NOVARTIS

Quantitative Safety and Epidemiology

Non-Interventional Final Report

SPP100A2417

A multi-database cohort study to assess the incidence rates of colorectal hyperplasia among hypertensive patients

REDACTED STUDY REPORT

Author

Document Status Final

Date of final version of the study report 14 December 2015

NIS Report Template Version 2.0 August-13-2014

PASS information	
Title	A multi-database cohort study to assess the incidence rates of colorectal hyperplasia among hypertensive patients
Version identifier of the final study report	V1
Date of last version of the final study report	14 December 2015
EU PAS register number	8295
Active substance	Aliskiren
Medicinal product	Rasilez®
Product reference	EU/1/08/491/001-080
Procedure number	EMEA/H/C/000964
Marketing authorization holder	Novartis Pharma
Joint PASS	No
Research question and objectives	The primary objective of this study was to assess the age- and sex-specific incidence rates of colorectal polyps, colorectal cysts, benign neoplasms and cancer among patients diagnosed with hypertension. The secondary objective was to assess the time from hypertension diagnosis to colorectal polyps, colorectal cysts, benign neoplasms and cancer diagnosis among patients diagnosed with hypertension.
Country(-ies) of study	The Netherlands, United Kingdom, Italy, Spain
Author	[redacted]

Marketing authorization holder

Marketing authorization holder(s)	Novartis Europharm Limited Frimley Business Park Camberley GU16 7SR United Kingdom
MAH contact person	

Table of contents

List of tables

List of figures

1 Abstract

Title

A multi-database cohort study to assess the incidence rates of colorectal hyperplasia among hypertensive patients

Keywords

Hypertension; incidence; colorectal cancer; polyps; risk

Rationale and background

Epidemiologic studies examining the association between hypertension and colorectal cancer show inconsistent results.

Research question and objectives

The primary objective of this study was to assess the age- and sex-specific incidence rates of colorectal polyps, colorectal cysts, benign neoplasms and cancer among patients diagnosed with hypertension.

The secondary objective was to assess the time from hypertension diagnosis to colorectal polyps, colorectal cysts, benign neoplasms and cancer diagnosis among patients diagnosed with hypertension.

Study design

Multi-database dynamic cohort study, in four European primary care databases. Study period: 01 January 2000 – 31 December 2013.

Setting

Study population: All adult patients (aged 18-79 years) with a diagnosis of arterial hypertension and continuous enrollment in the corresponding database for at least 1 year prior to the start of follow-up. One primary and two secondary study cohorts were defined:

1. Primary
    - Incident hypertension subjects (excluding subjects with prevalent hypertension).
2. Secondary
    - Incident and prevalent hypertension subjects
    - Incident hypertension subjects, who have had a colonoscopy that was negative for colorectal polyps, cysts and neoplasms.

Exclusion criteria: history of cancer (excluding non-melanoma skin cancer); 80 years of age or older at start of cohort entry and less than 30 days of follow-up.

Follow-up: From first-time recording of arterial hypertension (cohort entry) until earliest date of diagnosis of colorectal hyperplasia; end of enrollment in database; death; reaching age of 80 years; or end of study period (whatever comes first).

Subjects and study size, including dropouts

Table 1-1 presents the study size of the primary and secondary populations.

Table 1-1 Study size of primary and secondary analysis populations across data sources

	CPRD(UK)	HSD(IT)	IPCI(NL)	SIDIAP(ES)
Total source population	13,673,353	1,738,798	1,708,253	6,332,740
Population with 1 year of continuous data	10,924,766	1,627,136	1,672,154	6,175,913
Aged 18-79 years	8,517,175	1,501,405	1,327,309	4,963,766
Subjects with Hypertension Dx	1,353,403	376,760	225,038	1,006,441
Cancer-free at cohort entry	1,192,143	344,532	203,027	931,582
Less than 30 days of follow-up	6756	2201	869	6126
Cohort 2. Incident+Prevalent hypertension	1,185,387	342,331	202,158	925,456
Cohort 1. Incident hypertension	688,674	237,922	40,568	336,456
Cohort 3. Incident hypertension and negative colonoscopy	12,956	6,896	3,112	7,154

CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

Variables and data sources

Variables: Age, sex at index date, use of low-dose aspirin and antihypertensive drug exposure before cohort entry. The following variables were considered present if there was a code before or at cohort entry: inflammatory bowel disease, obesity, hereditary colon cancer syndromes, familial colon cancer syndromes, hyperlipidemia, hypertriglyceridemia, hypercholesterolemia, type 2 diabetes mellitus (T2DM), ischemic heart disease, angina pectoris, stroke/transient ischemic attack (TIA), smoking, alcohol use, body mass index (BMI), colonoscopy, irritable bowel syndrome (IBS), and cholecystectomy.

Data sources:

-The Clinical Practice Research Datalink (CPRD; formerly known as General Practice Research database [GPRD]) from the United Kingdom (UK)

-The Health Search/CSD Longitudinal Patient (HSD) database from Italy (IT)

-The Integrated Primary Care Information (IPCI) database from the Netherlands (NL)

- The Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP) from Spain (ES)

Results

Incidence rates of colorectal outcomes (polyps, cysts, benign neoplasm and cancer) were consistent in the four European data sources (Table 1-2). Median follow-up time ranged from 1.9 to 6 years.

Table 1-2 Incident rates of colorectal outcomes in the four European data sources

	CPRD (UK) IR/100,000 PY	HSD (IT) IR/100,000 PY	IPCI(NL) IR/100,000 PY	SIDIAP (ES) IR/100,000 PY
Incident Cohort				
Composite Event	348.2	517.5	510.3	628.2
Polyp	241.8	380.2	426.4	484.3
Cyst	6.4	2.9	0	2.3
Benign neoplasm	4.2	0	1.1	10.1
Colorectal Cancer	124.3	147.1	107.2	141.8
Incident+Prevalent Cohort				
Composite Event	343.4	506.8	567.7	589.6
Polyp	227.2	366.4	472.8	423.3
Cyst	8.8	3.0	0	4.3
Benign neoplasm	4.8	0	0.7	15.3
Colorectal Cancer	131.9	149.4	130.2	158.1
Incident Colonoscopy Cohort				
Composite Event	758.1	1016	1391.8	2223
Polyp	617.1	777.3	1306.1	2043
Cyst	8.1	12.1	0	0
Benign neoplasm	8.1	0	0	28.7
Colorectal Cancer	162	136.5	115.1	143.6

CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

Incidence rates of the outcomes were higher with increasing age, in males, with presence of comorbid diseases (such as obesity, type 2 diabetes mellitus, ischemic heart disease, IBD, IBS, hyperlipidemia); and in patients with prior use of drugs before cohort entry (any anti-hypertensive drug low-dose aspirin).

The one-year risk of the composite endpoints varied between 0.27% and 0.65%. The cumulative 5-year risk of the composite colorectal outcomes varied between 1.5% and 3% across data sources, and was higher with older ages, presence of IBD, for males compared to females, a prior colonoscopy positive for any of the study outcomes and people already using anti-hypertensive drugs (for other reasons such as β-blockers, diuretics or ACE inhibitors for heart failure, post-myocardial infarction, high coronary disease risk, diabetes, chronic renal disease, or stroke prevention).

Particularly in the first month and up to 3 months following hypertension diagnosis the incidence rates of outcomes were high, which could be explained with increased chances for diagnoses due to increased work up when someone is diagnosed with hypertension (Berksonian bias).In some databases this was stronger than in others.

Discussion

The higher incidences of the study outcomes in the Incident cohort with presence of comorbid diseases, is in line with results from the literature. The incidence of the study outcomes was particularly high in the first 3 months after hypertension diagnosis, suggesting the presence of Berksonian bias. The incidence of Colorectal Cancer (CRC) was consistent across the databases, and was even more similar across the data sources when standardizing to a common reference population. The increase in colorectal polyp and decrease in CRC incidence since 2005 in CPRD and HSD may be due to the implementation of CRC screening programs in the United Kingdom and Italy. Since patients were censored at the age of 80 years, the results may thus not be generalizable for hypertensive subjects aged 80 years or older.

Marketing Authorization Holder

Novartis Europharm Limited
Frimley Business Park
Camberley GU16 7SR
United Kingdom

Name(s) and Affiliation(s) of Principal Investigator(s)

[redacted]

2 List of abbreviations

BMI	Body Mass Index
CHMP	Committee for Medicinal Products for Human Use
CPRD	Clinical Practice Research Datalink
CRC	Colorectal Cancer
CRO	Contract Research Organization
DRI	Direct Renin Inhibitor
[redacted]	[redacted]
EMA	European Medicines Agency
ENCePP	European Network of Centres for Pharmacoepidemiology and Pharmacovigilance
EU	European Union
FAP	Familial adenomatous polyposis
FIT	Fecal Immunochemical Testing
GPP	Good Pharmacoepidemiology Practices
HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HSD	Health Search/CSD Longitudinal Patient
IBS	Irritable Bowel Syndrome
ICD	International Classification of Diseases
ICPC	International Classification for Primary Care
IPCI	Integrated Primary Care Information
IR	Incidence Rate
ISPE	International Society for Pharmacoepidemiology
MedDRA	Medical Dictionary for Regulatory Activities
NIS	Non-interventional Study
PASS	Post-Authorization Safety Study
PI	Principal Investigator
RMP	Risk Management Plan
RRE	Remote Research Environment
SIDIAP	Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària
STROBE	Strengthening the Reporting of Observational Studies in Epidemiology
T2DM	Type 2 Diabetes Mellitus
TIA	Transient Ischemic Attack
UK	United Kingdom
WHO	World Health Organization

3 Investigators

Department	Database	Name
[redacted]		

4 Other responsible parties

Not applicable.

5 Milestones

Table 5-1 Study milestones

Milestone	Planned date	Actual date	Comments
Registration in the EU PAS register	December 2014	27 December 2014	N/A
Protocol approval EMA	January 2015	9 January 2015	NA
Start of data collection	May 2015	1 May 2015	NA
End of data collection	June 2015	18th September 2015	Delayed due to SIDIAP re-run
Final report of study results	January 2015	14 December 2015	NA

6 Rationale and background

6.1 Purpose and rationale

In the context of the Rasilez Follow-up Measure (FUM) 025 regarding the carcinogenic potential of aliskiren (and of aliskiren in combination with hydrochlorothiazide [FUM 026]), the Committee for Medicinal Products for Human Use (CHMP) requested Novartis to perform a non-interventional study (NIS) assessing age- and sex-stratified incidence rates of colorectal hyperplasia in a hypertensive population to get background incidence rate data in this population. Based on that request, Novartis proposed a NIS with secondary use of data derived from various European healthcare data sources, namely the Clinical Practice Research Datalink (CPRD) from the United Kingdom (UK), the Health Search/CSD Longitudinal Patient (HSD) database from Italy, the Integrated Primary Care Information (IPCI) database from the Netherlands, and the Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP) from Spain.

6.2 Background

Aliskiren (Rasilez®) is the first orally active direct renin inhibitor (DRI) approved for the treatment of hypertension. Aliskiren's inhibitory effect on angiotensin I generation, through renin blockade, is highly specific and long-lasting (24 hours) (Fisher et al 2008). Its antihypertensive effect is similar to that determined for angiotensin converting enzyme inhibitors, angiotensin receptor blockers and thiazides (Angeli et al 2012, Musini et al 2008, Schmieder et al 2009).

Colorectal hyperplasia is listed as a potential risk in the aliskiren risk management plan (RMP), which is based upon pre-clinical findings in a rodent carcinogenicity study. However, these findings have not been confirmed in a 2-year marmoset study or in targeted clinical studies including a colonoscopy study. Therefore, the increased risk of colorectal hyperplasia observed in the rodent study may reflect high intraluminal drug concentrations in rats or could be a species-specific difference between rats and humans in response to aliskiren exposure.

Colorectal cancer (CRC) is a common type of cancer, being the third most commonly diagnosed cancer in males and the second in females. The incidence of CRC varies globally over 10-fold, mainly due to differences in dietary and environmental factors. According to data from various cancer registries worldwide, age-standardized CRC incidence rates may range from approximately 4.1 to 59.1 per 100,000 per year among males and 3.6 to 39.5 per 100,000 per year among females, respectively (Center et al 2009).

Apart from genetic susceptibility, other risk factors of CRC are inflammatory bowel disease (IBD) (Eaden et al 2001, Herrinton et al 2012), alcohol use, obesity and diabetes mellitus (Ahmed et al 2006, Esposito et al 2012).

Epidemiologic studies examining a potential association of hypertension with the development of CRC have reported inconsistent findings. Various published observational studies did not find an increased risk of CRC in patients with hypertension compared to normotensive patients (Aleksandrova et al 2011, Azoulay et al 2012, Kim et al 2007, Lever et al 1999, Lindgren et al 2005, Lindholm et al 2001, Negri et al 1999, Sturmer et al 2006). However, there is also evidence from other observational studies that hypertension might be

associated with an increased risk of CRC (Batty et al 2003, Othman and Zin 2008, Pelucchi et al 2010, Stocks et al 2008), especially in hypertensive patients with type 2 diabetes mellitus (T2DM) and obesity (Stocks et al 2008), or with some components of the metabolic syndrome (Kim et al 2007). Additionally, there is some evidence that the mortality due to CRC may be higher in hypertensive compared to normotensive patients (Batty et al 2003, Watanabe et al 2005). Again, the data on this topic are inconsistent, as a review including 10 longitudinal studies could only demonstrate an increased risk of mortality from cancer in general, but not of CRC specifically (Grossman et al 2002).

It is currently difficult to establish to what extent a potential increased risk of CRC during antihypertensive treatment with aliskiren may be related to the drug, as the underlying background risk of CRC in hypertensive patients is unknown. Specific data on the incidence rate of CRC or more general on colorectal hyperplasia in a population of hypertensive patients are not available in the published literature. This non-interventional study was therefore planned – as requested by CHMP – to provide additional details on the background incidence of colorectal hyperplasia in hypertensive patients in general. An additional non-interventional study based on US health claims data (Study SPP100A2418) assessed the risk of colorectal hyperplasia specifically in association with exposure to aliskiren and with other antihypertensive drugs. The results suggest that for GI cancer there is no statistical difference between aliskiren and other antihypertensives (HR 1.03; 95% CI 0.91-1.16) or between aliskiren and non-hypertensives (HR 1.10; 95% CI 0.94-1.28). Other antihypertensive therapy users had a significantly greater risk of GI cancer compared to non-hypertensive patients (HR: 1.58, 95% CI: [1.53, 1.62]). The HR for colorectal hyperplasia, when comparing aliskiren with other antihypertensives was 1.08 (95% CI 1.05-1.11); the HR of the aliskiren cohort compared with the non-hypertensive cohort was 1.12 (95% CI 1.09-1.16). Other antihypertensive therapy users had a significantly greater risk of GI cancer compared to non-hypertensive patients (HR: 1.42, 95% CI: [1.41, 1.43]). The risk of colorectal hyperplasia was higher in the aliskiren cohort as well as in the antihypertensive cohort, when compared to the non-hypertensive cohort. These results should be interpreted with caution considering the limitations of the study, which in summary are: patients aged 65 years and older were not equally represented in all cohorts (2% of the non-hypertensive patients were 65 years and older); the possibility of channeling bias and/or residual confounding by indication given that treatment is often decided on the basis of the severity of hypertension. Research question and objectives

The primary objective of this non-interventional study is to assess age- and sex-stratified incidence rates of colorectal hyperplasia among patients with diagnosed hypertension based on information from various European primary care data sources.

The secondary objective is to assess the hazard and time from hypertension diagnosis until colorectal hyperplasia diagnosis/occurrence. This includes estimating the one-year risk of colorectal neoplasms after hypertension diagnosis, stratified by sex.

7 Amendments and updates to the protocol

Number	Date	Section of study protocol	Amendment or update	Reason
1	14 Nov 2014	9.7 Data Analysis	Amendment	CHMP requested to assess the impact of excluding complicated hypertension persons from the cohort. A sensitivity analysis was performed by including prevalent hypertension patients.
1	26 March 2015	Annex 8 in SAP	Amendment	Case validation not in CPRD

8 Research methods

8.1 Study design

A dynamic cohort study in hypertensive patients was performed with secondary use of data derived from multiple databases.

Primary care databases (electronic medical records) were used to identify patients with an incident diagnosis of hypertension, which happens often in primary care. The databases included in the study (see Section 9.2 for details) contain information on demographics, diagnoses, and drug prescriptions. The clinical information captured by the databases is by different disease coding systems, such as the International Classification of Diseases (ICD) 9th or 10th revision (WHO),(Rahme et al 2004) International Classification for Primary Care (ICPC) (Lamberts et al 1992), or READ (Chisholm 1990).

Harmonization of corresponding disease codes of the primary outcome was performed across coding systems (ICD-9th revision, ICD-10th revision, ICPC-, READ-coding system). A code list used for identification of colorectal hyperplasia can be found in Annex 1.

8.2 Setting

The following four European general practice databases were used:

1. The Clinical Practice Research Datalink (CPRD; formerly known as General Practice Research Database [GPRD]) from the United Kingdom (Jick et al 1991, Jick et al 2003)
2. The Health Search/CSD Longitudinal Patient (HSD) database from Italy (Filippi et al 2005)
3. The Integrated Primary Care Information (IPCI) database from the Netherlands (Vlug et al 1999)
4. The Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP) from Spain (Bolibar et al 2012)

Details on the individual databases are given in the subsections below. The databases have been selected based on their geographic location, the availability of population and primary care data based information plus their recognized reputation in the area of drug utilization and

safety research. Multiple countries are included in order to provide international data. All of the participating databases are part of the "EU-ADR Alliance", a stable collaboration framework for running drug safety studies in a federated manner, especially when the participation of several electronic health care record databases is required (Coloma et al 2011).

All of the chosen databases comply with European Union (EU) guidelines on the use of medical data for medical research and have been validated for pharmaco-epidemiological research. Concerning the validity of a hypertension diagnosis in the participating databases; several pharmaco-epidemiological studies have been performed combining data from IPCI and HSD. This includes a study on the prevalence and treatment of hypertensive patients (Sturkenboom et al 2008). A study on incident hypertensive patients by physician diagnosis has been performed in CPRD (Burke et al 2006). For SIDIAP, there has been a recent study on the validity of cardiovascular diseases and their risk factors in SIDIAP database (Ramos et al 2012). Colorectal cancer has been studied as outcome in CPRD (Garcia-Rodriguez and Huerta-Alvarez 2001, Hong et al 2013, van Staa et al 2005, Yang et al 2004), and IPCI (van Soest et al 2008), but not yet in HSD and SIDIAP.

In some studies (Garcia-Rodriguez and Huerta-Alvarez 2001, van Soest et al 2008) extensive case validation was performed to validate the diagnostic codes for CRC. In CPRD this yielded evidence that the validity of CRC diagnoses is high: over 95% of computer-recorded incident CRC diagnoses were validated by review of original medical records or by confirmation of the diagnosis by the GP (Garcia-Rodriguez and Huerta-Alvarez, 2001)). In IPCI, the positive predictive value of the ICPC colorectal cancer code was 90% (van Soest et al 2008). Other subcategories (polyps, benign neoplasms and cysts) have not been validated.

Concomitant or prior medications entered into the database were identified on the basis of the World Health Organization (WHO) Drug Reference List. Medical history/current medical conditions and adverse events were identified on initial basis of the Medical dictionary for regulatory activities (MedDRA) terminology but were not recoded or recorded as such.

Drugs were coded according to the Anatomical Chemical Therapeutic system in IPCI, HSD and SIDIAP, and according to BNF/multilex codes in CPRD.

8.2.1 Clinical Practice Research Datalink

The CPRD from the UK collates the computerized medical records of general practitioners (GPs) in the UK who act as the gatekeepers of healthcare and maintain patients' life-long electronic health records. As such they are responsible for primary healthcare. The data recorded in the CPRD include demographic information, prescription details, clinical events, preventive care, specialist referrals, hospital admissions, and major outcomes, including death. The majority of the data are coded in READ Codes (Booth, 1994); however, additional text is also available, which can improve the sensitivity and specificity of data. Validation of data with original records (specialist letters) is also available.

Importantly, CPRD operates a careful and continual quality control procedure that ensures that only practices that are "up-to-standard" (UPS) are included in the research dataset. The dataset is generalizable to the UK population based upon age, sex, socioeconomic class and national geographic coverage.

There are currently approximately 13.2 million patients (acceptable for research purposes) – of which 5.7 million are active (still alive and registered with the GP practice) – in approximately 680 practices. Data include demographics, all GP/healthcare professional consultations (phone, letter, email, in surgery, at home), diagnoses and symptoms, laboratory test results, treatments, including all prescriptions, all data referrals to other care, hospital discharge summary (date and ICD-9 code), hospital clinic summary, preventive treatment and immunizations, death (date and cause). Drug and event coding dictionaries are available upon request.

CPRD is listed under the European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP) resources database.

(www.encepp.eu/encepp/resourcesDatabase.jsp).

8.2.2 Health Search/CSD Longitudinal Patient

The Italian arm of the study used the Health Search CSD Longitudinal Patient Database (HSD), a longitudinal observational database that is representative of the Italian general population. It was established in 1998 by the Italian College of General Practitioners (Filippi et al 2005). The HSD contains data from computer-based patient records from a selected group of GPs covering a total of 1.5 million patients located throughout Italy. These GPs voluntarily agreed to collect data for the database and attend specified training courses. The database includes information on the age, sex, and identification of the patient, and GP registration information, which is linked to prescription information, clinical events and diagnoses, hospital admission, and causes of death. All diagnoses are coded according to ICD-9-CM. Drug names are coded according to the ATC classification system. To be included in the study, GPs must have provided data for at least 1 year and meet standard quality criteria pertaining to: levels of coding, prevalence of well-known diseases, and mortality rates (Cricelli et al 2003). The HSD complies with EU guidelines on the use of medical data for research. The HSD has been used as data source for a number of peer-reviewed publications on the prevalence of disease conditions, drug safety and prescription patterns in Italian primary care (Cazzola et al 2011). Approval for use of data is obtained from the Italian College of General Practitioners.

HSD is listed under the ENCePP resources database. (www.encepp.eu/encepp/resourcesDatabase.jsp).

8.2.3 Integrated Primary Care Information database

[redacted]. IPCI is a longitudinal observational database that contains data from computer-based patient records of a selected group of GPs throughout The Netherlands, who voluntarily chose to supply data to the database. Collaborating practices are located throughout The Netherlands and the collaborating GPs are comparable to other GPs in the country according to age and sex. In the Netherlands, all citizens are registered with a GP practice, which forms the point of care and acts as a gatekeeper in a two-way exchange of information with secondary care. The medical records of each patient can therefore be assumed to contain all relevant medical information including medical findings and diagnosis

from secondary care. The IPCI database is representative for the Dutch population regarding age and sex (Voordouw et al 2004).

The database contains information on about 1.6 million patients. This is the cumulative number of patients who have ever been part of the dynamic cohort of patients who have been registered. ICPC is the coding system for patient complaints and diagnoses, but diagnoses and complaints can also be entered as free text. Prescription data such as product name, quantity prescribed, dosage regimens, strength and indication are entered into the computer (Vlug et al 1999). The National Database of Drugs, maintained by the Royal Dutch Association for the Advancement of Pharmacy, enables the coding of prescriptions, according to the ATC classification scheme recommended by the World Health Organization (WHO).

As this is a primary care database, information on specialist prescribing, drug dispensing and actual drug intake is missing.

IPCI is listed under the ENCePP resources database. (www.encepp.eu/encepp/resourcesDatabase.jsp).

8.2.4 Electronic clinical records in primary care as a source of information for epidemiological research (SIDIAP)

GPs play an essential role in the public health care system of Spain, as they are responsible for primary health care, long-term prescriptions and specialist and hospital referrals. The Spanish public health care system covers more than 98% of the population. The SIDIAP Database comprises of electronic medical records of a representative sample of patients attended by GPs in Catalonia (North-East Spain), covering a population of more than 5.1 million patients (about 80% of the total of 7.5 million population of Catalonia) from approximately 274 primary care practices with 3,414 participating GPs. The SIDIAP data comprises the clinical and referral events registered by primary care health professionals (GPs and nurses) and administrative staff in electronic medical records, comprehensive demographic information, prescription and corresponding pharmacy invoicing data, specialist referrals, primary care laboratory test results, and hospital admissions and their major outcomes. Health professionals gather this information using ICD-10 codes, and structured forms designed for the collection of variables relevant for primary care clinical management, such as country of origin, sex, age, height, weight, body mass index, tobacco and alcohol use, blood pressure measurements, and blood and urine test results. Only GPs who meet quality control standards can participate in the SIDIAP database. Encoding personal and clinic identifiers ensures the confidentiality of the information in the SIDIAP Database. Recent reports have shown the SIDIAP data to be useful for epidemiological research (Garcia-Gil Mdel et al 2011).

SIDIAP is listed under the ENCePP resources database. (www.encepp.eu/encepp/resourcesDatabase.jsp).

8.3 Subjects

8.3.1 Source population

The source population for this study consisted of all subjects that were registered with general practitioners who contributed data to the four participating databases between January 1st

2000 and December 31st 2013. In order to create the study cohort and apply the inclusion and exclusion criteria, follow-up period for each member of the source population was first created. Start of follow-up was the latest date of the following: first registration in the database+365 days, 18th birthday, or start of the study period (1/1/2000). Follow-up ended at the earliest of the following dates: end of the study period, 80th birthday, transferring out, or death.

Within this source population of subjects 18-79 years of age with at least 365 days of valid data in the database, three different study cohorts were defined, one primary and two secondary.

1. A cohort of incident hypertension patients (Cohort 1) which is the primary cohort.
2. A cohort of incident and prevalent hypertension patients (Cohort 2), secondary cohort.
3. A cohort of incident hypertension patients, whom have had a colonoscopy which was negative for colorectal cancer, polyps and cysts (Cohort 3), secondary cohort.

8.3.2 Inclusion criteria

The cohorts could be generally defined as adult patients (aged 18-79 years) in the individual databases with a diagnosis of arterial hypertension between January 01st 2000 and December 31st 2013.

Specific inclusion criteria for the three different cohorts were:

Cohort 1 (primary analyses): incident hypertension patients

- Patients with the first diagnosis of arterial hypertension occurring after start of follow-up without a preceding event of complication of hypertension.

Cohort 2: incident and prevalent hypertension patients

Cohort 2 contains all patients of Cohort 1 plus patients fulfilling one of the following criteria:

- a diagnosis of arterial hypertension before start of follow-up or a recorded complication of hypertension prior the date of the first diagnosis of arterial hypertension
- a recorded complication of hypertension without any diagnosis of arterial hypertension

Complications of hypertension included: hypertensive heart disease (Healey and Connolly, 2003) i.e. cardiomegaly; hypertensive retinopathy; hypertensive encephalopathy or hypertensive renal disease [glomerulosclerosis or nephrosclerosis]. In Annex 2 codes for identification of complications of hypertension are given. If an incident hypertension subject developed a complication of hypertension during follow-up, the subject was not considered as a prevalent case.

Cohort 3: incident hypertension patients with a negative colonoscopy

- all inclusion criteria for Cohort 1

- furthermore, a colonoscopy negative for the study outcomes (colorectal polyps, hyperplasia, cysts and neoplasm) within the period 10 years prior to and 91 days after the date of hypertension diagnosis.

This cohort is a subset of cohort 1. These timeframes are chosen according to colorectal cancer surveillance and follow-up timings as recommended in guidelines (Lieberman et al 2012); a time frame up to 10 years may be chosen when a colonoscopy turned out to be negative, whereas an interval of around 5 years should be chosen when there were polyps < 10 mm removed at the previous colonoscopy. Cohort entry starts at time of fulfilling all inclusion criteria into this cohort.

This colonoscopy cohort was included because of the potential of Berksonian bias (diagnostic bias) as patients with a diagnosis in routine clinical care (hypertension) may be more easily/sooner referred for clinical check-up and diagnostic procedures as compared to subjects without the disease (non-hypertensive subjects). Presence of Berkson bias is likely in the hypertensive cohort given the fact they were routinely followed in medical care and often may have other/advanced cardiovascular diseases that require antiplatelet therapy. The use of antiplatelet medication may enhance the likelihood of colorectal bleeding, for which they might undergo a colonoscopy. The chance of finding any of the primary endpoints is therefore higher in a hypertension cohort as compared to the general non-hypertensive population. In order to look at the effect of potential Berksonian bias, a subgroup analysis in hypertensive subjects who had a 'negative' colonoscopy was conducted. Only in a population with a recorded 'negative' colonoscopy one can be certain that patients were event-free at a certain point in time..

8.3.3 Exclusion criteria

From each of the three study cohorts, patients with one or more of the following conditions were excluded:

- Patients who did not have a diagnosis of arterial hypertension but used antihypertensive drugs exposure for other indications, such as portal hypertension or pulmonary hypertension)
- A history of colorectal malignant cancer prior to cohort entry (see Section 8.3.1 – 8.3.2 for definition of cohort entry)
- A history of cancer (excluding non-melanoma skin cancer and colorectal cancer) prior to cohort entry (see Section 8.3.1 – 8.3.2 for definition of cohort entry)
- Less than 30 days of follow-up after cohort entry (see Section 8.3.1 – 8.3.2 for definition of cohort entry)

Exclusion criteria for Cohort 1 and Cohort 3 (incident hypertension)

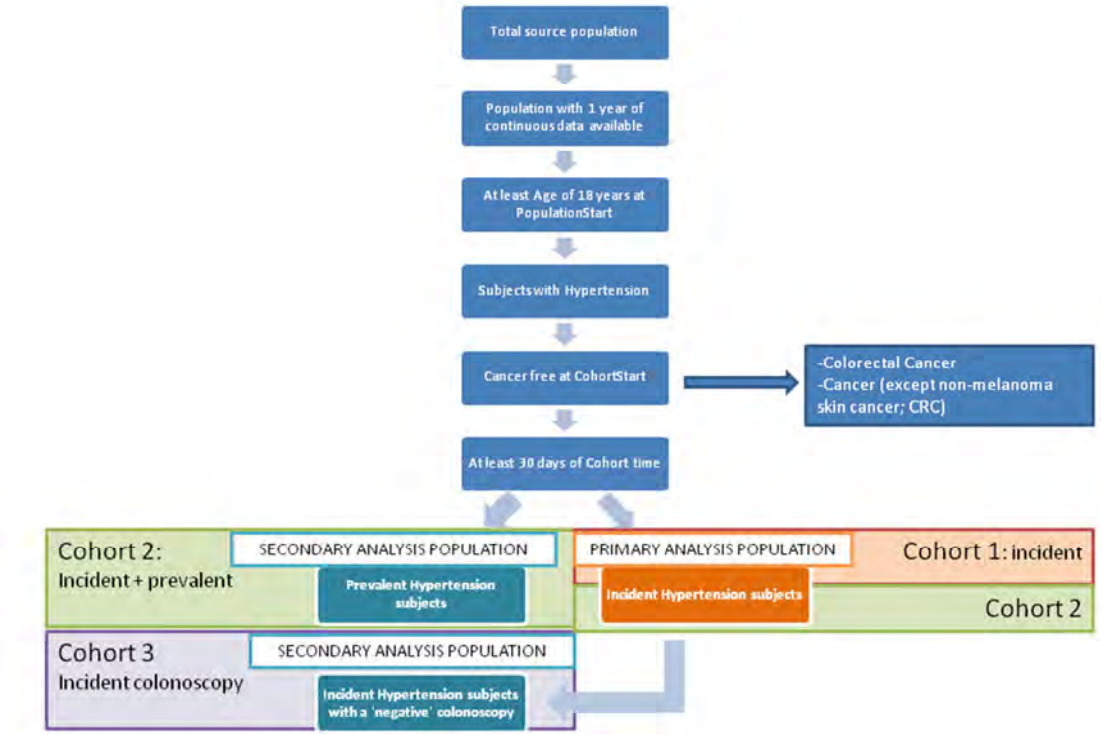
- a diagnosis of arterial hypertension before start of follow-up (see Section 8.3.1) .
- a diagnosis of complications of hypertension preceding the first diagnosis of arterial hypertension

The sequence of hypertension to complications of hypertension may take years to decades and also due to co-morbid occurrence of other diseases, such as obesity and type 2 diabetes

mellitus (T2DM), inclusion of patients who are identified with a 'hypertension' complications are advanced in their disease and if severity of hypertension would increases the risk of colorectal neoplasms this would lead to an overestimation of risk and unclear hazard function. By using an inception cohort (new hypertension patients), patients at a more homogenous stage are included and the risk over time is estimated better.

Starting from the overall population, the number of subjects meeting these specific inclusion and exclusion criteria are presented in the flow chart in Figure 8-1 .

Figure 8-1 Flow chart of study population



In order to create the correct cohort entry date operationally and distinguish between the certainty of hypertension diagnosis the following mutually exclusive groups were created within each of the three cohorts (Figure 8-2):

1. At least 1 diagnosis code of arterial hypertension and no measurement with elevated blood pressure (recorded 4 weeks before or after diagnosis code) and no evidence of antihypertensive treatment (diuretics, beta-blockers, calcium antagonists, ACE-inhibitor, angiotensin receptor blocker; on or up to four weeks after the date of the diagnosis). *The date of the diagnosis code recorded in the database was used as start date.* This case definition likely captures untreated and mild new onset hypertension, however, it could capture misclassified hypertension subjects as no additional evidence in the form of treatment or blood pressure measurement of hypertension is present around the time of the diagnosis.
2. Documentation of at least 1 diagnosis code of arterial hypertension **plus a measurement** with elevated blood pressure (defined as systolic blood pressure ≥ 140 mmHg and/or diastolic ≥ 90 mmHg) recorded 4 weeks before or after diagnosis code and no antihypertensive treatment (diuretics, beta-blockers, calcium antagonists, ACE-inhibitor, angiotensin receptor blocker; on or up to four weeks after the date of the diagnosis). This approach would allow capturing 'untreated' or less severe hypertensive patients who may have started their hypertension follow-up and care by lifestyle changes rather than treatment. *The date of diagnosis code recorded in the database* was *used as cohort entry.*
3. Documentation of at least 1 diagnosis code of arterial hypertension **plus start of treatment** of any of the following antihypertensive drugs (diuretics, beta-blockers, calcium antagonists, ACE-inhibitor, angiotensin receptor blocker) on or up to four weeks after the date of the diagnosis and no recorded blood pressure measurement indicative of hypertension within 4 weeks. A four week timeframe is chosen to be more conservative and not miss any starting treatments for hypertension and take any delay in recording of the prescriptions into account. This case definition captures the treated hypertensive patients. *The date of diagnosis code recorded in the database was used as cohort entry.*
4. Documentation of at least 1 diagnosis code of arterial hypertension **plus a measurement** with elevated blood pressure (defined as systolic blood pressure ≥ 140 mmHg and/or diastolic ≥ 90 mmHg) recorded 4 weeks before or after diagnosis code hypertension **plus start of treatment** of any of the following antihypertensive drugs (diuretics, beta-blockers, calcium antagonists, ACE-inhibitor, angiotensin receptor blocker) on or up to four weeks after the date of the diagnosis. This case definition captures patients with more certain hypertension as measurement and treatment ensures the diagnosis and therapy at start of initial hypertension diagnosis. *The date of diagnosis code recorded in the database was used as cohort entry.*

Table 8-1 Mutually exclusive groups of hypertensive subjects entering any of the three hypertension cohorts

Classified according to above mentioned criteria in group:	Hypertension *code*	Elevated blood pressure *Measurement*:	*Antihypertensive treatment*	Assumed to reflect:
1	Yes	No	No	Untreated, less severe, possibly misclassified hypertension

Classified according to above mentioned criteria in group:	Hypertension *code*	Elevated blood pressure *Measurement*:	*Antihypertensive treatment*	Assumed to reflect:
2	Yes	Yes	No	Untreated, less severe hypertension
3	Yes	No	Yes	Treated hypertension
4	Yes	Yes	Yes	Certain (more severe) hypertension

8.3.4 Primary source of data. Primary analysis population

An example of inclusion and follow-up in the three cohorts is provided in Figure 8-2 for the primary study cohort, in Figure 8-3 for the prevalent+incident cohort (secondary analysis) and in Figure 8-4 for the negative colonoscopy cohort.

Some issues can be observed from these graphs that require explanation. In the prevalent+incident cohort the hypertension diagnosis of prevalent hypertension patients is prior to start of follow-up. Since the exact date of hypertension onset is not known in all cases (all history may not be available) cohort entry was the start of follow-up in this instance.

In the negative colonoscopy cohort the onset of hypertension was combined with the negative colonoscopy to define cohort entry. Cohort entry was the diagnosis of hypertension in the incident cohort, if a negative colonoscopy was present in the data in the period 10 years (3653 days) prior to first hypertension diagnosis. To accommodate more recent colonoscopies the patients were allowed to have a colonoscopy within the first 3 months after first hypertension diagnosis, and then the date of colonoscopy was used as cohort entry.

Figure 8-2 Hypertension cohort entry – incident cases

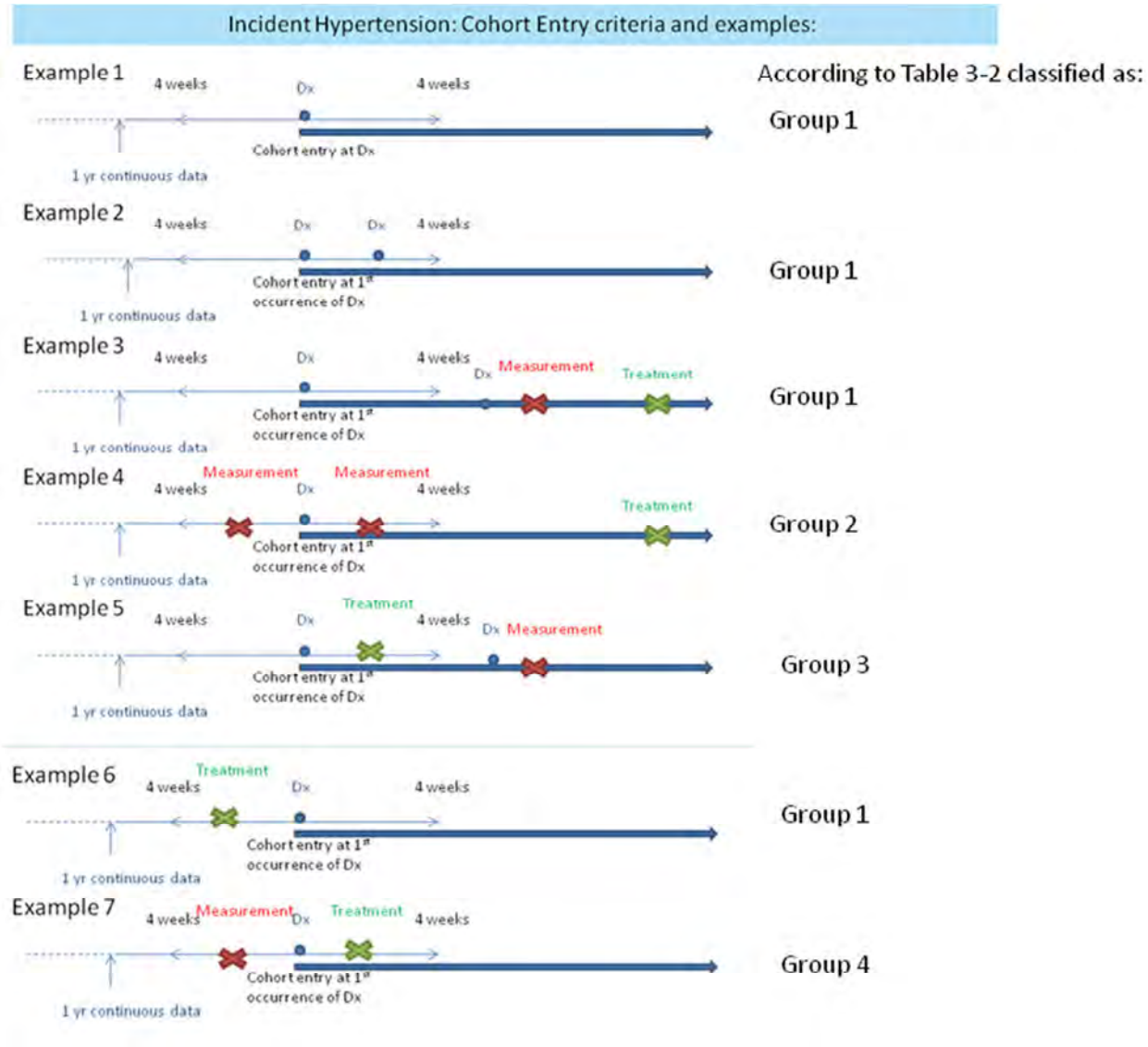


Figure 8-3 Hypertension cohort entry – prevalent cases

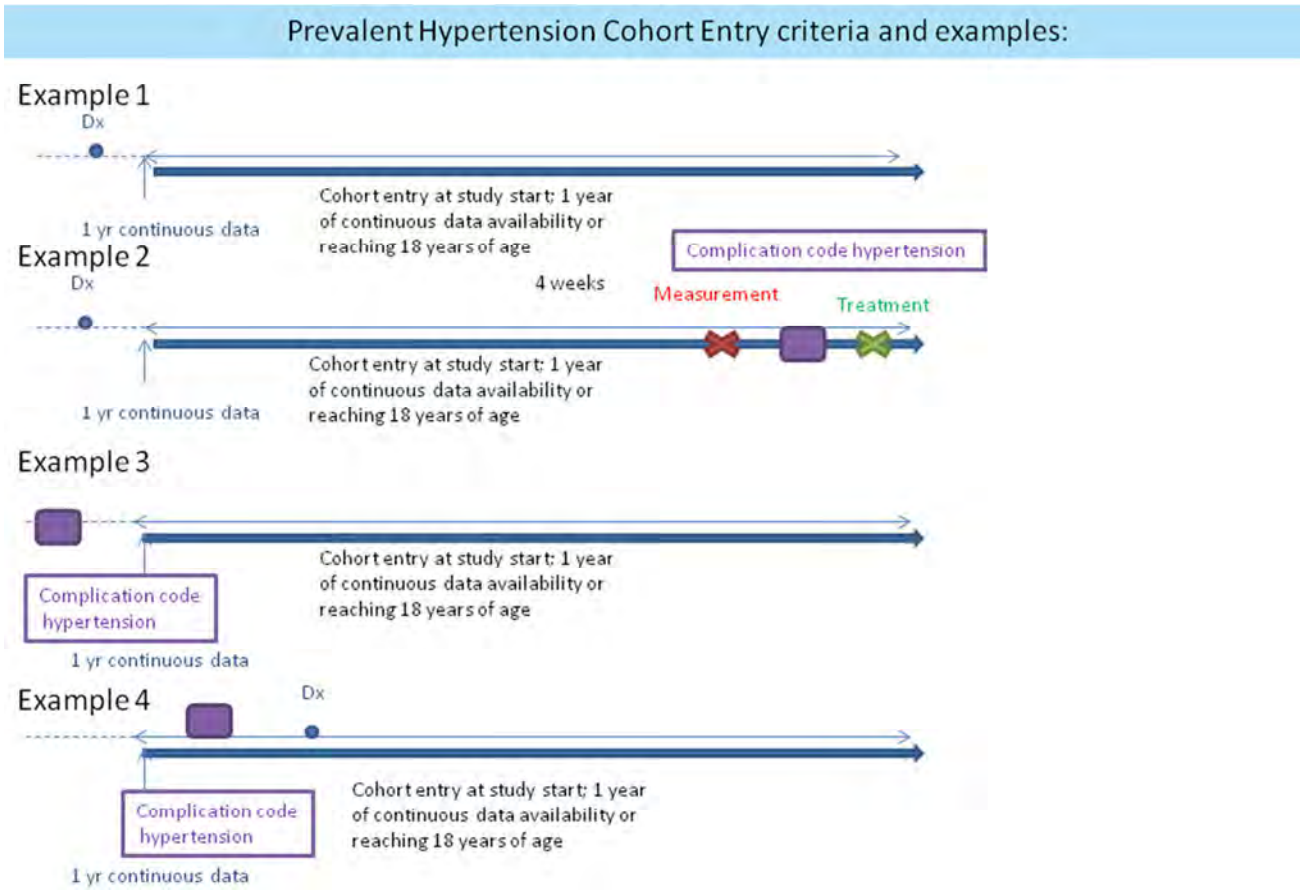
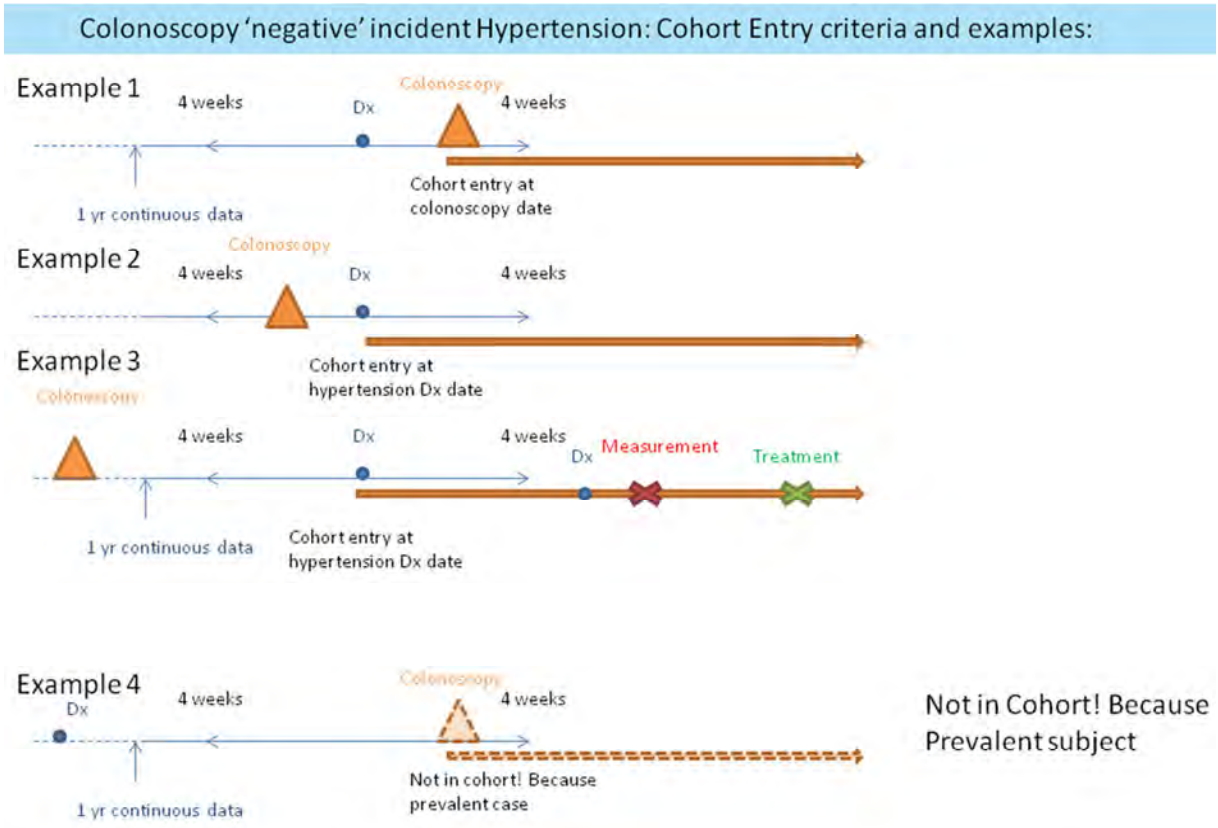


Figure 8-4 Incident Hypertension subjects with a 'negative' colonoscopy



8.4 Variables

Two datasets were created: one dataset for the (Cohort 1 and Cohort 2), and one dataset for Cohort 3. In addition to follow-up times and periods, these files comprised the time till event, the type of outcome, certainty group, hypertensive treatment, and co-morbidity: Inflammatory Bowel Disease, hereditary cancer syndromes, negative colonoscopy, cholecystectomy, obesity, use of low dose aspirin, history of ischemic heart disease, history of cerebrovascular disease, hyperlipidemia, hypertriglyceridemia, hypercholesterolemia, Type 2 diabetes mellitus, smoking, excessive alcohol use. Annex 4 describes the technical details for each of the variables.

The endpoints of interest were defined as a diagnosis corresponding to colorectal hyperplasia defined as any of the following diagnoses and/or conditions:

- Cysts = Colorectal cysts: cysts of the colon, rectum and perineum, other than pilonidal sinus or cysts
- Polyps = colorectal polyps (with/without hyperplasia and irrespective of the type or size of polyp): Polyps in the colon or rectum, these are precursor lesions of subsequent colorectal adenomas and cancer

- Benign Neoplasms = Benign Colorectal neoplasms
- Colorectal Cancer = Malignant Colorectal Neoplasms with or without metastases
  Malignant lesions of the colon, rectum, such as adenomas and adenocarcinomas

Part of the case definition of colorectal hyperplasia also included: any ulcerations or colonic bleeding related or due to colorectal malignancy or cell dysplasia. Carcinoma in situ was therefore also included. These specific conditions of colorectal hyperplasia are included as they have been included as outcome definition in the SPP100A2418 study. Codes corresponding to colorectal hyperplasia are shown in Annex 1.

Tumor size and tumor type (adenocarcinoma or other more rare histopathological types of cancer) was not taken into account, neither polyp size or polyp type (such as a distinction between tubular, villous or sessile polyps).

8.5 Data sources and measurement

The data sources used in the study are described in Section 8.2.

8.6 Bias

Three cohorts were analyzed: (1) incident subjects with hypertension; (2) incident and prevalent subjects with hypertension and (3) incident hypertension subjects that had undergone a colonoscopy that was negative for any of the study outcomes. In this way, it can not only demonstrate that prevalent hypertension will inflate the incidence rates, as along with the duration of hypertension the patient is more likely to develop colorectal outcomes given the common risk factors for hypertension and colorectal outcomes. Secondly, by including the third cohort of subjects with a 'negative' colonoscopy only from this group it can be certain that they did not have any outcome before cohort entry; as the outcomes of interest in the study require a diagnostic procedure to be certain of presence and absence of the outcomes.

Information on some of the covariables might have been missing. This is particularly relevant for the following risk factors: body mass index (BMI), alcohol use and smoking. It is likely that in more recent years, more information on body mass index is available as compared to beginning of 2000. BMI values were classified into categories, and considered if there was no information available as missing information and as a separate category ('unknown BMI'). It is likely that only information on severe alcohol abuse is recorded by the GP, therefore any recording of alcohol problems was considered as 'excessive' use, and if no recording is mentioned it was be classified as 'normal alcohol use. The same may be true for smoking, as patients may seek help from their GP to quit smoking, rather than that they record patients do not smoke at all.

8.7 Study size

Table 8-2 Study size of primary and secondary analysis populations across data sources

	CPRD (UK)	HSD (IT)	IPCI (NL)	SIDIAP (ES)
Total source population	13,673,353	1,738,798	1,708,253	6,332,740
Population with 1 year of	10,924,766	1,627,136	1,672,154	6,175,913

continuous data				
Aged 18-79 years	8,517,175	1,501,405	1,327,309	4,963,766
Subjects with Hypertension Dx	1,353,403	376,760	225,038	1,006,441
Cancer-free at cohort entry	1,192,143	344,532	203,027	931,582
Less than 30 days of follow-up	6756	2201	869	6126
Study cohorts				
Cohort 2. Incident+Prevalent hypertension	1,185,387	342,331	202,158	925,456
Cohort 1. Incident hypertension	688,674	237,922	40,568	336,456
Cohort 3. Incident hypertension and negative colonoscopy	12,956	6,896	3,112	7,154

CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

8.8 Data transformation

This is described in Section 8.4.

8.9 Statistical methods

SAS for Windows Version 9.3 was used for the analyses.

8.9.1 Main summary measures

Descriptive analyses used proportions for categorical variables and mean and standard deviations for continuous variables (e.g., age). Incidence rates and 95% confidence intervals were obtained by dividing the number of events over the number of person-years accumulated. Incidence Rate Ratios (IRR) with 95% confidence intervals were calculated by Poisson regression. Standardized incidence rates were calculated by applying the observed incidence rates to the reference population incidence rate. To estimate the time to disease, Kaplan Meier curves were used.

8.9.2 Main statistical methods

Incidence rates of Colorectal Outcomes

Incidence rates were calculated for all three hypertension cohorts. Incidence rates were calculated by dividing the number of incident cases by the total number of person-years at risk within the study population.

The event types for which IRs were given are:

- Composite endpoint (any of the outcomes, whichever is first reported)
- Rates of polyps
- Rates of cysts
- Rates of benign neoplasm
- Rates of colorectal cancer (CRC)

When calculating IRs for specific events person time was censored at time of this event.

IRs were provided for the total cohorts, as well as specific for age, sex, calendar year and by presence of co-morbidities (Inflammatory Bowel Disease, Hereditary Cancer syndromes, Negative Colonoscopy, Cholecystectomy, Obesity, Use of Low Dose Aspirin, History of Ischemic heart disease, History of Cerebrovascular disease, Hyperlipidemia, Hypertriglyceridemia, Hypercholesterolemia, Type 2 Diabetes Mellitus, Smoking, Excessive Alcohol use) and BMI category (underweight; normal; overweight; obese; severe obese; unknown).

Standardization of the overall incidence rates was done to be able to compare the incidence rates across the databases participating in the current study, and for external comparison with other studies. Age and gender were used for standardization to the world reference population (Eurostat 2013)). By stratifying the results by five years age group and sex it is possible to put our results and incidence rates next to those that are reported in literature. In addition, IR were also standardized to the CPRD population. The incidence rates obtained in the study were applied to the age distribution (person-time or persons) of these populations; this yielded the number of expected patients in these populations per year of which the incidence rates may be obtained if they were hypertensive. This can then be compared with the known/observed incidence rate in these countries/databases.

Incidence Rate Ratios

Univariate Poisson regression was used to calculate relative rates (IRR) for gender, age, calendar year , BMI categories and comorbidities with 95 % confidence intervals for each of the four endpoints, as well as for the composite endpoint in all three cohorts (Incident; Incident+Prevalent; Incident colonoscopy). Poisson regression uses a log-linear model in the covariates

For calendar year trends over time were investigated by using year as continuous covariable.

The IRR for comorbities were adjusted for gender and age.

Kaplan Meier analysis

Survival analysis (Kaplan Meier estimates) were used to analyze time from hypertension to event. Kaplan Meier curves were created for each of the four specific events as well as for the composite endpoint in the three cohorts (Incident; Incident+Prevalent; Incident colonoscopy). The analysis was stratified by Gender, Age Category (18-44; 45-64; 65-79), Inflammatory Bowel Disease, Colonoscopy negative and use of any antihypertensive drug before Cohort Start (any antihypertensive drug is defined as the combination of Use of Diuretics, Beta blockers, Calcium Antagonists, ACE Inhibitors, AT II Antagonists and Renin Inhibitors before Cohort Start).

Non-informative Kaplan Meier curves (because of low number of events and/or small subgroups) are not presented.

8.9.3 Missing values

This is described in Section 8.6. For smoking, BMI and colonoscopy a separate category of ‘unknown’ was considered.

8.9.4 Sensitivity analyses

Not applicable.

8.9.5 Amendments to the statistical analysis plan

Not applicable.

8.10 Quality control

Creation of the cohorts, all data transformation and generation of variables and the creation of the files with aggregated person years and events (to calculate incidences) was double programmed in Jerboa and in SAS. An initial quality check module in Jerboa allowed databases to check their data on for instance outliers and errors in coding. Programming for the tables and analyses was done in SAS.

9 Results

All Tables and Figures of the three cohorts and for the five outcomes by databases are presented in a stand-alone document [Final Tables.pdf]. In this report the most important findings were summarized.

9.1 Participants

In Figure 8-2 it is described by database how the three different cohorts were formed.

9.2 Descriptive data

A full description of the general characteristics of the subjects in the three cohorts is shown in in the stand-alone document [Final Tables.pdf].

In most cohorts for CPRD, HSD, IPCI and SIDIAP patients had confirmed hypertension since the diagnosis code often was accompanied by an elevated blood pressure measurements and a prescription. In HSD and SIDIAP a third of the cohort had a diagnosis code plus a prescription without a recorded blood pressure measurement. This observation was seen for all three cohorts, though for the Incident+Prevalent cohort the majority proportion shifted towards entry via diagnosis plus prescriptions only.

Incident cohort

A summary of important variables for the incident cohort are presented in Table 9-1. In CPRD and SIDIAP there was a slight majority of males (51.5% and 55%, respectively) whereas in HSD and IPCI there were slightly more females (51.9% and 52.6%, respectively). The proportion of subjects entering the incident cohort with a prior diagnosis of colorectal polyp was consistent across the databases (1.05% in SIDIAP; up to 1.45% in IPCI).

Use of anti-hypertensive drugs (for other indications than hypertension) prior to cohort entry was high in the incident cohort; with the use of diuretics being mostly used and second often used were ACE-inhibitors, prior use of Renin inhibitors was null. The vast majority of subjects with a BMI value available were either overweight or obese; a finding seen in all

databases. Hyperlipidemia was present in between 15.5% (HSD) and 31.6% (SIDIAP) of incident hypertension subjects.

Table 9-1 General Characteristics of primary study cohort

	CPRD (UK)	HSD (IT)	IPCI (NL)	SIDIAP (ES)
N Total	688,674	237,922	40,568	336,456
Age 18-44	16.1%	15.2%	15.5%	14.3%
Age 45-64	52.1%	51.5%	54.3%	54.1%
Age 65-79	31.8%	33.2%	30.3%	31.5%
Female	48.5%	51.9%	52.6%	45.0%
Male	51.5%	48.1%	47.4%	55%
Prior diagnoses of colorectal polyp	1.10%	1.16%	1.45%	1.05%
Prior use of diuretics	15.5%	15.5%	18.4%	21.2%
Prior use of ACE inhibitors	11.7%	13.11%	13.5%	22.3%
Prior use of Renin inhibitors	-	0%	0.1%	0.2%
Obesity	21.5%	9.7%	14.8%	25.5%
Hyperlipidemia	21.7%	15.5%	22%	31.6%

CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

Other cohorts

Generally, mean ages were comparable across the different datasets, with older aged subjects in the Incident+Prevalent Cohort than in the Incident Cohort. In the same way, subjects in the Incident Colonoscopy Cohort (Cohort 3) had a 3-4 years higher mean age than in the Incident+Prevalent Cohort (Cohort 2). Prevalence of comorbidities were higher in the Incident+Prevalent cohorts than in Incident and Incident Colonoscopy (Cohort 3) cohorts. In the Incident Colonoscopy cohort the presence of IBD was more prevalent than in the Incident Cohort (between 1.34% and 9.09% in Incident Colonoscopy Cohort versus between 0.27% and 1.56% in the Incident Cohort). This was also seen for IBS with a higher prevalence of IBS in the Incident Colonoscopy cohort than in the Incident Cohort.

9.3 Follow-up

Time in cohort in the different cohorts by database is shown in Table 9-2. Follow-up was longest for the incident+prevalent cohort since cohort entry started at the start of follow-up for patients with prevalent hypertension, whereas it started at the time of first diagnosis during follow-up for the incident cohorts.

Table 9-2 Time in Cohort in different cohorts by database

	CPRD (UK)	HSD (IT)	IPCI (NL)	SIDIAP (ES)
Incident Cohort				

	CPRD (UK)	HSD (IT)	IPCI (NL)	SIDIAP (ES)
Median (years)	5.4	6.2	1.9	3.4
Q1 – Q3	2.6-8.8	3.1-9.5	0.9-3.2	1.7-5.1
Incident+Prevalent Cohort				
Median (years)	6.5	6.7	2.6	5.8
Q1 – Q3	3.1-10.4	3.3-10.5	1.4-3.9	3.0-7.0
Incident Colonoscopy Cohort				
Median (years)	5.4	4.2	1.6	2.1
Q1 – Q3	2.6-8.6	2.1-7.1	0.8-2.9	1.0-3.7

CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

9.4 Incidence rates

The number of events within each cohort is shown in the stand-alone document [Final Tables.pdf]. In Table 9-3 an overview of the incidence rates is provided. Of note is that there were no benign neoplasms recorded in HSD and no cysts in IPCI.

Table 9-3 Incidence Rates (per 100,000 person years) of outcomes in different cohorts by database

	CPRD (UK)	HSD (IT)	IPCI (NL)	SIDIAP (ES)
Incident Cohort				
Composite Event	348.21	517.49	510.29	628.22
Polyp	241.85	380.16	426.37	484.34
Cyst	6.36	2.94	0	2.25
Benign neoplasm	4.17	0	1.10	10.05
Colorectal Cancer	124.34	147.10	107.18	141.76
Incident+Prevalent Cohort				
Composite Event	343.40	506.81	567.66	589.62
Polyp	227.17	366.43	472.83	423.33
Cyst	8.81	3.03	0	4.26
Benign neoplasm	4.81	0	0.71	15.32
Colorectal Cancer	131.94	149.39	130.18	158.03
Incident Colonoscopy Cohort				
Composite Event	758.09	1015.99	1391.77	2223.25
Polyp	617.07	777.33	1306.09	2043.14
Cyst	8.06	12.08	0	0
Benign neoplasm	8.05	0	0	28.68
Colorectal Cancer	161.96	136.52	115.06	143.55

CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

Incidence rates of colorectal cancer were higher in the cohort that included prevalent subjects (Incident+Prevalent cohort) than the incident cohort. For the other outcomes, the incidence rates were comparable between the incident and Incident+Prevalent cohort. In the incident colonoscopy cohort, the incidence rates of all outcomes were higher than in the incident cohort.

9.5 Time trend analyses

Figure 9-1 and Figure 9-2 present the IRs of Polyps and CRC. Table 9-4 and Table 9-5 present the personyears that were used for the calculation of these IR.

Across the calendar years of hypertension diagnoses, incidences of polyps and CRC did not increase or decrease substantially. However, the following patterns were observed:

For the incident cohort:

- The incidence rate in polyps seemed to increase after the year 2008, particularly in HSD, SIDIAP and IPCI. For CPRD, the increase was not very prominent (Figure 9-1).
- The incidence rate of CRC seemed to decrease in HSD and CPRD since the year 2004-2005. The incidence rate in IPCI is less stable but in general a decreasing trend can be seen. The incidence rate of CRC in SIDIAP seemed to be stable over calendar time (Figure 9-2).

Figure 9-1 Incidence Rate of Polyps in Incident Hypertension Cohort by database over calendar years

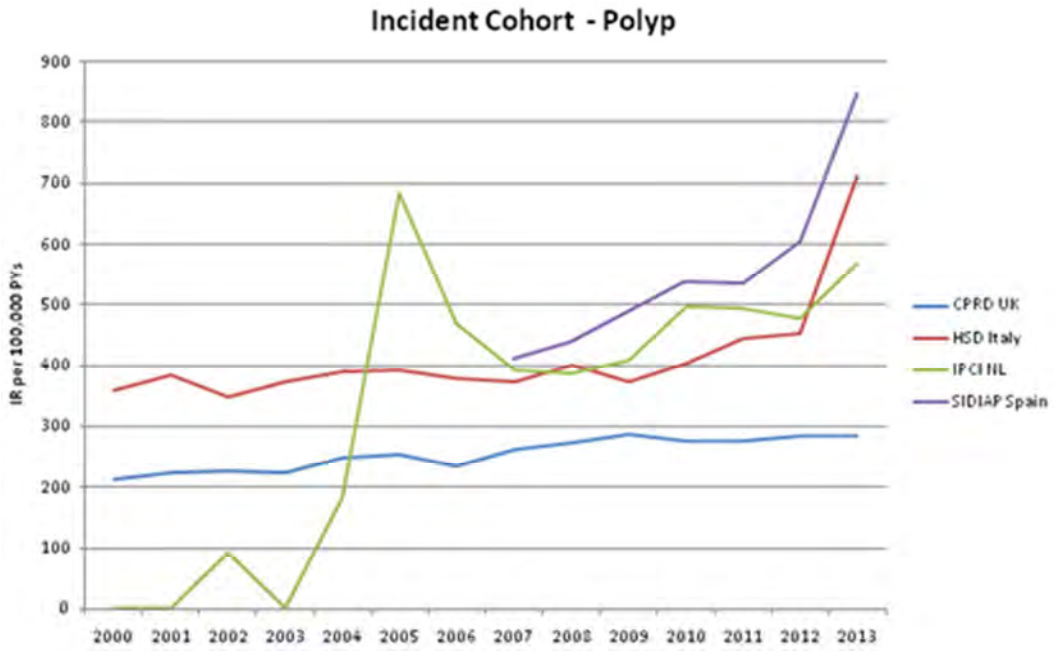


Table 9-4 Personyears for calculation of the incidence Rates of Polyps in Incident Hypertension Cohort by database over calendar years

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
CPRD	4.72	4.81	4.88	4.95	4.88	3.87	3.10	2.54	2.01	1.51	1.12	0.73	0.48	0.16
HSD	2.31	1.98	1.83	1.46	1.44	1.50	1.20	0.92	0.79	0.62	0.47	0.31	0.18	0.06
IPCI	0.02	0.02	0.01	0.01	0.01	0.01	0.04	0.06	0.12	0.15	0.16	0.15	0.10	0.04
SIDIAP								3.19	2.42	2.10	1.66	1.10	0.72	0.24

Person Years X 100,000

Figure 9-2 Incidence Rate of CRC in Incident Hypertension Cohort by database over calendar years

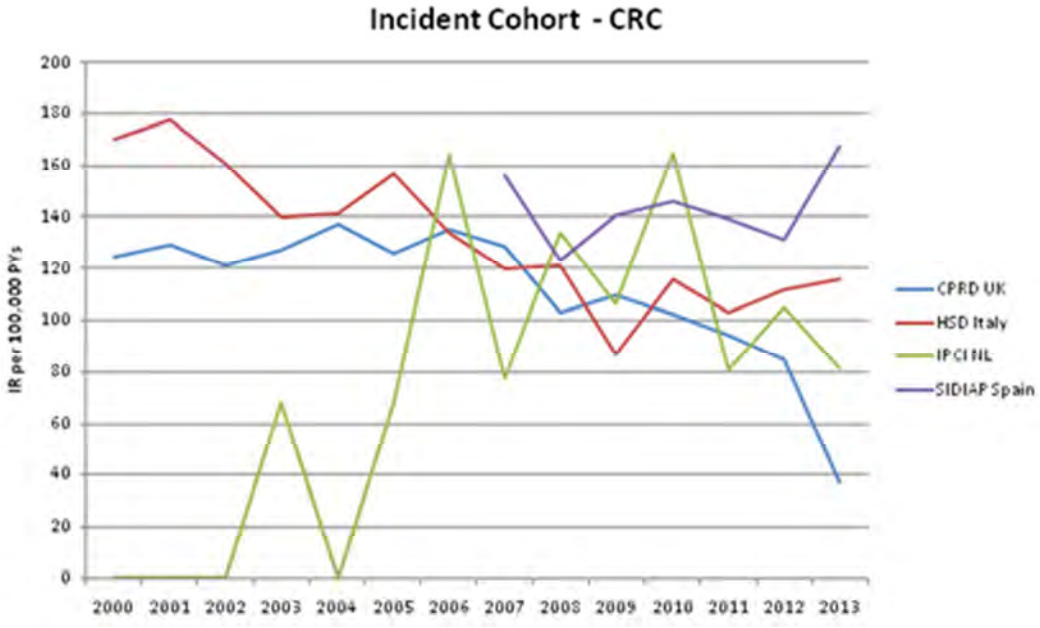


Table 9-5 Personyears for the Incidence Rates of CRC in Incident Hypertension Cohort by database over calendar years

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
CPRD	4.74	4.83	4.90	4.97	4.91	3.89	3.11	2.55	2.02	1.51	1.12	0.73	0.48	0.16
HSD	2.34	2.01	1.85	1.48	1.46	1.52	1.22	0.93	0.80	0.62	0.48	0.31	0.18	0.06
IPCI	0.02	0.02	0.01	0.01	0.01	0.01	0.04	0.06	0.12	0.15	0.16	0.15	0.10	0.04
SIDIAP								3.21	2.44	2.12	1.67	1.10	0.73	0.24

Person Years X 100,000

9.6 Standardization

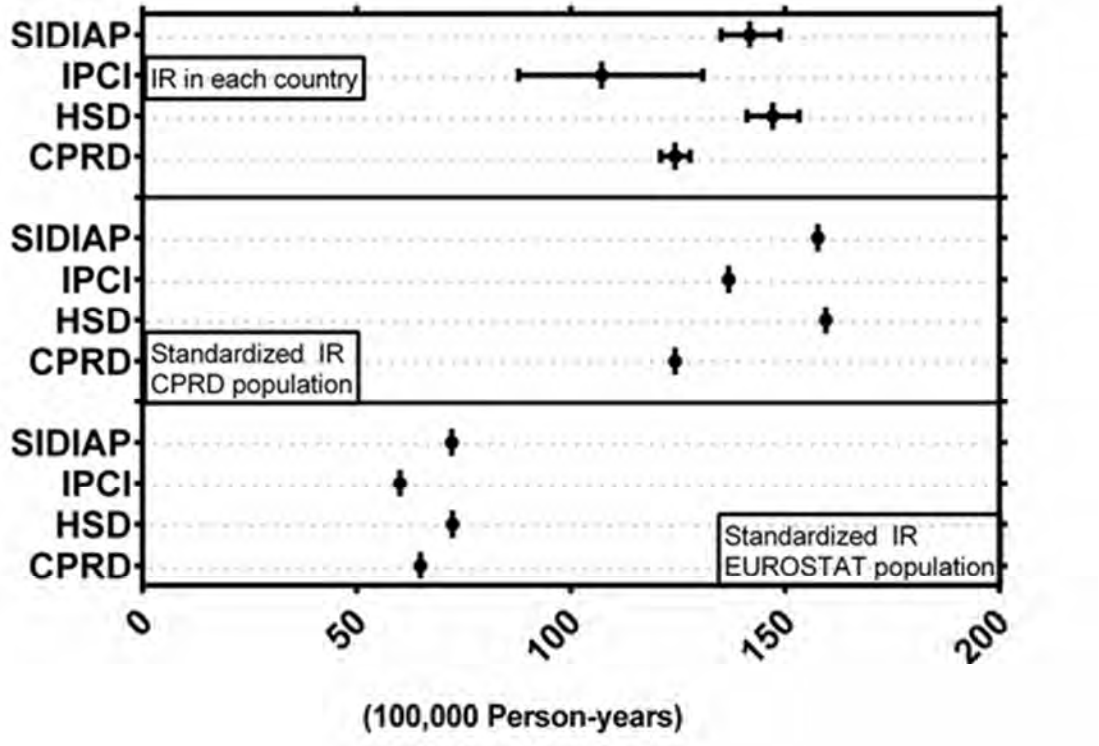
Since the underlying age-sex distributions differ between databases, a standardization of the observed incidence rates was performed by estimating standardized incidence rates (SIR). To perform this standardization, the IR was recalculated as if the age distribution in the population studied was equal to the age distribution in the reference population. For the reference population, both a European general population (Eurostat) for which has an incidence rate of CRC, as well as an internal reference (CPRD) hypertensive population were considered. Since the cohorts in the current study are restricted to hypertensive subjects, these are on average older than the reference Eurostat population. Standardizing our observed rates to the Eurostat population should therefore result in much lower incidence rates.

Colorectal cancer

In the primary study cohort the SIRs for colorectal cancer across the databases varied between 60 and 72 per 100,000 person-years in incident hypertension subjects when standardized to the population (Figure 9-3).

When standardized to the age- and sex-distribution of the CPRD population, rates for colorectal cancer varied between 137 (IPCI) to 160 (HSD) per 100,000 person-years in incident hypertension subjects.

Figure 9-3 Incidence Rate and Standardized Incident Rate of CRC in Incident Hypertension Cohort by database

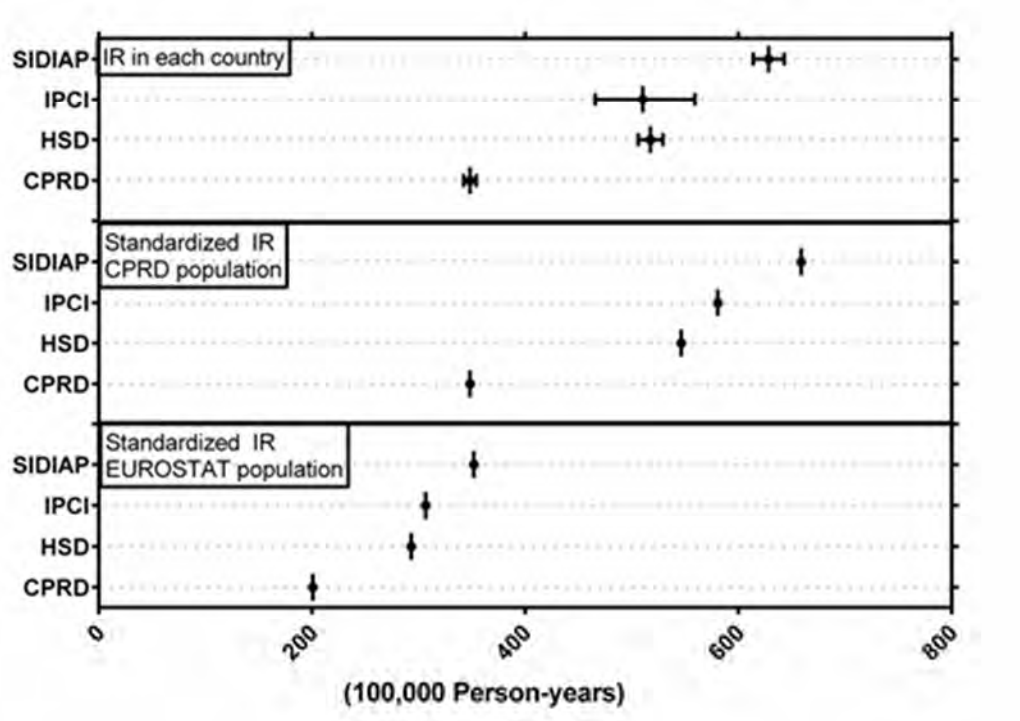


Similar patterns were found for the cohort that included prevalent subjects as well, except that SIRs were slightly lower in this cohort than in the incident cohort. In contrast, in the incident colonoscopy cohort the SIRs were higher than in the incident cohort with SIRs for colorectal cancer using the Eurostat population varying between 59.3 (HSD) and 101 (IPCI) per 100,000 person-years. When using the CPRD as reference population the SIRs increased with standardized incidence rates between 115(IPCI) and 166 (SIDIAP) per 100,000 person-years.

Composite endpoint

For the composite endpoint incidence rates varied substantially between databases. They lowered and the variation was smaller when standardizing to Eurostat population (general population) and the rates increased when standardizing to CPRD population (population with hypertension). CPRD rates were lowest (Figure 9-4).

Figure 9-4 Incidence Rate and Standardized Incident Rate of Composite Event in Incident Hypertension Cohort by database



9.7 Risk factors

Colorectal Cancer

Univariately, older age, male gender and most co-morbidities were associated with an increased rate of CRC. Prior use of antihypertensive drugs, which may be considered a proxy (as its causal association was not investigated in this study) for underlying cardiovascular morbidity was associated with a small increase in risk of CRC in CPRD and SIDIAP but not in other databases. Presence of other cardiovascular diseases/proxies showed a similar pattern, type 2 diabetes mellitus was associated with increased adjusted IRRs for colorectal cancer ([Final Tables.pdf]] and Table 9-6). Associations were in the same direction in all databases but mostly not significant in HSD and IPCI because of smaller sizes.

Table 9-6 Sex-Age-Adjusted Incidence Rate Ratios of Colorectal Cancer in Incident Cohort

	CPRD UK			HSD Italy			IPCI NL			SIDIAP Spain		
	IRR	95% LL	95% UL	IRR	95% LL	95% UL	IRR	95% LL	95% UL	IRR	95% LL	95% UL
History of any hypertensive drug before entry	1.10	1.03	1.16	1.05	0.97	1.15	1.04	0.70	1.56	1.13	1.02	1.24
History of low dose aspirin	1.12	1.04	1.21	0.97	0.85	1.10	1.27	0.66	2.46	1.05	0.92	1.21
Obesity	1.09	1.01	1.17	0.82	0.69	0.96	0.71	0.36	1.41	1.06	0.94	1.18
Hyperlipidemia	1.12	1.05	1.20	1.06	0.95	1.18	1.16	0.74	1.81	1.06	0.96	1.18
HyperCholesterolemia	1.14	1.07	1.22	1.05	0.94	1.18	1.19	0.78	1.83	1.03	0.93	1.14
Type 2 Diabetes	1.16	1.06	1.27	1.24	1.10	1.40	1.01	0.49	2.08	1.35	1.20	1.52
Past Smoking	1.23	1.15	1.33	1.18	0.99	1.42	1.21	0.69	2.15	1.03	0.88	1.20
Alcohol excess	1.38	1.13	1.68	1.53	0.85	2.78	0.52	0.07	3.71	1.33	0.91	1.95

*Use of any antihypertensive drug (ACE-inhibitors; ATII antagonists; Beta-blockers, Calcium Antagonists, Diuretics before cohort entry, Clinical Practice Research Datalink (CPRD), United Kingdom (UK), Health Search/CSD Longitudinal Patient (HSD) database, Italy (IT), Integrated Primary Care Information (IPCI), the Netherlands (NL), Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària (SIDIAP), Spain (ES)

When comparing the estimated adjusted IRRs in the Incident+Prevalent cohort with the corresponding IRR obtained in the Incident cohort, in general similar observations are seen with higher adjusted IRRs when comorbidities are present than when they are not present. Also, the IRRs were quite similar in both cohorts (tables in stand-alone document only [Final Tables.pdf]).

For the incident colonoscopy cohort the same was observed: adjusted IRRs with presence of comorbid diseases being higher than without presence of comorbidities, though the IRRs are affected by the smaller sample size of this cohort which can be seen in the wider confidence intervals (tables in stand-alone document only [Final Tables.pdf]).

Composite Endpoint

Univariately, older age, male gender and most co-morbidities were associated with an increased rate of specific outcomes and the composite event. Prior use of antihypertensive drugs, which may be considered a proxy (as its causal association was not investigated in this study) for underlying cardiovascular morbidity was associated with a small increase in risk of CRC but not in HSD. Presence of other cardiovascular diseases/proxies showed a similar pattern, type 2 diabetes mellitus was associated with increased adjusted IRRs for colorectal cancer ([Final Tables.pdf] and Table 9-7). Hereditary cancer, history of IBS and moreover a negative colonoscopy were consistently associated with the composite endpoint. Associations were in the same direction in all databases but mostly not significant in HSD and IPCI because of smaller sizes.

Table 9-7 Sex-Age-Adjusted Incidence Ratios of Composite event in Incident Cohorts

	CPRD UK			HSD Italy			IPCI NL			SIDIAP Spain		
	IRR	95% LL	95% UL	IRR	95% LL	95% UL	IRR	95% LL	95% UL	IRR	95% LL	95% UL
*History of antihypertensive drug	1.18	1.14	1.22	0.98	0.94	1.03	1.30	1.08	1.57	1.16	1.11	1.22
History of low dose aspirin	1.22	1.17	1.28	1.07	1.00	1.15	1.21	0.87	1.68	1.10	1.03	1.18
History of Inflammatory Bowel Disease	1.66	1.44	1.92	1.17	0.85	1.59	1.08	0.51	2.27	1.32	0.89	1.98
Hereditary cancer	11.57	4.34	30.83	.	.	.	1.81	0.25	12.86	.	.	.
History of IBS	1.49	1.39	1.59	1.23	1.04	1.46	1.38	1.06	1.81	1.64	1.29	2.08
Negative colonoscopy	2.06	1.91	2.23	1.92	1.72	2.16	2.27	1.77	2.91	3.51	3.14	3.92
Cholecystectomy	1.34	1.24	1.46	1.30	1.09	1.55	1.71	1.17	2.50	1.75	0.97	3.16
Obesity	1.09	1.05	1.14	1.04	0.96	1.12	1.01	0.77	1.32	0.96	0.91	1.01
History of Ischemic Heart Disease	1.19	1.12	1.25	1.09	0.98	1.21	1.35	1.06	1.73	1.07	0.96	1.20
History of Cerebrovascular disease	1.13	1.04	1.22	1.19	1.05	1.34	1.10	0.79	1.54	1.13	1.00	1.29
Hyperlipidemia	1.25	1.20	1.30	1.16	1.09	1.22	1.57	1.29	1.92	1.25	1.19	1.31
Hypertriglyceridemia	1.14	0.94	1.40	1.19	0.99	1.42	1.02	0.26	4.10	1.21	1.08	1.36
HyperCholesterolemia	1.28	1.23	1.33	1.13	1.07	1.21	1.53	1.26	1.85	1.24	1.18	1.31
Type 2 Diabetes	1.12	1.05	1.18	0.96	0.89	1.04	0.86	0.60	1.25	1.11	1.04	1.18
Current Smoking	1.32	1.26	1.39	1.24	1.13	1.36	1.59	1.20	2.11	1.20	1.13	1.29
Past Smoking	1.35	1.29	1.42	1.34	1.22	1.48	1.45	1.10	1.92	1.20	1.12	1.29
Alcohol excess	1.50	1.34	1.68	1.70	1.26	2.28	1.77	1.09	2.87	1.07	0.88	1.30

*Use of any antihypertensive drug (ACE-inhibitors, ATII antagonists, Beta-blockers; Calcium Antagonists, Diuretics) before cohort entry as a proxy of cardiovascular co-morbidity, CPRD: Clinical Practice Research Datalink, ES: Spain, HSD: Health Search/CSD Longitudinal Patient database, IPCI: Integrated Primary Care Information, IT: Italy, NL: the Netherlands, SIDIAP: Sistema d'Informació per al Desenvolupament de la Investigació en Atenció Primària, UK: United Kingdom

9.8 Other analyses

9.8.1 Kaplan Meier analyses

Kaplan Meier figures are created for all 5 outcomes and for all three 3 cohorts. The figures are included in the stand-alone document [Final Tables.pdf]. In Table 9-8 the one-year and 5-year risks are presented.

Table 9-8 One and 5-year risks of outcomes in Incident Cohorts

	CPRD UK		HSD Italy		IPCI NL		SIDIAP Spain	
	1 yr	5 yr	1 yr	5 yr	1 yr	5 yr	1 yr	5 yr
Composite	0.27%	1.52%	0.59%	2.40%	0.47%	2.53%	0.65%	3.06%
Polyps	0.19%	1.03%	0.42%	1.76%	0.39%	2.12%	0.50%	2.36%

	CPRD UK		HSD Italy		IPCI NL		SIDIAP Spain	
Cysts	0.01%	0.03%	0.01%	0.01%	0.00%	0.00%	0.01%	0.01%
Benign neoplasm	0.00%	0.02%	0.00%	0.00%	0.00%	0.00%	0.02%	0.05%
CRC	0.08%	0.55%	0.18%	0.69%	0.10%	0.58%	0.14%	0.71%

Composite Event

For the composite event 5-year risks differed from 1.5% and 3% between the databases in the incident cohort, one year risks are much lower (Table 9-8). In general, these curves were significantly lower (lower event-free follow-up) for males (except for IPCI not significantly); older age groups (65-79 years lower event-free follow-up than 45-64 years; and 18-44 years); prior use of any antihypertensive drug (as proxy for existing cardiovascular disease); history of IBD (due to smaller number of IBD subjects not significantly for SIDIAP; HSD and IPCI); a previous colonoscopy positive for any of the study outcomes. For the Incident+Prevalent cohort similar patterns were seen (all available in stand-alone document [Final Tables.pdf]).

For the Incident Colonoscopy cohort cumulative risks were higher than in the Incident cohort; overall 5 year risks for the composite event varied between 2.5% and >10%. Stratified Kaplan Meier curves are affected by the smaller number of events in each stratum, but similar increase risks as described above for the Incident Cohort were seen by male sex, older age, and prior use of any antihypertensive drug and history of IBD.

Polyp

In the Incident Cohort curves the cumulative 5-year risks ranged between 1% and 2%, one year risks were all below 1%. Presence of male sex, older age; prior use of any antihypertensive drug (as proxy for existing cardiovascular disease); history of IBD; and a previous colonoscopy positive for any of the study outcomes provided higher cumulative risks. These findings are less pronounced than seen for the composite event (see description above).

For the Incident+Prevalent cohort similar observations as described for the Incident cohort are seen. Again, in the incident colonoscopy cohort stratified curves are affected by the smaller number of events occurring in this cohort providing less pronounced effects of the risk factors on the Kaplan Meier curves. Also, the cumulative 5-year risks for polyps in the Incident Colonoscopy cohort were much higher than in the incident cohort, namely varying between 2% and 10% (stand-alone document [Final Tables.pdf])

Cyst, Benign Neoplasm

Given the small number of events for the outcomes Cyst and Benign Neoplasm, only Kaplan Meier curves for the total population by databases are shown and not by stratifying factors as these curves are not informative. One year and 5- year risks are low, which may be based on lack of recording in the databases.

Colorectal Cancer

In the Incident Cohort and in the Incident+Prevalent Cohort consistent results were observed. The 5-year cumulative risk of CRC in the incident Cohort was less than 1%, consistent across

the databases. Cumulative risks for CRC were higher for males; older aged subjects; subjects with prior use of any hypertensive drugs and with a prior positive colonoscopy.

The increased risk of CRC for presence of inflammatory bowel disease was less pronounced; this may be because of the relatively 'short' follow-up time in the databases whereas CRC generally takes decades to develop and because the number of IBD subjects was rather limited. Again, in the Incident colonoscopy cohort stratified curves are affected by the smaller number of events occurring in this cohort providing less pronounced effects of the risk factors on the Kaplan Meier curves.

9.8.2 Incidence rate over time

Incidences of the outcomes since the time of hypertension diagnosis (and thus cohort entry) are shown in the figures in separate documents ([IRperPeriod.pdf]). Incidences of outcomes shortly after hypertension diagnosis may possibly be explained by Berkson bias (diagnostic bias) as patients with a diagnosis in routine clinical care (hypertension) may be more easily/sooner referred for clinical check-up and diagnostic procedures as compared to subjects without the disease (non-hypertensive subjects).

Incident Hypertension Cohort and Incident + Prevalent Cohort

It is indeed seen that in the first 30 days after hypertension diagnosis the incidence is very high in all databases, and subsequently decreases in the second month, but appears to be increasing again after this period; likely illustrating the actual increase in incidence of outcomes in hypertensive subjects. For the outcomes: composite event, polyps and CRC, the initial peak in the first 30 days after cohort entry is more prominent in HSD and SIDIAP than in IPCI and CPRD.

Figures for Cyst and Benign neoplasm outcomes are derived from a small number of events occurring in the three cohorts, as can been seen by the wider 95% confidence intervals.

Incident Colonoscopy Cohort

For the composite event and polyps for IPCI and HSD a peak in the incidence in the first 3 months was seen, which was less prominent in CPRD and SIDIAP. In all four databases, in the first month after cohort entry the incidence rate of CRC was high, which decreased since time of hypertension diagnosis. Also, the curves are less stable in the Incident colonoscopy cohort because of the restricted sample size in this cohort.

9.8.3 Validation of Outcomes

Validation of outcomes was performed in a random sample of events in IPCI. The positive predictive values are shown in [[Final Tables.pdf]]. From the definite and probable cases of this random sample a training set was created that was trained by the machine learner and subsequently applied on the remaining set of potential cases. The remaining classified cases were again manually validated to ensure the date of diagnosis and the evidence of the diagnosis. The final cases were included in the study.

9.9 Adverse events/adverse reactions

Not applicable.

10 Discussion

10.1 Key results

Incidence rates of colorectal outcomes (polyps, cysts, benign neoplasm and cancer) were consistent in the four European data sources. Incidence rates of the outcomes were higher with increasing age, in males, with presence of comorbid diseases (such as obesity, type 2 diabetes mellitus, ischemic heart disease, IBD, IBS, hyperlipidemia and proxies for underlying cardiovascular disease such as history of any anti-hypertensive drugs, low-dose aspirin). These incidence rates were adjusted for age and sex. This study was not set up to assess a potential causal association between these drugs and the outcomes and should not be interpreted as such.

The incidence of colorectal cancer was higher in the cohort that included incident and prevalent hypertension subjects, whereas for the other specific outcomes the incidence rates were comparable between the incident cohort and the incident plus prevalent cohort. The incidence rates of all outcomes were higher in the incident colonoscopy cohort than in the incident cohort, consistent with the fact that possible reasons leading to a colonoscopy represented a risk factor for the outcomes.

The cumulative 5-year risk of colorectal outcomes varied between 1% and 3% and was higher with older ages, presence of IBD, for males compared to females, a prior colonoscopy and use of any anti-hypertensive drug.

Particularly in the first month or up to 3 months following hypertension diagnosis the incidence rates of outcomes were high, which may at least partially be explained by increased medical attention and diagnostic work up and therefore chance of diagnosis.

10.2 Limitations

Misclassification of the primary endpoint (colorectal hyperplasia) may be present since not all the outcomes were validated. This was the most important reason to consider case validation (see Annex 3 for validation protocol) in IPCI. For CPRD the validity of CRC has been studied previously with a PPV of 98% (sensitivity 92% and specificity 99%) (Dregan et al 2012). Also, in a second study case validation was performed to validate the diagnostic codes for CRC in CPRD with evidence that the validity of CRC diagnoses is high: over 95% of computer-recorded incident CRC diagnoses were validated by review of original medical records or by confirmation of the diagnosis by the GP (Garcia-Rodriguez and Huerta-Alvarez, 2001). Without case validation in other databases, the incidence of colorectal neoplasms may be underestimated due to inclusion of false positive cases. In SIDIAP the sensitivity of CRC codes was shown to be high (87.7%) (Garcia-Gil et al 2014) In addition, the date of the outcome could be misclassified. Misclassification in IPCI was mitigated by considering the date of recording of first symptoms leading to a colorectal neoplasm diagnosis (e.g. melena, change in stool frequency and/or consistency, colonoscopy referral); if this information was

not available/retrievable, the date of diagnosis was used. This way, a risk of potential misclassification of the date of outcome was reduced.

As the risk of CRC increases with higher age, reaching age of 80 years was considered as censoring point. This is relevant as subjects 80 years and older are more likely to be moved to a nursing home and thus the ability to capture outcomes reliably in this age group is uncertain. Generalization of the results is therefore not possible for hypertensive subjects aged 80 or older.

Berkson bias is likely present in the incident hypertensive cohort given the fact that patients are routinely followed in medical care. In addition, the use of antiplatelets medication may enhance the likelihood of colorectal bleeding, for which they might undergo a colonoscopy. The chance of finding any of the primary endpoints is therefore higher in a hypertension cohort as compared to the general population. Therefore, an analysis as well on the cohort of hypertensive subjects who had a 'negative' colonoscopy was performed. Only in a population with a recorded 'negative' colonoscopy one can be certain that they were event-free at a certain point in time. At cohort entry it may very well occur that subjects do have polyps, but have not yet been investigated/diagnosed since they are asymptomatic. To identify true 'incident' colorectal polyps, a subject must have had a 'negative' colonoscopy at an earlier point in time. Also, the analysis of incidence of outcomes per period since the diagnosis of hypertension showed that in the initial 3 months hypertensive subjects have a high incidence of the outcomes; this is much more likely because of the diagnostic routine care rather than an actual increase because of the hypertension.

As mentioned above, it cannot be certain that at cohort entry patients are 'free' of any colon hyperplasia or polyps, since for this evaluation/conclusion reports from colonoscopy are required. However, since many people will not have undergone a colonoscopy, only those with documented colorectal hyperplasia were excluded, but subjects with 'silent/unknown hyperplasia' will falsely still remain. Therefore, a cohort with hypertension subjects that had a 'negative' colonoscopy was included. Subsequently, it was observed that results in general were similar to the incident cohort, although the incidence rates of the outcomes in the incident colonoscopy cohort were higher. This is expected as the underlying population in this cohort all had a higher a-priori chance of the outcome, as they had some indication to undergo a colonoscopy and is a selected group of patients. From prior studies it is known that a colonoscopy screening population is different from the general population with regards to prevalence of colorectal outcomes (Brenner et al 2014).

Screening and surveillance programs for colorectal cancer, for instance by means of examining occult blood in feces or by sigmoidoscopy or colonoscopy, are aimed to identify subjects at an early stage of colorectal cancer development: thus, at the stage of colon polyps. Nationwide colorectal cancer screening has been started in European countries, including Italy, the United Kingdom, the Netherlands and Spain (Schreuders et al 2015).

- In Italy, nationwide screening (mostly done by Fecal Immunochemical Testing [FIT]; but in some regions in Northern Italy done by sigmoidoscopy), (Giorgi Rossi et al 2015, Reggiani-Bonetti et al 2013,, Bonetti et al 2013) roll-out phase is ongoing. When these programs have been started differs per region, and to which extent screening areas are covered by HSD is something to explore.

- In the United Kingdom, a phased nationwide screening program is implemented since 2005-2006, using the guaiac based test for fecal occult blood (gFOBT), the roll-out is complete. In 2014 the screening has switched to FIT testing, which is more sensitive to detect adenoma and cancer than gFOBT even in a single sample.
- In the Netherlands, since January 2014 a phased nationwide screening program was started using the FIT test, the roll-out phase of screening is still ongoing.
- In Spain, a FIT-based screening program in certain regions of Spain was started in 2000, the roll-out phase of screening is still ongoing (Ascunce et al 2010).

The detection rate of polyps increases by screening programs, which became visible after calculating the incidence rates of the endpoints separately over calendar time. A decrease in the incidence of CRC was observed after 2004-2005 in CPRD and HSD and also an increase in the incidence of polyps between 2008 and 2011 particularly in IPCI, HSD and SIDIAP. This may be the result of implementation of colorectal screening strategies.

A non-hypertensive population was not included in the study and therefore it was not possible to do a direct comparison between the hypertension and a non-hypertension population However, standardizing the observed incidence rates with a standard European reference population allows generalizing the overall incidence rates of outcomes in each of the cohorts. However the Berkson bias (differential diagnostic work-up between both populations) cannot be ruled out.

None of the databases participating in the current study is able to identify actual drug intake. This is a limitation of the study.

A limitation of this study is the completeness of available data. Some of the blood pressure measurements or follow-up of hypertensive patients in the databases may have been missed. In spite of this the majority of patients entered the cohort based on a hypertension diagnosis code, plus an elevated blood pressure measurement, plus a prescription. Recording of smoking status and alcohol use is likely underreported in the databases; as it was observed that between 25% and 71% of study subjects in the incident cohort did not have a measurement of smoking. Excess alcohol intake use is only recorded when a subject may seek contact to the GP for problems related to alcohol use or when rehabilitation is desired by the patient, being most likely alcohol abuse. A similar explanation holds for smoking status; if subjects seek help to quit or diminish smoking this is likely to be recorded, as it will be recorded regularly for patients with COPD, asthma or other lung diseases.

The detail of definitions used in this study for the colorectal outcomes may not reflect coding in clinical practice; at colonoscopy either nothing, or a polyp (including non-adenoma polyps and adenomas) or colorectal cancer is found. The separation of the outcomes into cysts and benign neoplasms has been done based on the clinical study SPP100A2418, however, these outcomes did not occur often in the databases, as was expected.

10.3 Interpretation

The observation that the incidence of colorectal polyps and cancer was higher in subjects with the comorbid conditions present (such as obesity, T2DM, use of drugs) has been observed in prior studies (Batty et al 2003, Othman and Zin 2008, Pelucchi et al 2010, Stocks et al 2008,

Kim et al 2007) Also, the findings that well acknowledged risk factors for CRC such as IBD (Eaden et al 2001, Herrinton et al 2012) alcohol use, obesity and T2DM (Ahmed et al 2006, Esposito et al 2012) were associated with the outcome in this study, are reassuring about the validity of the study results. The PRAC did not request a direct comparator, but in order to put the data in perspective a standard reference population (comprising hypertensive subjects as well) was used to generalize the incidence rates of outcomes based on a sex- and age-distribution from the European standard population. From these analyses it was observed that the standardized incidence rates varied between 65 to 72 per 100,000 person-years in the incident cohort. The CRC incidence rates vary worldwide (based on data from cancer registries) between 4.1-59 per 100,000 person-years among males and 3.6-39.5 among females (Center et al 2009) The crude incidence rate of CRC based on the European Globocan cancer registry is 31.3 per 100,000 person-years (WHO (2012) Standardizing our hypertensive population to an age-and sex-distribution of a European standard population shows that the incidence rate of CRC is higher in the hypertensive population compared to a reference population. Though comparison of the SIRs in the current study (sex- and age-standardized) with the age-standardized incidence rates of the countries may not be ideal (since the rate from the Globocan cancer registry is not standardized by sex), however, it gives some insight in the higher incidence of outcomes in hypertensive populations, if the difference in incidences is not explained by differential sex distributions between our study populations and the Globocan population (Table 10-1).

Table 10-1 Standardized Incidence Rates of Colorectal Cancer by country

Country	Globocan (WHO, 2012)		Current Study	
	Crude rate per 100,000 person-years	Age standardised rate per 100,000 person-years	Database	Age-sex standardized rate per 100,000 person-years (standardized to Eurostat)
EU-28	68.0	31.3		
UK	64.9	30.2	CPRD	64.85
Italy	78.9	33.9	HSD	72.38
NL	83.3	40.2	IPCI	60.09
Spain	68.9	33.1	SIDIAP	72.23

A post-hoc analysis of the STOP-Hypertension-2 trial, which enrolled elderly patients with severe hypertension (systolic blood pressure of ≥180 mmHg and/or diastolic blood pressure ≥ 105 mmHg), compared the number of observed cancer cases among the 6,614 enrolled patients with the number of expected cancer cases based on the Swedish background rate (Lindholm et al 2001) The standardized incidence ratio for colon cancer was 1.10 (95%CI: 0.86-1.40) and for rectum cancer 0.99 (95%CI: 0.67-1.41). The study population in this trial is older (mean age 76 years) and a larger proportion is females (67%) compared to our study. It demonstrates that even in patients with more severe hypertension and with older age, the number of observed colon or rectum cancers was not more than what was expected based on the underlying background rate of the source population from Sweden (Lindholm et al 2001).

Many studies so far have investigated the risk of separate components of the metabolic syndrome or the joint effect of the components from metabolic syndrome on colorectal adenomas (Kim et al 2007) and colorectal cancer (Ahmed et al 2006, Aleksandrova et al 2011, Pelucchi et al 2010, Stocks et al 2008, Sturmer et al 2006, Esposito et al 2012). From these

studies, the conclusions vary with different components of the metabolic syndrome considered to be most important for an increased risk of CRC. Some report hypertension and being overweight (Pelucchi et al 2010); having at least 3 components of the metabolic syndrome (Ahmed et al 2006); abdominal obesity and abnormal glucose metabolism (Aleksandrova et al 2011); obesity, hypertension and hyperglycemia (Stocks et al 2008); overweight and diabetes (Sturmer et al 2006) as risk factors. However, it is not possible to separate risks of the different components into single entities. Presence of comorbid conditions and components in the metabolic syndrome confound the risk of cancer and the combined presence of several risk factors may constitute a higher risk than separate risk factors. Also, in other studies hypertension was not associated with an increased risk of colon cancer (Batty et al 2003, Negri et al 1999, Grove et al 1991, Sturmer et al 2006). It was observed that incidences of the colorectal outcomes were higher in hypertensive patients with presence of the comorbid conditions such as ischemic heart disease, cerebrovascular disease, IBD, IBS, cholecystectomy, hyperlipidemia, type 2 diabetes mellitus, excessive alcohol use, smoking, use of low-dose aspirin, use of any anti-hypertensive drug, with increasing age and for males. These observations are in line with the results reported in literature as described above.

The screening for CRC has been implemented in the United Kingdom since 2005-2006, since 2000 in Spain and in Italy the timing differs per region. This is probably reflected in the decreasing incidence rates of CRC after 2005 in CPRD and HSD from (126.1 per 100,000 Person-Years in 2005 to 37.2 per 100,000 Person-Years in 2013 in CPRD; and 156.6 per 100,000 Person-Years in 2005 to 115.5 per 100,000 Person-Years in 2013 in HSD). Consistently with the screening outcomes it was noticed an increase in the incidence of polyps in all countries between 2008 and 2013. This is probably an effect of the CRC screening programs (Giorgi Rossi et al 2015, Reggiani-Bonetti et al 2013, Schreuders et al 2015, Ascunce et al 2010) The fact that decrease in incidence of CRC in Spain was not observed despite the implementation of the screening programme in 2000 in Spain, might be due to the fact that data from SIDIAP was available since 2007, long after introduction of the screening.

Since colorectal polyps may be present asymptomatically, studies that investigate the occurrence of polyps are mainly prevalence studies, as one can never be certain about the ‘incidence’ or ‘new development’ of polyps in subjects. It is therefore difficult to compare results from our study with others.

10.4 Generalizability

CRC is a fairly clear outcome of which some information from cancer registries about the occurrence and incidence of the disease in the separate countries is available. Although the crude incidence rates may vary, given the potential differential sex- and age-distribution in each database, the incidence rates of CRC in the incident cohort are consistent across the databases (Table 9-3). The fact that the incidence rates of CRC across databases are indeed comparable across the data sources, is also seen in the Standardized Incidence Rates when using both the Eurostat population and the CPRD population to standardize the observed incidence rates to. Differences in incidence rates of CRC became smaller by standardizing to a common population.

A cohort study conducted in CPRD assessed the association between use of angiotensin receptor blockers and cancer, which included CRC (Azoulay et al 2012). The study cohort

included subjects that were prescribed an antihypertensive agent between January 1995 and December 2008. The proportion of subjects entering the cohort by use of the separate classes is comparable with the proportion of subjects on drugs in the incident cohort of CPRD in our study (e.g. 4.2% on angiotensin receptor blockers vs 3.1% in our study; 49.2% on diuretics and/or beta-blockers vs 26.8% on diuretics and 24.8% on beta-blockers in our study; 25.7% of the subjects was exposed to ACE-inhibitors at cohort entry vs 17.7% in our study). This emphasizes the validity and generalizability of the data that was obtained within CPRD. It also demonstrates that the results seen for the other databases, which are in line with CPRD, may be generalizable to other studies including hypertensive subjects.

As the risk of CRC increases with higher age, reaching age of 80 years was considered as censoring point. This is relevant as subjects 80 years and older are more likely to be moved to a nursing home and thus the underlying source population in the databases within this age group is not well defined. Generalization of the results is therefore not possible for hypertensive subjects aged 80 or older.

11 Other information

Not applicable.

12 Conclusion

Incidence rates of colorectal outcomes (polyps, cysts, benign neoplasm and cancer) were consistent in the four European data sources especially after standardization; rates of benign neoplasm and cysts were low, and most events were due to polyps and CRC. Incidence rates of the outcomes (adjusted for age and sex) were higher with increasing age, in males, in presence of comorbid diseases (such as obesity, type 2 diabetes mellitus, ischemic heart disease, IBD, IBS, hyperlipidemia) and proxies for cardiovascular disease such as prior use of drugs (any anti-hypertensive drugs, low-dose aspirin).

The incidence of colorectal cancer was higher in the cohort that included incident and prevalent hypertension subjects, whereas excluding prevalent hypertensive patients had less impact on the rate of the other outcomes. The incidence rates of all outcomes were higher in the incident colonoscopy cohort than in the incident cohort, consistent with the fact that a colonoscopy was a risk factor for the outcomes.

The cumulative 5-year risk of colorectal outcomes varied between 1% and 3% and was higher with older age, with presence of IBD, in males, with a prior colonoscopy positive for any of the study outcomes, and with proxies for cardiovascular co-morbidity such as prior use of any anti-hypertensive drug.

Particularly in the first month or up to 3 months following hypertension diagnosis the incidence rates of outcomes were high, which may be explained by increased medical attention and opportunities for diagnosis.

13 References

[Eurostat (2013)] Population on 1 January 2013 by five years age group and sex (Internet) Available from: <http://ec.europa.eu/eurostat/web/population-demography-migration-projections/population-data/database> (Accessed 16th March 2015).

[WHO (2013)] Collaborating Centre for Drug Statistics Methodology, Guidelines for ATC classification and DDD assignment (Internet) Available from: <http://www.whocc.no/atcddd/> (Accessed September 23, 2013).

[Ahmed RL, Schmitz KH, Anderson KE, et al. (2006)] The metabolic syndrome and risk of incident colorectal cancer. Cancer; 107(1): 28-36.

[Aleksandrova K, Boeing H, Jenab M, et al. (2011)] Metabolic syndrome and risks of colon and rectal cancer: the European prospective investigation into cancer and nutrition study. Cancer Prev Res (Phila); 4(11): 1873-83.

[Angeli F, Reboldi G, Mazzotta G, et al. (2012)] Safety and efficacy of aliskiren in the treatment of hypertension and associated clinical conditions. Curr Drug Saf; 7(1): 76-85.

[Ascunce N, Salas D, Zubizarreta R, et al. (2010)] Cancer screening in Spain. Ann Oncol; 21(3).

[Azoulay L, Assimes TL, Yin H, et al. (2012)] Long-term use of angiotensin receptor blockers and the risk of cancer. PLoS One; 7(12): e50893.

[Batty GD, Shipley MJ, Marmot MG, et al. (2003)] Blood pressure and site-specific cancer mortality: evidence from the original Whitehall study. Br J Cancer; 89(7): 1243-7.

[Bolíbar B, Fina Avilés F, Morros R, et al (2012)] SIDIAP database: electronic clinical records in primary care as a source of information for epidemiologic research Med Clin (Barc); 138(14):617-21.

[Booth N (1994)] What are the Read Codes? Health Libr Rev; 11(3): 177–182.

[Brenner H, Hoffmeister M & Jansen L (2014)] Comparisons of colorectal cancer mortality between screening participants and the general population are strongly biased unless an incidence-based mortality approach is used. J Clin Epidemiol; 67(2): 184-9.

[Burke TA, Sturkenboom MC, Lu SE, et al. (2006)] Discontinuation of antihypertensive drugs among newly diagnosed hypertensive patients in UK general practice. J Hypertens; 24(6): 1193-200.

[Cazzola M, Puxeddu E, Bettoncelli G, et al. (2011)] The prevalence of asthma and COPD in Italy: a practice-based study. Respir Med; 105(3): 386-91.

[Center MM, Jemal A, Smith RA, et al. (2009)] Worldwide variations in colorectal cancer. CA Cancer J Clin; 59(6): 366-78.

[Chisholm J (1990)] The Read clinical classification. BMJ; 300(6732): 1092.

[Coloma PM, Schuemie MJ, Trifiro G, et al. (2011)] Combining electronic healthcare databases in Europe to allow for large-scale drug safety monitoring: the EU-ADR Project. Pharmacoepidemiol Drug Saf; 20(1): 1-11.

[Cricelli C, Mazzaglia G, Samani F, et al. (2003)] Prevalence estimates for chronic diseases in Italy: exploring the differences between self-report and primary care databases. J Public Health Med; 25(3): 254-7.

[Dregan A, Moller H, Murray-Thomas T, et al. (2012)] Validity of cancer diagnosis in a primary care database compared with linked cancer registrations in England. Population-based cohort study. Cancer Epidemiol; 36(5): 425-9.

[Eaden JA, Abrams KR & Mayberry JF (2001)] The risk of colorectal cancer in ulcerative colitis: a meta-analysis. Gut; 48(4): 526-35.

[Esposito K, Chiodini P, Colao A, et al. (2012)] Metabolic syndrome and risk of cancer: a systematic review and meta-analysis. Diabetes Care; 35(11): 2402-11.

[Filippi A, Vanuzzo D, Bignamini AA, et al. (2005)] The database of Italian general practitioners allows a reliable determination of the prevalence of myocardial infarction. Ital Heart J; 6(4): 311-4.

[Fisher ND, Jan Danser AH, Nussberger J, et al. (2008)] Renal and hormonal responses to direct renin inhibition with aliskiren in healthy humans. Circulation; 117(25): 3199-205.

[Garcia-Gil M, Elorza JM, Banque M, et al. (2014)] Linking of primary care records to census data to study the association between socioeconomic status and cancer incidence in Southern Europe: a nation-wide ecological study. PLoS One; 9(10).

[Garcia-Gil Mdel M, Hermosilla E, Prieto-Alhambra D, et al. (2011)] Construction and validation of a scoring system for the selection of high-quality data in a Spanish population primary care database (SIDIAP). Inform Prim Care; 19(3): 135-45.

[Garcia-Rodriguez LA & Huerta-Alvarez C (2001)] Reduced risk of colorectal cancer among long-term users of aspirin and nonaspirin nonsteroidal antiinflammatory drugs. Epidemiology; 12(1): 88-93.

[Giorgi Rossi P, Vicentini M, Sacchettini C, et al. (2015)] Impact of Screening Program on Incidence of Colorectal Cancer: A Cohort Study in Italy. Am J Gastroenterol; 110(9): 1359-66.

[Grossman E, Messerli FH, Boyko V, et al. (2002)] Is there an association between hypertension and cancer mortality? Am J Med; 112(6): 479-86.

[Grove JS, Nomura A, Severson RK, et al. (1991)] The association of blood pressure with cancer incidence in a prospective study. Am J Epidemiol; 134(9): 942-7.

[Healey JS & Connolly SJ (2003)] Atrial fibrillation: hypertension as a causative agent, risk factor for complications, and potential therapeutic target. Am J Cardiol; 91(10A): 9G-14G.

[Herrinton LJ, Liu L, Levin TR, et al. (2012)] Incidence and mortality of colorectal adenocarcinoma in persons with inflammatory bowel disease from 1998 to 2010. Gastroenterology; 143(2): 382-9.

[Hong JL, Meier CR, Sandler RS, et al. (2013)] Risk of colorectal cancer after initiation of orlistat: matched cohort study. BMJ; 347: f5039.

[Jick H, Jick SS & Derby LE (1991)] Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. BMJ; 302(6779): 766-8.

[Jick SS, Kaye JA, Vasilakis-Scaramozza C, et al. (2003)] Validity of the general practice research database. Pharmacotherapy; 23(5): 686-9.

[Kim JH, Lim YJ, Kim YH, et al. (2007)] Is metabolic syndrome a risk factor for colorectal adenoma? Cancer Epidemiol Biomarkers Prev; 16(8): 1543-6.

[Lamberts H, Wood M & Hofmans-Okkes IM (1992)] International primary care classifications: the effect of fifteen years of evolution. Fam Pract; 9(3): 330-9.

[Lever AF, Hole DJ, Gillis CR, et al. (1999)] Is cancer related to hypertension or to its treatment? Clin Exp Hypertens; 21(5-6): 937-46.

[Lieberman DA, Rex DK, Winawer SJ, et al. (2012)] Guidelines for colonoscopy surveillance after screening and polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer. Gastroenterology; 143(3): 844-57.

[Lindgren AM, Nissinen AM, Tuomilehto JO, et al. (2005)] Cancer pattern among hypertensive patients in North Karelia, Finland. J Hum Hypertens; 19(5): 373-9.

[Lindholm LH, Anderson H, Ekbom T, et al. (2001)] Relation between drug treatment and cancer in hypertensives in the Swedish Trial in Old Patients with Hypertension 2: a 5-year, prospective, randomised, controlled trial. Lancet; 358(9281): 539-44.

[Musini VM, Fortin PM, Bassett K, et al. (2008)] Blood pressure lowering efficacy of renin inhibitors for primary hypertension. Cochrane Database Syst Rev(4): CD007066.

[Negri E, Bosetti C, La Vecchia C, et al. (1999)] Allergy and other selected diseases and risk of colorectal cancer. Eur J Cancer; 35(13): 1838-41.

[Othman NH & Zin AA (2008)] Association of colorectal carcinoma with metabolic diseases; experience with 138 cases from Kelantan, Malaysia. Asian Pac J Cancer Prev; 9(4): 747-51.

[Pelucchi C, Negri E, Talamini R, et al. (2010)] Metabolic syndrome is associated with colorectal cancer in men. Eur J Cancer; 46(10): 1866-72.

[Rahme E, Barkun AN, Adam V, et al. (2004)] Treatment Costs to Prevent or Treat Upper Gastrointestinal Adverse Events associated with NSAIDs. Drug Safety; 27(13): 1-21.

[Ramos R, Ballo E, Marrugat J, et al. (2012)] Validity for use in research on vascular diseases of the SIDIAP (Information System for the Development of Research in Primary Care): the EMMA study. Rev Esp Cardiol (Engl Ed); 65(1): 29-37.

[Reggiani-Bonetti L, Di Gregorio C, Pedroni M, et al. (2013)] Incidence trend of malignant polyps through the data of a specialized colorectal cancer registry: clinical features and effect of screening. Scand J Gastroenterol; 48(11): 1294-301.

[Schmieder RE, Philipp T, Guerediaga J, et al. (2009)] Long-term antihypertensive efficacy and safety of the oral direct renin inhibitor aliskiren: a 12-month randomized, double-blind comparator trial with hydrochlorothiazide. Circulation; 119(3): 417-25.

[Schreuders EH, Ruco A, Rabeneck L, et al. (2015)] Colorectal cancer screening: a global overview of existing programmes. Gut; 64(10): 1637-49.

[Stocks T, Lukanova A, Johansson M, et al. (2008)] Components of the metabolic syndrome and colorectal cancer risk; a prospective study. Int J Obes (Lond); 32(2): 304-14.

[Sturkenboom MC, Dieleman JP, Picelli G, et al. (2008)] Prevalence and treatment of hypertensive patients with multiple concomitant cardiovascular risk factors in The Netherlands and Italy. J Hum Hypertens; 22(10): 704-13.

[Sturmer T, Buring JE, Lee IM, et al. (2006)] Metabolic abnormalities and risk for colorectal cancer in the physicians' health study. Cancer Epidemiol Biomarkers Prev; 15(12): 2391-7.

[Van Soest EM, Van Rossum LG, Dieleman JP, et al. (2008)] Proton pump inhibitors and the risk of colorectal cancer. Am J Gastroenterol; 103(4): 966-73.

[Van Staa TP, Card T, Logan RF, et al. (2005)] 5-Aminosalicylate use and colorectal cancer risk in inflammatory bowel disease: a large epidemiological study. Gut; 54(11): 1573-8.

[Vlug AE, Van Der Lei J, Mosseveld BM, et al. (1999)] Postmarketing surveillance based on electronic patient records: the IPCI project. Methods Inf Med; 38(4-5): 339-44.

[Voordouw AC, Sturkenboom MC, Dieleman JP, et al. (2004)] Annual revaccination against influenza and mortality risk in community-dwelling elderly persons. JAMA; 292(17): 2089-95.

[Watanabe Y, Ozasa K, Ito Y, et al. (2005)] Medical history of circulatory diseases and colorectal cancer death in the JACC Study. J Epidemiol; 15 Suppl 2: S168-72.

[WHO (2013)] Classification of Diseases (Internet) Available from: <http://www.who.int/classifications/icd/en/> (Accessed September 23, 2013).

[WHO (2012)] Globocan 2012. Estimated Cancer Incidence, Mortality and Prevalence Worldwide in 2012 (Internet) Available at: <htpp//globocan.iarc.fr/Pages/online.aspx> (Accessed September 22, 2015).

[Yang YX, Hennessy S & Lewis JD (2004)] Insulin therapy and colorectal cancer risk among type 2 diabetes mellitus patients. Gastroenterology; 127(4): 1044-50.

Annex 1 – Colorectal outcomes Definitions

Colorectal Outcomes

Outcomes of interest

- Cysts = Colorectal cysts: cysts of the colon, rectum and perineum, other than pilonidal sinus or cysts
- Polyps = Colorectal polyps (with/without hyperplasia and irrespective of the type or size of polyp): Polyps in the colon or rectum, these are precursor lesions of subsequent colorectal adenomas and cancer.
- Benign Neoplasms = Benign Colorectal neoplasms
- Colorectal Cancer = Malignant Colorectal Neoplasms with or without metastases. Malignant lesions of the colon, rectum, such as adenomas and adenocarcinomas.

Part of the case definition of colorectal hyperplasia was: any ulcerations or colonic bleeding related or due to colorectal malignancy or cell dysplasia. Carcinoma in situ was therefore also included.

Tumor size and tumor type (adenocarcinoma or other more rare histopathological types of cancer) was not taken into account, neither polyp size or polyp type (such as a distinction between tubular, villous or sessile polyps).

The following labels codes of disease have been mapped.

Label	ICD9CM	ICD-10	READ codes	ICPC	EventType
COLORECTAL POLYPS					
Colonic Polyps	V18.51 V12.72	K63.5	B713.11 B713.13 B713811 J578.00 J578.11	D78	POLYP
Bengin neoplasm of rectum and anal canal (adenomatous anal and rectal polyp)	211.4				POLYP
Anal and rectal polyp	569.0	K62.1 D12.8	J570.00 J570000 J570100 J570z00		POLYP
Adenomas			BB5N.00 BB5N.11 BB5N000 BB5Nz00		POLYP

Polypectomies			68W2200 771G400 7722.11 7722.12 7726212 771G400 7M0Ey11		POLYP
Adenomatous polyps			BB5L.00 BB5L000 BB5L011 BB5L100 BB5L200 BB5L300 BB5Lz00 BB5N.00 BB5N.11 BB5N000 BB5N200 BB5N211		POLYP
COLORECTAL CYSTS AND HYPERPLASIA					
Cyst of perineum			7C16500 7D08.00	D78	CYST
Carcinoma in situ		D01.0 D01.1 D01.2 K38.0	B803.00 B803000 B803100 B803200 B803300 B803400 B803600 B803700 B803z00 B804.00 B804000 B804100 B804z00		CYST
Dysplasia of anus	569.84 569.85				CYST

COLORECTAL BENIGN NEOPLASM					
Benign neoplasm of colon	211.3	D12x	B718300 B718400 B713.00 B713.12 B713000 B713100 B713200 B713300 B713400 B713600 B713700 B713800 B713900 B713z00 B714.00 B714000 B714100 B714111 B714z00	D78	BNEOPL
Benign neoplasm of rectum/anal canal	211.4				BNEOPL
COLORECTAL CANCER					
Malignant neoplasm colon/rectum		C19x	B575.00 B575z00 B902.00 68W2400	D75	CRC
Malignant neoplasm colon	153x	C18x	B13..00 B130.00 B131.00 B132.00 B133.00 B134.00 B134.11 B136.00 B137.00 B138.00 B13y.00 B13z.00 B13z.11 B575000 B902400 ZV10014		CRC

Malignant neoplasms rectum	154x	C19x	B14..00 B140.00 B141.00 B141.11 B141.12 B14y.00 B14z.00 B575100 ZV10017 B902500		CRC
Adenocarcinomas			BB5M.00 BB5M000 BB5M100 BB5Mz00 BB5N100		CRC

x means all codes falling under this category

Annex 2 – Hypertension definition

Arterial Hypertension

Definition of arterial hypertension

Arterial hypertension was identified via diagnosis codes, prescriptions/dispensings and blood pressure measurements (elevated blood pressure defined as systolic blood pressure ≥ 140 mmHg and/or diastolic ≥ 90 mmHg).

Complications of hypertension

Complications of hypertension was considered as proxy for 'prevalent' hypertension, as the subject most likely has had the diagnosis of arterial hypertension before. Complications of hypertension include hypertensive heart disease, kidney disease, encephalopathy and retinopathy.

The following codes were used for identification of arterial hypertension by Diagnosis Codes:

Label	ICD9CM	ICD-10	READ codes	ICPC	EventType
Essential Hypertension	401.x	I10	G20..00 G200.00 G201.00 G20z.00		HTN
Essentiele hypertensie zonder orgaanbeschadiging				K86x	HTN
Pre-existente hypertensie in zwangerschap				W81.01	HTN
H/O: hypertension			14A2.00		HTN
Hypertension monitoring			662..12		HTN
Good hypertension control			6627.00		HTN
Poor hypertension control			6628.00		HTN
Hypertension:follow-up default			6629.00		HTN
Moderate hypertension control			662b.00		HTN
Hypertension six month review			662c.00		HTN
Hypertension annual review			662d.00		HTN
Hypertension treatm. started			662F.00		HTN
Hypertensive treatm.changed			662G.00		HTN
Hypertension treatm.stopped			662H.00		HTN
On treatment for hypertension			662O.00		HTN
Hypertension 9 month review			662P000		HTN
Trial reduction of antihypertensive therapy			662q.00		HTN

Trial withdrawal of antihypertensive therapy			662r.00		HTN
Lifestyle advice regarding hypertension			67H8.00		HTN
High cost hypertension drugs			7Q01.00		HTN
Other specified high cost hypertension drugs			7Q01y00		HTN
Antihypertensive therapy			8B26.00		HTN
Patient on maximal tolerated antihypertensive therapy			8BL0.00		HTN
Hypertension clinical management plan			8CR4.00		HTN
Hypertension treatment refused			8I3N.00		HTN
Hypertensive disease			G2...00		HTN
BP - hypertensive disease			G2...11		HTN
Systolic hypertension			G202.00		HTN
Diastolic hypertension			G203.00		HTN
Hypertension NOS			G20z.11		HTN
Secondary Hypertension			G24..00 G240.00 G240z00 G241.00 G241z00 G244.00 G24z.00 G24zz00		HTN
Stage 1 hypertension			G25..00 G25..11		HTN
Severe hypertension			G26..00 G26..11		HTN
Hypertension resistant to drug therapy			G27..00		HTN
Stage 2 hypertension (NICE - Nat Ins for Hth Clin Excl 2011)			G28..00		HTN
Other specified hypertensive disease			G2y..00		HTN
Hypertensive disease NOS			G2z..00		HTN
Hypertensive crisis			G672.11		HTN
[X]Hypertensive diseases			Gyu2.00		HTN
[X]Other secondary hypertension			Gyu2000		HTN
[X]Hypertension secondary to other renal disorders			Gyu2100		HTN

	CODES FOR COMPLICATIONS OF HYPERTENSION				
Hypertensive heart disease	402x	I11x	G21..00 G210.00 G210000 G210100 G211.00 G211000 G211100 G21z.00 G21z000 G21z011 G21z100 G21zz00	k87x	COMPLHTN
Hypertensive heart and renal disease	404x	I13x	G23..00 G230.00 G231.00 G232.00 G233.00 G23z.00		COMPLHTN
Cong heart fail, hypertens, age, diab, stroke 2 risk score			38DE.00		COMPLHTN
Hypertensive encephalopathy	437.2	I67.4	G672.00		COMPLHTN
Hypertensive retinopathy	362.11		F421300	F83.02	COMPLHTN
Hypertensive renal disease	403x 405.01 405.11 405.91	I12x I15x	G22..00 G220.00 G221.00 G222.00 G22z.00 G22z.11 G240000 G241000 G24z000		COMPLHTN

x means all codes falling under this category

Annex 3 – Validation of outcomes

Event Definition:

Colorectal cysts: Cysts in the colon, sigmoid, perineum or rectum occur very rarely. Piloinidal sinus cyst occur more often, however, these derive from a different origin (hair growth) and thus are not part of the case definition.

Colorectal polyps: Polyps in the colon or rectum, these are precursor lesions of subsequent colorectal adenomas and cancer. A polyp of the colon refers to a protuberance into the lumen from the normally flat colonic mucosa. Polyps are usually asymptomatic but may ulcerate and bleed, cause tenesmus if in the rectum, and, when very large, produce intestinal obstruction. Colonic polyps are usually classified as non-neoplastic, hamartomatous, neoplastic (adenomas and carcinomas), serrated (which can be neoplastic or non-neoplastic), and submucosal (which can be neoplastic or non-neoplastic). Adenomas of the colon and rectum were also considered in this group.

Benign neoplasms: Benign neoplasms of the colon, sigmoid or rectum other than polyps and adenomas. Also ulcerations or cell dysplasia or hyperplasia other than polyps are considered as a benign neoplasm.

Colorectal Cancer: Colorectal cancer is a malignant disease where there is a neoplasm located in the colon, sigmoid or rectum. CRC is common and highly lethal disease. It is ranked 2nd and 3rd most common type of cancer among men and women in the Netherlands. The lifetime risk to develop CRC is 5%-6% for adults. First symptoms leading to CRC diagnosis are often occult bleeding, however most subjects remain asymptomatic for a long period. This means that diagnosis is often in a late stage of disease.

The normal sequence of CRC development is from dysplasia, to polyp, to adenoma to colorectal cancer. For this reason, every patient should be considered to be reviewed for all outcomes.

Case Classification:

- **Definite Case**: It is a definite case. This case fulfills with the definition of outcomes provided in this document or other additional supporting information that strongly suggests/confirms the diagnosis and according to the criterion of the assessor that reviewed the patient file it is a definite case. In General Practitioner (GP) databases, if the confirmation from a specialist or a letter from a hospital is available, it is considered strong evidence to classify the subject as a definite case.
- **Probable case**: there is some indication of the event but there is no strong evidence to support the diagnosis. For example, there is a code but there is no additional evidence to confirm it as a case. In GP databases, the presence of the code or free text mentioning the event but without confirmation from the specialist, hospital letter or the criteria mentioned in the definition is considered as a probable case.

- **Doubtful case:** the information found in the electronic medical record suggests for an alternative diagnosis, there are doubts about the diagnosis and not additional information is available to definitely rule-out or classify it in the other categories.
- **No case:** An event that does not fulfill with the definition of PC provided in this document and according to the criteria of the medical doctor that reviewed the patient file this is not a case. Prevalent cases are considered as no cases.

Criteria for Colorectal polyp: Diagnosis of a colorectal polyp or adenoma with additional evidence consisting of:

- Report from colonoscopy
- Histological diagnosis of polyp or adenoma (biopsies, polypectomy)

Criteria for Colorectal cyst: Recording of a cyst in colon, rectum or sigmoid.

- Report from colonoscopy
- Radiological confirmation (CT scan, MRI, ultrasound)

Criteria for Benign neoplasm: In case of a benign tumor (**not** malignant), which is not classified as a colorectal polyp. So it does **not** include adenomas or polyps. However, it does include benign neoplasms of colon and of rectum and anal canal. Carcinoma in situ is considered as colorectal benign neoplasm as well.

Criteria for Colorectal cancer: Diagnosis of colorectal cancer with additional evidence consisting of:

- Histological diagnosis of CRC (biopsies, surgical resection specimen)
- Serological marker (CEA)
- Radiological confirmation (CT scan, PET scan)
- Chemotherapy or radiation therapy administration for CRC
- or death related to CRC.

If the record is a recurrence of the CRC, please record this as a recurrent cancer. If recurrence, and if possible, please locate the first time mentioning of the initial CRC (and do not record it as a recurrent cancer).

General comments

If the date of event is unknown please fill in date: 01-01-1900.

If you know the year of the event, but not the day and not the month, please fill in: 01-06-[year]

If you know the year and month of the event, but not the day, please fill in: 15-[month]-[year]

Event Date is the date on which the event occurred.

Annex 4 - Variable descriptions and possible values

Variable name	Description	Possible values / Calculation
YearHTN	Calendar Year of first diagnosis of Hypertension	Calendar Year of first event HTN
MonthHTN	Calendar Month of first diagnosis of Hypertension	Calendar Month of first event HTN
HistComplHTN	Is there a diagnosis of complicated Hypertension	Yes if there is any event Complhtn with Date <= CohortStart
		No otherwise
AgeHTN	Age at first diagnosis of Hypertension	For Incident cases: Time between BirthDate and CohortStart
		For Prevalent cases: missing value.
AgeStart	Age at start cohort	Time between BirthDate and CohortStart
AgeStartCat	AgeStart categorical	18 <= AgeStart <=44
		45 <= AgeStart <=64
		65 <= AgeStart <=79
PrevInc	Prevalent or Incident Hypertension at CohortStart	Prevalent if CohortStart > date of first event HTN or if an event COMPLHTN is preceding all events of HTN
		Incident if CohortStart = date of first event HTN and no previous event COMPLHTN
BPmeasurement	Is there a measurement of high BP 28 days<=CohortStart<=28 days	Yes if there is a measurement SBP >= 140 or a measurement DBP >= 90 in the period 28 days<=CohortStart<=28 days
		No if there is no such measurement
HypTreatment	Is there hypertension treatment 365days<=CohortStart<=28 days	Yes if there is treatment of diuretics (ATC C03*, C07BA*; C07BB*; C07BG*; C07CA*; C07CB*; C07CG*; C07DA*; C09BA*; C09DA*; C09DX01; C09DX03), beta-blockers (C07*) calcium antagonists (C08*; C09BB*; C09DB*; C09DX01; C09DX03; C09XA53; C09XA54), ACE-inhibitor (C09A*; C09B*), angiotensin receptor blocker (C09C*; C09D*) or renin inhibitor (ATC C09X*) started or ongoing 365days<=CohortStart<=28 days
		No if there is no such treatment
HistDiuretics	Was the patient using Diuretics < CohortStart?	Yes if patient was using diuretics (ATC C03*, C07BA*; C07BB*; C07BG*; C07CA*; C07CB*; C07CG*; C07DA*; C09BA*; C09DA*; C09DX01; C09DX03) < CohortStart
		No if not
HistBetaBl	Was the patient using Beta-blockers < CohortStart?	Yes if patient was using beta-blockers (C07*) < CohortStart
		No if not
HistCalcAnt	Was the patient using Calcium Antagonists < CohortStart?	Yes if patient was using calcium antagonists (C08*; C09BB*; C09DB*; C09DX01; C09DX03; C09XA53; C09XA54) < CohortStart

Variable name	Description	Possible values / Calculation
HistACEInh	Was the patient using ACE-inhibitors < CohortStart?	No if not Yes if patient was using ACE-inhibitor (C09A*; C09B*) < CohortStart
HistATIIAnt	Was the patient using ATII antagonists < CohortStart?	No if not Yes if patient was using diuretics (angiotensin receptor blocker (C09C*; C09D*) < CohortStart
HistReninInh	Was the patient using Renin Inhibitors < CohortStart?	No if not Yes if patient was using or renin inhibitor (ATC C09X*) < CohortStart
HistPolyp	Colorectal polyp < CohortStart	No if not Yes, if date of any Event Polyp < CohortStart Otherwise no
TimePolyp	Time between CohortStart and date of first event POLYP after or at CohortStart	Time between CohortStart and date of first event POLYP after or at CohortStart. Missing (value -1) if no event POLYP after or at CohortStart.
HistCyst	Colorectal cyst < CohortStart	Yes, if date of any Event Cyst < CohortStart Otherwise no
TimeCyst	Time between CohortStart and date of first event CYST after or at CohortStart	Time between CohortStart and date of first event of CYST after or at CohortStart Missing value (-1) if no event CYST after or at CohortStart
HistBneopl	Colorectal benign neoplasm	Yes, if date of any Event Bneopl < CohortStart Otherwise no
TimeBneopl	Time between CohortStart and date of first event BNEOPL after or at CohortStart	Time between CohortStart and date of first event of BNEOPL after or at CohortStart Missing value (-1) if no event BNEOPL after or at CohortStart
TimeCRC	Time between CohortStart and date of first event CRC	Time between CohortStart and date of first event of CRC after or at CohortStart Missing value (-1) if no event CRC after or at CohortStart
HistIBD	Covariable Inflammatory bowel disease <= CohortStart	Yes, if any event of IBD <= CohortStart Otherwise, no
HistHercan	Covariable Hereditary colon cancer syndrome <= CohortStart	Yes, if any event of Hercan <= CohortStart Otherwise, no
HistIBS	Covariable Irritable bowel syndrome <= CohortStart	Yes, if any event of IBS <= CohortStart Otherwise, no
COLONSTATUS	Status of Colonoscopy at CohortStart This variable is only relevant for *CohortHyp.txt*	Value of most recent measurement of COLON before or at CohortStart. Possible values: NEGATIVE POSITIVE UNKNOWN NONE if not any measurement of COLON before or at CohortStart.
TimeToColonNeg	Time between CohortStart and	Time between last date of COLON = NEG

Variable name	Description	Possible values / Calculation
	Colonoscopy NEGATIVE This variable is only relevant in *CohortColon.txt*	before or at CohortStart andCohortStart
HistChol	Covariable Cholecystectomy <= CohortStart	Yes, if any event of CHOL <= CohortStart Otherwise, no
Obesity	Obesity (codes + BMI values)	Yes, if any event of OBES or measurement of BMI with value greater or equal to 30 in the period 365 days <=CohortStart<= 30 days In case an event OBES and measurement BMI on the same day: the code OBES will prevail, even when the BMI value is inconsistent with the code OBES.
UseLDA	Use of Low-dose aspirin 365 days<=CohortStart<=30 days	Yes if patient was using low-dose aspirin (ATC B01AC06) 365 days<=CohortStart<=30 days No if not
HistIschheart	History of ischemic heart disease<= CohortStart	Yes, if any event of Ischheart <= CohortStart Otherwise, no
HistCerebrodis	History of transient ischemic attack or stroke (ischemic and hemorrhagic) <= CohortStart	Yes, if any event of Cerebrodis <= CohortStart Otherwise, no
Hyperlip	Hyperlipidemia (Drugs and Disease Codes) 365 days<=CohortStart<=30 days	Yes if patient was using a drug for hyperlipidemia (ATC C10A*; C10B*) or has any event of HYPERLIP 365 days<=CohortStart<=30 days No if not
Hypertrig	Hypertriglyceridemia (Drugs and Disease Codes) 365 days<=CohortStart<=30 days	Yes if patient was using a drug for Hypertriglyceridemia (ATC C10AB*; C10BA03; C10AB04) or has any event of HYPERTRIG 365 days<=CohortStart<=30 days No if not
Hyperchol	Hypercholesterolemia (Drugs and Disease Codes) 365 days<=CohortStart<=30 days	Yes if patient was using a drug for Hypercholesterolemia (ATC C10A*; C10B*) or has any event of HYPERCHOL 365 days<=CohortStart<=30 days No if not
T2DM	Diabetes Mellitus Type 2 (Drugs and Disease Codes) 365 days<=CohortStart<=30 days	Yes if patient was using a drug for type 2 diabetes mellitus (ATC A10*) or has any event of T2DM 365 days<=CohortStart<=30 days No if not
BMICat	BMI category at CohortStart	Value of measurement of BMI closest to CohortStart in the period of 365 days <= CohortStart<=365 days. If there are BMI measurements at the same time distance before and after CohortStart, the measurement before CohortStart will prevail. If no BMI measurement is available, BMI could also be calculated based on LENGTH and WEIGHT. If a calculated BMI and a reported

Variable name	Description	Possible values / Calculation
		BMI are on the same day, the reported BMI will prevail. Possible values: UNDER if BMI <18 NORMAL if BMI >=18 and < 25 OVER if BMI >=25 and <30 OBESE if BMI >=30 and <35 SEVOB if BMI >=35 UNKNOWN if no measurement of BMI available.
Smoking	SMOKING status	Value of the measurement of SMOKING closest to CohortStart, measured in the period 365 days<=CohortStart<=365 days. Possible values: CURRENT PAST NEVER UNKNOWN if no measurement of SMOKING available. If there are SMOKING measurements at the same time distance before and after CohortStart, the measurement before CohortStart will prevail.
ALCOEXCESS	Excessive alcohol use	Any event of ALCOEXCESS in the period 365 days<=CohortStart<=365 days. Possible values: YES if an event of ALCOEXCESS is present NO if no event of ALCOEXCESS is present
CohortTime	Follow-up time in cohort	CohortEnd-CohortStart