Ceugnart L, Heid P, Lenczner G, Maire A, et al. Artificial intelligence and breast screening: French Radiology Community position paper. Diagn Interv Imaging. 2019;100(10):553-66. doi: 10.1016/j.diii.2019.08.005. PubMed PMID: 31523026.

160. Wang A, Lee B, Patel S, Whitaker E, Issaka RB, Somsouk M. Selection of patients for large mailed fecal immunochemical test colorectal cancer screening outreach programs: A systematic review. Journal of medical screening. 2021;28(4):379-88. doi: 10.1177/0969141321997482. PubMed PMID: 33683155.

161. Sequist TD, Zaslavsky AM, Colditz GA, Ayanian JZ. Electronic patient messages to promote colorectal cancer screening: a randomized controlled trial. Arch Intern Med. 2011;171(7):636-41. doi: 10.1001/archinternmed.2010.467. PubMed PMID: 21149743.

162. Muller D, Logan J, Dorr D, Mosen D. The effectiveness of a secure email reminder system for colorectal cancer screening. AMIA Annu Symp Proc. 2009;2009:457-61. PubMed PMID: 20351899.

163. Seitz JF, Lapalus D, Arlotto S, Gentile S, Ettori F, Rinaldi Y, et al. Colorectal cancer screening by fecal immunochemical test or colonoscopy in France: how many people are actually covered? Focus on the Provence-Alpes-Cote d'Azur region. European journal of gastroenterology & hepatology. 2022;34(4):405-10. doi: 10.1097/MEG.0000000000002338. PubMed PMID: 34882643.

164. Benamouzig R, Uzzan B. Aspirin to prevent colorectal cancer: time to act? Lancet (London, England). 2010;376(9754):1713-4. doi: 10.1016/S0140-6736(10)61509-7. PubMed PMID: 20970845.

165. Delluc C, Benamouzig R. [Hormone replacement therapy and colorectal cancer]. Gastroenterol Clin Biol. 2008;32(1 Pt. 1):19-28. doi: 10.1016/j.gcb.2007.12.008. PubMed PMID: 18341973.

166. Kaufmanis A, Vincelet C, Koivogui A, Delattre-Massy H, Deyra J, Ait Hadad H, et al. Devenir en deuxième campagne de dépistage organisé du cancer colorectal d'un test négatif en première campagne selon le dosage d'hémoglobine fécale. In: (SNFGE) editor. Journées Francophones d'Hépato-Gastroentérolgie et d'Oncologie Digestive (JFHOD); MARCH 2018; Paris. Paris, France2019. p. P.395.

167. Robaszkiewicz M. Intérêt des scores pour la prédiction du risque de cancer colorectal. In: METIER LG, editor. Dépistage et prévention des cancers colorectaux. ROUEN, FRANCE2018. p. 99-101.