BMJ Open

Cohort profile: The MCC-Spain follow-up on colorectal, breast and prostate cancers. Study design and initial results

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-031904
Article Type:	Cohort profile
Date Submitted by the Author:	24-May-2019
Complete List of Authors:	Alonso, Jessica ; University of Cantabria – IDIVAL, Santander, Spain Molina, Antonio J; Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain. Jiménez-Moleón, Jose Juan; University of Granada, Department of Preventive Medicine and Public Health, School of Medicine; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Pérez-Gómez, Beatriz ; National Center for Epidemiology, Carlos III Institute of Health; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Martin, Vicente; 2- Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain. Moreno, Victor; Catalan Institute of Oncology (ICO-IDIBELL), Cancer Prevention and Control Program; University of Barcelona, Clinical Sciences, Faculty of Medicine Amiano, Pilar; Public Health Division of Gipuzkoa, BioDonostia Research Institute; Consortium for Biomedical Research in Epidemiology and Public Health (CIBER Epidemiología y Salud Pública, CIBERESP) Ardanaz, Eva de Sanjose, Silvia ; L'Hospitalet de Llobregat, Barcelona Salcedo-Bellido, Inmaculada; Universidad de Granada, Preventive Medicine and Public Health+34 Fernandez-Tardon, Guillermo; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Alguacil, Juan; CIBER de Epidemiología y Salud Pública (CIBERESP), Salas, Dolores; Generalitat Valenciana Conselleria de Sanitat Marcos-Gragera, Rafael; Institut Catala d' Oncologia, Epidemiology Unit and Girona Cancer Registry; Universitat de Girona, Research Group on Statistics, Econometrics and Health (GRECS) Chirlaque, Maria Dolores; Murcia Cancer Registry Aragonés, Nuria; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Castaño-Vinyals, Gemma ; Instituto de Salud Global Barcelona Pollán, Marina; Centro Nacional de Epidemiologia. Instituto de Salud Carlos III, Area de Epidemiologia Ambiental y Cancer; CIBER Epidemiologia y Salud Publica (CIBERESP), Kogevinas, Manolis; Instituto de Salud Global de Barcelona, Llorca, Javier; University of Cantabria,
Keywords:	Cohort, Epidemiology < TROPICAL MEDICINE, colorectal cancer, breast

	cancer, prostate cancer, MCC-Spain

Cohort profile: The MCC-Spain follow-up on colorectal, breast and prostate cancers. Study design and initial results

Authors: Jessica Alonso-Molero*[1], Antonio J Molina*[2], Jose J Jiménez-Moleón*[3,4], Beatriz Pérez-Gómez[3,5,6], Vicente Martín[2,3], Víctor Moreno[3,7,8], Pilar Amiano[3,9], Eva Ardanaz[3,10,11], Silvia de Sanjosé[3,7], Inmaculada Salcedo[3,4], Guillermo Fernández-Tardón[3,12], Juan Alguacil[3,13], Dolores Salas[3,14], Rafael Marcos-Gragera[3,15], M Dolores Chirlaque[3,16], Nuria Aragonés[3,17], Gemma Castaño-Vinyals[3,18,19,20], Marina Pollán[3,5,6], Manolis Kogevinas[3,18,19,20], Javier Llorca[1,3].

*Equal contribution

Affiliations:

1- University of Cantabria-IDIVAL, Santander, Spain.
2- Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain.
3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain.
4- Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain.
5- Cancer and Environmental Epidemiology Unit, National Center for Epidemiology, Carlos III Institute of Health, Madrid, Spain.
6- Cancer Epidemiology Research Group, Oncology and Hematology Area, IIS Puerta de Hierro (IDIPHIM), Madrid, Spain.
7- Cancer Prevention and Control Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Spain.
8- Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain.
9- Public Health Division of Gipuzkoa, BioDonostia Research Institute, San Sebastian, Spain.
10- Navarra Public Health Institute, Pamplona, Spain.
11- IdiSNA, Navarra Institute for Health Research, Pamplona, Spain.
12- Oncology Institute, University of Oviedo, Oviedo.
13- Centro de Investigación en Recursos Naturales, Salud y Medio Ambiente (RENSMA), Universidad de Huelva, 21004 Huelva, Spain.
14- Área de Cáncer y Salud Pública, FISABIO-Salud Pública, 46020 Valencia, Spain.
15- Unitat d'Epidemiologia i Registre de Càncer de Girona (UERCG), Pla Director d'Oncologia, Institut Català d'Oncologia, Institut d'Investigaciò Biomèdica de Girona (IdIBGi), Universitat de Girona, Girona, Spain.
16- Department of Epidemiology, Regional Health Authority, IMIB-Arrixaca, Murcia University, Murcia, Spain.
17- Department of Health, Epidemiology Section, Public Health Division, Madrid.
18- ISGlobal, Barcelona, Spain.

19- Universitat Pompeu Fabra (UPF), Barcelona, Spain.

20- IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain.

Correspondence:

Jéssica Alonso-Molero

Facultad de Medicina, Universidad de Cantabria

Avda. Herrera Oria s/n

39011 Santander (Cantabria)

Spain

ORCiD: https://orcid.org/0000-0002-1939-8798

e-mail: alonsomoleroj@gmail.com

Word count: 3144

ABSTRACT

Purpose: From 2016, the MCC-Spain has turned towards the identification of factors associated with cancer prognosis; in this regard, inception cohorts on colorectal, breast and prostate cancers has been assembled using the incident cases originally recruited.

Participants: 2140 new cases of colorectal cancer, 1732 of breast cancer and 1112 of prostate cancer were initially recruited in 12 Spanish provinces; all cancers were incident and pathologically confirmed. Follow-up was obtained for 2097 (98%), 1685 (97%) and 1055 (94.9%) patients, respectively.

Findings to date: Information gathered at recruitment included medical history, lifestyle and environmental exposures. Biological samples were obtained, and 80% patients were genotyped using a commercial exome array. The follow-up was performed in three ways: reviewing medical records; phoning the patients for carrying out a quality-of-life interview, and consulting the Spanish National Death Index. Ninety seven percent recruited patients have been followed-up in 2017 or 2018; patient-years of follow-up were 30914. Most colorectal cancers (52%) were at clinical stage II or less at recruitment; 819 patients died in the follow-up; five-year survival was better for women (74.4%) than for men (70.0%). 71% breast cancers were diagnosed at stages I or II; 206 women with breast cancer died in the follow-up; five-year survival was 90.7%. 49% prostate cancers were diagnosed at stage II and 32% at stage III; 119 patients with prostate cancer died in the follow-up (5-year survival: 93.7%).

Future plans: MCC-Spain has built three prospective cohorts on highly frequent cancers; the information collected at recruitment would allow to investigate clinical, lifestyle, environmental and genetic variables as putative prognosis factors.

KEYWORDS:

Cohort, epidemiology, colorectal cancer, breast cancer, prostate cancer, MCC-Spain

STRENGTHS AND LIMITATIONS OF THIS STUDY:

1. Efficiency when converting cases recruited in the case-control phase of MCC-Spain in prospective cohorts on three of the most frequent cancers in Spain.

2. Control about differential misclassification bias, given that: Firstly, patients were not aware of the hypotheses. Secondly, interviewers were familiar with the case-control study, not with the cohort design as it was decided later; therefore, if interviewers have introduced some bias, it could have been differential between cases and controls, but not among the cases.

3. Multicentre studies could introduce heterogeneity in both the information gathered and the way patients are treated.

4. So me participating patients have been lost; we have tried to minimise it by searching information in three ways: medical records, phone calls and IND; however, we cannot rule out that some patients without follow-up could have died.

INTRODUCTION

The MCC-Spain began as a case-control study in 2008, started by the Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), on both genetic and environmental exposures associated with colorectal, female breast, prostate and gastric cancers and chronic lymphocytic leukaemia. Its design has been published elsewhere[1]; it recruited 10,183 incident cases and controls between 2008 and 2013 in 12 Spanish provinces.

From 2016, the MCC-Spain has turned towards the identification of factors associated with cancer prognosis; in this regard, inception cohorts on colorectal, breast and prostate cancers has been assembled using the incident cases originally recruited between 2008 and 2013, and their prospective follow-up has been carried out in 2017-2018.

Tumour size, node infiltration, metastasis, histology, clinical stage, cancer subtype and first-line treatment are main features of survival studies[2–5] In this sense, the specific aims of this study are to frame a prospective cohort with colorectal, breast and prostate cancer cases; to study the treatment used for each cancer, its effects and the factors that probably have influence over it; and to create models with genetic, epidemiological and clinical-pathological data to predict the survival, treatment response, and toxicity.

In this article, we report the design of the follow-up study, the main description of all three cohorts and the preliminary results on survival.

COHORT DESCRIPTION AND METHODS

MCC-Spain is a population-based multicase-control study. Recruitment began in September 2008 and finished in December 2013. It was carried out in 12 Spanish provinces (Asturias, Barcelona, Cantabria, Girona, Granada, Gipuzkoa, Huelva, León, Madrid, Murcia, Navarra, and Valencia). The following subsections “Patient recruitment”, “Information at recruitment and biological samples”, “Genotyping” and “Initial clinical information” refer to that initial case-control phase; this is hereby summarized as it is the base for the present follow-up cohort study; more detailed information can be found elsewhere[1] The subsections “Cohort inception and follow-up” and “Statistical analysis” refer to the present cohort phase.

Patient recruitment and Public Involvement Statement

The MCC-Spain study called up 10183 subjects; they were between 20 and 85 years old, had resided in the catchment area for at least 6 months before the recruitment and were able to answer the epidemiological questionnaire. For the recruitment, study personnel contacted newly diagnosed cancer cases in the 21 collaborating hospitals; the types of cancer recruited in each hospital was locally decided according to the hospital characteristics. Cases were identified as soon as possible after the diagnosis; only histologically confirmed incident cases were included. From here on, we only refer to the recruited cases of colorectal (2140 cases), breast (1738 cases) and prostate (1112 cases) cancers; their distribution by province and hospital appears in Table 1.

All patients were recruited long before patient and public involvement came into consideration; therefore, patients were not formally involved. They are being informed on the project's main results via flyers.

Table 1. Provinces and hospital of recruitment

Province	Hospital	Colorectal cancer	Breast cancer	Prostate cancer
Asturias	Hospital de Cabueñes	77 (3.60%)	70 (4.03%)	16 (1.44%)
Barcelona	Hospital Clinic	69 (3.22%)	47 (2.70%)	53 (4.77%)
Barcelona	Hospital de Bellvitge – ICO	375 (17.52%)	109 (6.27%)	-
Barcelona	Hospital del Mar	222 (10.37%)	136 (7.83%)	152 (13.67%)
Barcelona	Hospital Germans Trias i Pujol	30 (1.40%)	-	199 (17.90%)
Cantabria	Hospital Universitario Marqués de Valdecilla	151 (7.06%)	141 (8.11%)	175 (15.74%)
Gipuzkoa	Hospital Donostia	119 (5.56%)	126 (7.25%)	-
Gipuzkoa	Instituto Oncológico		100 (5.75%)	-
Girona	Hospital Dr. Josep Trueta	-	21 (1.21%)	-
Girona	Hospital Santa Caterina	-	26 (1.50%)	-
Granada	Hospital San Cecilio	164 (7.66%)	-	64 (5.76%)
Huelva	Hospital Infanta Elena	16 (0.75%)	24 (1.38%)	16 (1.44%)
Huelva	Hospital Juan Ramón Jiménez	55 (2.57%)	84 (4.83%)	36 (3.24%)
León	Hospital de León	390 (18.22%)	226 (13.00%)	-
Madrid	Hospital La Paz	110 (5.14%)	164 (9.44%)	155 (13.94%)
Madrid	Hospital Ramón y Cajal	122 (5.70%)	177 (10.18%)	160 (14.39%)
Murcia	Hospital Morales Messeguer	34 (1.59%)	-	-
Navarra	Complejo Hospitalario de Navarra A (Hospital de Navarra)	76 (3.55%)	112 (6.44%)	-
Navarra	Complejo Hospitalario de Navarra B (Virgen del Camino)	49 (2.29%)	114 (6.56%)	-
Valencia	Hospital Dr. Peset	25 (1.17%)	4 (0.23%)	-
Valencia	Hospital La Fe	56 (2.62%)	57 (3.28%)	86 (7.73%)

Information at recruitment and biological samples

Information about sociodemographic, personal and familial medical history, use of drugs, reproductive history, physical activity, environmental and occupational exposures was gathered using a standardized questionnaire administered by trained personnel in a face-to-face interview. Diet information was obtained using a validated semi-quantitative frequency-food questionnaire filled by the participants. Biological samples were obtained, including peripheral blood (from 96% participants), toenail, hair (from 79% and 84% participants, respectively), urine or tumour biopsies. Regarding peripheral blood, 27ml were aliquoted

in whole blood, plasma, serum and cellular fraction for DNA extraction and stored at -80ºC. Saliva was collected from people unable to donate a blood sample.

Genotyping

From 80% participants, a genotype of exome was made using the Illumina® Infinium HumanExome. In addition to the about 250,000 exome variants included in the original beadcheap, 6000 SNPs previously found in GWAS or localized in metabolic pathways of interest were added upon MCC-Spain researchers' request. MCC-Spain has recently obtained funding for carrying out a GWAS with all the participants and to launch an analysis on circulant miRNA in breast cancer patients.

Initial clinical information

Trained personnel reviewed the medical records in order to collect information on pathology characteristics, tumour extension, clinical data, first-line treatment and recurrence. For colorectal cancer cases, we documented the first biopsy, the tumour location, surgical piece dimensions, histological type according to the ICD-O-3 version, TNM status, carcinoembryonic antigen levels and first-line treatment (surgery extension -if done-; neoadjuvant, adjuvant or palliative chemotherapy or radiotherapy). For breast cancers, we obtained information on tumour location, differentiation's degree, immuno-histochemical characteristics (hormonal receptors, Erb-B2), TNM status and first-line treatment (mastectomy / conservative surgery; neoadjuvant, adjuvant or palliative hormonotherapy, chemotherapy or radiotherapy; target-directed therapy such as transtuzumab). For prostate cancer cases, we gathered information on tumour location, Gleason score, D'Amico classification, TNM status, PSA levels and first-line treatment (none, surgery, hormonotherapy, chemotherapy or radiotherapy; including, when appropriate, the therapy purpose -neoadjuvant, adjuvant or palliative.) TNM status for all three tumours was classified according to the TNM-6th edition.

Cohort inception and follow-up

Colorectal, breast and prostate cancer cases recruited in the previous phase were used to incept three cancer-specific cohorts. Follow-up was carried-out between 2017 and 2018 by reviewing medical records. For colorectal cancer patients, we collected data on TNM status at recruitment, first-line treatment, surgical margins, patient status after first-line treatment (free of disease, partial response, progression, relapse or stable disease), appearance of second primary tumour, and current patient's vital status. For breast cancer patients, we gathered information on histological grade at diagnosis, Nottingham index, complete clinical/pathological remission, grade of response to treatment (according to the Miller and Payne system or similar classifications), relapse, second primary tumour, and current patient's vital status. For prostate cancer patients, the information assembled included PSA concentration, Gleason grade and biopsy characteristics at diagnosis; pathological characteristics of the surgical specimen, first-line treatment, clinical response to first-line treatment (stable disease / progression or relapse / unknown), chemical relapses, relapse clinical characteristics (local / metastatic and its location), second primary tumour, and current patient's vital status. Some of these data were obtained in order to double check the clinical information collected at recruitment.

The National Death Index (Índice Nacional de Defunciones -IND-) was consulted to realize the vital status of patients whose last contact with the hospital had occurred 3 or more months before our revision of his/her medical record. The IND is a nation-wide data-base supported by the Spanish Ministry of Health; it is intended to allow the researchers to establish the vital status of patients under study[6]

Patients alive at the follow-up were contacted by phone and asked to complete specific quality of life questionnaires: SF-12[7] (colorectal, breast and prostate cancers), FACT-Colorectal Symptom Index (FCSI)[8] (colorectal cancer), FACT/NCCN Breast Symptom Index[9] (breast cancer) and -for prostate cancer- the Charlson Comorbidity Index[10], the FACT-P questionnaire[11] and the International Prostate Symptom Score (I-PSS)[12]

Statistical analysis

Data are described using absolute frequencies with percentages and means with standard deviations. To obtain preliminary survival results, patients died by any cause before the end of follow-up were classified as events, and censored otherwise. Time of follow-up was the difference between date of diagnosis and date of death or date of last contact with the hospital or the researchers. Survival probabilities were obtained using unadjusted Kaplan-Meier estimators.

Ethics

The protocol of MCC-Spain was approved by the Ethics committees of the participating institutions [1] At recruitment, all participants were informed about the study objectives and signed an informed consent, which also included the authorization for following-up the patient via medical records or phone calls; only participants agreeing in being followed-up were included in the inception cohorts. Confidentiality of data is secured by removing personal identifiers in the datasets. The database was registered in the Spanish Agency for Data Protection, number 2102672171. Permission to use the study database (Individual-level deidentified patient data) will be granted to researchers outside the study group, after revision and approval of each request by the Steering Committee. Any kind of collaboration are encouraged.

FINDINGS TO DATE

The MCC-Spain has provided results on the effects of different risk factors as night shift work or chronotype [13,14]; use of antihypertensive medication [15], adherence to the Western dietary patterns [16]; physical activity [17]; use of environmental and genetic factors to elaborate a model to stratify the risk of colorectal cancer [18]; or to evaluate the adherence to nutrition-based guidelines and its association with the prevention of cancer [19], among others.

Adding to the aforementioned, initial results of the follow-up are showed in this work. Table 2 displays the main characteristics of the patients; Table 3 details specific information of each tumour; Table 4 describes first-line treatment.

Table 2. Main characteristics of the followed patients

Variable	Category	Colorectal cancer (n = 2097)	Breast cancer (n = 1685)	Prostate cancer (n =1055)
Age (mean±sd)		66.98 (±10.85)	56.5 (±12.6)	65.86 (±7.38)
Gender	**Women**	763 (36.39%)	1685 (100%)	-
	Men	1334 (63.61%)	-	1055 (100%)
Postmenopausal	**Yes**	-	1095 (65.0%)	-
	No	-	589 (35.0%)	-
	Missing	-	1 (0.1%)	-
Histology (specific types in each tumour)		Adenocarcinoma 1882 (89.75%)	Ductal 1276 (75.7%)	Adenocarcinoma (acinar) 1053 (99.91%)
		Mucinous adenocarcinoma 125 (5.96%)	Lobular 110 (6.5%)	Others 2 (0.09%)
		Signet ring cells adenocarcinoma 12 (0.57%)	Paget disease 19 (1.1%)	-
		Others 4 (0.19%)	Others 280 (16.6%)	-
		Unknow 74 (3.53%)	-	-
Tumour size	**T0**	98 (4.67%)	23 (1.4%)	-
	T1	125 (5.96%)	861 (51.1%)	227 (21.52%)
	T2	283 (13.49%)	424 (25.2%)	521 (49.38%)
	T3	1172 (55.89%)	73 (4.3%)	98 (9.29%)
	T4	319 (15.21%)	39 (2.3%)	8 (0.76%)
	Tis	-	109 (6.5%)	-
	Missing	100 (4.77%)	156 (9.3%)	196 (18.58%)
	Not evaluable	-	-	5 (0.47%)
Node infiltration	**N0**	1193 (56.89%)	877 (52.0%)	271 (25.69%)
	N1	515 (24.56%)	441 (26.2%)	9 (0.85%)
	N2	286 (13.64%)	186 (11.0%)	-
	N3	-	5 (0.3%)	-
	Missing	103 (4.91%)	176 (10.4%)	224 (21.23%)
	Not evaluable	-	-	551 (52.23%)
Metastasis	**No**	1721 (82.07%)	1376 (81.7%)	532 (50.43%)
	Yes	330 (15.74%)	41 (2.4%)	17 (1.61%)

	Missing	46 (2.19%)	268 (15.9%)	215 (20.38%)
	Not evaluable	-	-	291 (27.58%)
Clinical stage	**0**	77 (3.67%)	-	-
	I	338 (16.12%)	702 (41.7%)	367 (34.79%)
	II	673 (32.09%)	479 (28.4%)	496 (47.01%)
	III	569 (27.13%)	179 (10.6%)	132 (12.51%)
	IV	330 (15.74%)	22 (1.3%)	17 (1.61%)
	Missing	110 (5.25%)	303 (18.0%)	43 (4.08%)

Table 3. Specific information for each cancer

Specific information for colorectal cancer			Specific information for breast cancer			Specific information for prostate cancer		
			Oestrogen receptor	**Positive**	1398 (83.0%)		**1 (Gleason score = 6)**	449 (42.56%)
				Negative	244 (14.5%)			
				Missing	43 (2.6%)		**2 (Gleason score = 3+4)**	299 (28.34%)
Location	**Right colon**	566 (26.99%)	**Progesterone receptor**	**Positive**	1237 (73.4%)			
				Negative	401 (23.8%)		**3 (Gleason score =4+3)**	120 (11.37%)
	Left colon	719 (34.29%)		**Missing**	47 (2.8%)	**Gleason grade**		
	Rectum-sigma	791 (37.72%)	**Her2**	**Positive**	294 (17.4%)		**4 (Gleason score = 8)**	83 (7.87%)
				Negative	1250 (74.2%)			
	Unknown	21 (1%)		**Missing**	141 (8.4%)		**5 (Gleason score 9 or 10)**	65 (6.16%)
			Intrinsic subtype	**Luminal A**	997 (59.2%)			
				Luminal B	331 (19.6%)			
				Her2	81 (4.8%)		**Missing**	39 (3.70%)
				Basal-like	130 (7.7%)			
Differentiation's degree	**I**	520 (24.8%)		**Luminal ONI***	91 (5.4%)	**PSA** (ng/ml)**		11.51 (±16.28)
	II	1100 (52.46%)		**Non-luminal ONI***	13 (0.8%)			
	III	247 (11.78%)		**Missing**	42 (2.5%)	**D'Amico**	**Low risk**	325 (30.81%)
	Not evaluable	230 (10.97%)	**Grade**	**I**	329 (19.5%)		**Intermediate risk**	425 (40.28%)
				II	520 (30.9%)			
				III	355 (21.1%)		**High risk**	284 (26.92%)
				Missing	481 (28.5%)		**Missing**	21 (1.99%)

*ONI: Otherwise Not Identified. **PSA: Prostate Specific Antigen

Table 4. First-line treatment

Treatment	Category	Colorectal cancer	Breast cancer	Prostate cancer
None (active surveillance)		-	-	38 (3.6%)
Surgery		Total: 1999 (95.3%)	Conservative: 1231(73.1%)	Prostatectomy: 639 (61.4%)
		Resection: 1800 (85.8%)		
		Palliative: 127 (6.1%)	Mastectomy: 454 (26.9%)	
		No resection: 61 (2.9%)		
		Others: 11 (0.5%)		
Chemotherapy	**Neoadjuvant**	427 (20.4%)	200 (11.9%)	1 (0.1%)
	Adjuvant	1024 (48.8%)	664 (39.4%)	1 (0.1%)
	Palliative	67 (3.2%)	25 (1.5%)	7 (0.7%)
Radiotherapy	**Neoadjuvant**	401 (19.1%)	5 (0.3%)	227 (21.5%)
	Adjuvant	82 (3.9%)	1132 (67.2%)	36 (3.4%)
	Palliative	5(0.2%)	21 (1.2%)	2 (0.2%)
Endocrine therapy	**Yes**	-	1023 (60.7%)	Adjuvant to surgery: 19 (1.8%)
				Adjuvant to radiotherapy: 99 (9.4%)
				Neoadjuvant: 102 (9.7%)
				Palliative: 69 (6.5%)
	No	-	662 (39.3%)	689 (65.3%)
Others (specify for each tumour)	**Endoscopy**	Complete resection: 107 (5.1%)	-	-
		Non-complete resection: 62 (3.0%)		
	Her2-targeted therapy	-	152(9.0%)	-
	Cryotherapy	-	-	21 (2.0%)
	Transurethral resection	-	-	4 (0.4%)

Colorectal Cancer

Out of 2140 patients with colorectal cancer, 2097 (98%) have been followed. They were 67±10.9 years-old on average at recruitment; 1334 (63.4%) were men. The first case was recruited on 18th of March 2007 and the follow-up was closed on 23rd of August 2018, accounting for 12813.8 person-years of follow-up. 819 (39.1%) cases died in this period; linearized mortality rate was 6.4 per 100 patient-years (95% CI: 6.0 – 6.8).

Most cases (1882, 90%) were adenocarcinoma; the most frequent location was rectum-sigma (37.7%) and the less frequent right colon (27%). 52% patients were at clinical stage II or less; in 110 patients (5.3%) we could not establish the clinical stage. 52.5% cancers were moderately differentiated (grade II) and 24.8% well differentiated (grade I).

Surgery was carried out in 1999 colorectal cancer patients; it was for palliative purposes in 127 patients (6.1%). 169 patients were treated via endoscopy, reaching complete resection in 107 of them. 1518 (72.4%) patients received chemotherapy; most of them (1451) for adjuvant or neoadjuvant purposes; 488 (23.2%) received radiotherapy (401 neoadjuvant, 82 adjuvant and only 5 palliative).

Five-year survival probability estimated via Kaplan-Meier was 71.6% (95% CI: 69.6 – 73.5) (Figure 1a). Survival was higher in women (74.4%, 95% CI: 71.0 – 77.2) than in men (70.0%, 95% CI: 67.5 – 72.4) ($p<0.001$) (Figure 1b). Five-year survival probability was 85.2% (81.0 – 88.6) in patients diagnosed in stage I, 84.0% (81.0 – 86.6) in stage II, 73.4% (69.6 – 76.9) in stage III and 27.6% (22.9 – 32.5) in stage IV (Figure 2a).

Breast Cancer

The maximum span for breast cancer follow-up was nine and a half years (from 13th July 2007 to 22nd March 2017). Follow-up was obtained for 1685 out of 1738 breast cancer patients (97%), adding 10931 person-years; 206 patients died in the follow-up; the linearized mortality rate was 1.9 per 100 patient-years (95% CI: 1.6 – 2.2).

Women with breast cancer were 56.5±12.6 year-old on average at recruitment; 65% were postmenopausal. The most usual type of tumour was ductal (75.7%), followed by lobular (6.5%). Most breast cancers were diagnosed at early stages (71% at stages I or II) and only 41 (2.4%) had metastasized at the time of diagnosis. 83% cancers were oestrogen receptor positive, 73.4% progesterone receptor positive and 17.4% Her2 positive. Regarding intrinsic subtypes, 997 (59.2%) could be classified as luminal A, 331 (19.6%) as luminal B, 81 (4.8%) as Her2 and 130 (7.7%) as basal-like. According to grade of differentiation, moderately differentiated accounted for 30.9% breast cancers; well differentiated and bad differentiated accounted for about 20% cancers each. Grade could not be obtained from medical records in 481 patients (28.5%).

Conservative surgery was performed in 1231 (73.1%) patients and mastectomy in the remaining 454 (26.9%). Adjuvant or neoadjuvant chemotherapy was administered to 50.3% patients, while radiotherapy was used in 1158 women (68.7%), endocrine therapy was used in 1023 women (60.7%) and Her2-targeted therapy in 152 patients (9.0%). Kaplan-Meier 5-year survival with breast cancer was 90.7% (95% CI: 89.2

– 92.0) (Figure 1c). Women diagnosed in stage I had 97% (95.5 – 98.1) 5-year survival probability, 91.9% (89.1 – 94.1) in stage II, 84.1% (77.8 – 88.7) in stage III and 38.5% (18.6 – 58.2) in stage IV (Figure 2b).

Prostate Cancer

A total of 1112 men with prostate cancer were recruited and 1055 (94.9%) have been followed-up; the first patient was included on 26th January, 2008 and the end of follow-up was on 13th July, 2018, adding 7169.6 person-year of follow-up. Patients were 65.9 years-old on average at recruitment. 119 patients died in the follow-up, making the linearized mortality rate 1.7 per 100 patient-years (95% CI: 1.4 – 2.0).

Almost all prostate cancers (99.9%) were adenocarcinoma; 496 (47%) were diagnosed at stage II and 132 (12.5%) at stage III. The level of PSA gives an average of 11.5±16.3 ng/ml. Considering the Gleason score, 42.6% prostate cancers were well differentiated (Gleason grade = 1, i.e. Gleason score = 6); 28.3% were at Gleason grade 2 (Gleason score = 3+4), and only 14.0% were bad differentiated (Gleason grade 4 or 5; Gleason score ≥8); Gleason grade could not be established in 17.4% patients. D'Amico classification system results in 31.4% patients with low-risk cancer, 41.1% intermediate and 27.4% high-risk cancer.

Thirty-eight prostate cancer patients were not treated medically at the beginning, being followed by active surveillance; prostatectomy was performed in 61.4% cases; radiotherapy in 265 patients (25.1%) and endocrine therapy in 289 patients (27.4%). A small number of patients were treated via trans-urethral resection, cryotherapy or chemotherapy. Five-year survival probability by Kaplan-Meier was 93.7% (95% CI: 92.0 – 95.1) (Figure 1d). Survival probability 5 years after being diagnosed was 94.5% (88.1 – 97.5) for patients in stage I, 95.6% (93.3 – 97.2) in stage II, 92.4% (88.5 – 95.0) in stage III and 70.5 (42.8 – 88.6) in stage IV (Figure 2c).

DISCUSSION

In this article, we have described how three cohorts on colorectal, breast and prostate cancers have been assembled from patients originally recruited for a case-control study, which makes 97% patients followed-up and accounts for more than 30,000 person-years. This is a main achievement of a network settled within the CIBERESP in 12 Spanish provinces. The study is population based and included only incident cancers; the amount of detailed information recorded as well as the availability of biological samples at recruitment will allow the identification of genetics, environmental, lifestyle and clinical prognosis factors in three frequent cancers in Spain. In this regard, a remarkable feature of the study is the feasibility of studying cancer risk factors as putative prognosis factors; for example, risk factors already analysed in the case-control phase have been diet, circadian cycle disruption, some drugs, endocrine disruptors, artificial light or proximity to green spaces; information regarding these risk factors was recorded at recruitment and is available for a prognosis factor analysis in the follow-up.

Personalized medicine will require the integration of huge amounts of information from different -omics; in this regard, MCC-Spain already participates in international consortiums as Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO; https://www.fredhutch.org/en/labs/phs/projects/cancer-prevention/projects/gecco.html), Breast Cancer Association Consortium (BCAC;

http://bcac.ccge.medschl.cam.ac.uk/) and Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL; http://practical.icr.ac.uk/blog/), where MCC-Spain would contribute to study interactions among the putative prognosis factors in vast population samples.

One of the main strengths of this study is its efficiency: When converting cases recruited in the case-control phase of MCC-Spain in prospective cohorts on three of the most frequent cancers in Spain, we take advantage of the recruitment itself and both the information and the samples collected at recruitment, making it possible to build the three cohorts at the only marginal cost of the follow-up. Had we had to assemble new cohorts on these cancers, however, would have require doubling the resources we have already expended in the case-control phase.

Obtaining information on personal history, occupational exposures, diet, physical exercise or other lifestyle components is somewhat subjective as both patients and interviewers could be prone to be influenced by their feelings or beliefs about the hypotheses under study, eventually leading to differential misclassification bias. This could hardly have occurred in this study: Firstly, patients were not aware of the hypotheses. Secondly, interviewers were familiar with the case-control study, not with the cohort design as it was decided later; therefore, if interviewers have introduced some bias, it could have been differential between cases and controls, but not among the cases, which would make more robust the results obtained in this cohort study.

This study has also some weaknesses: Firstly, multicentre studies are double edged; they are needed in order to include many patients, but they could introduce heterogeneity in both the information gathered and the way patients are treated. In this regard, the analysis of prognosis factors should be adjusted for the hospital of recruitment. Secondly, some participating patients have been lost; we have tried to minimise it by searching information in three ways: medical records, phone calls and IND; however, we cannot rule out that some patients without follow-up could have died. It is noteworthy that -due to the small number of patients without follow-up- the maximum bias it could introduce in our survival estimates is 2% for colorectal cancer, 3% for breast cancer and 5% for prostate cancer.

Summarizing, the MCC-Spain study has assembled three cohorts with about 4,700 cancer patients accounting for 30,000 patient-years of follow-up, with only 3% patient withdrawals. The information gathered at recruitment will allow to prospectively investigate clinical, lifestyle, environmental and genetic variables as prognosis factors in colorectal, breast and prostate cancers in Spain.

DATA STATEMENT

Permission to use the study database (Individual-level deidentified patient data) will be granted to researchers outside the study group, after revision and approval of each request by the Steering Committee. Any kind of collaboration are encouraged.

ACKNOWLEDGEMENTS

We thank all the subjects who participated in the study and all MCC-Spain collaborators.

INFORMED CONSENT:

Informed consent was obtained from all individual participants included in the study.

AUTHOR STATEMENT:

Jessica Alonso-Molero, Antonio J Molina, Jose J Jiménez-Moleón have contributed to the conception and design of the study, considering the same contribution. All authors have acquired data and have been involved in drafting the manuscript. All of them read and approved the final manuscript.

FUNDING:

The study was partially funded by the "Accion Transversal del Cancer", approved on the Spanish Ministry Council on the 11th October 2007, by the Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PS09/00773-Cantabria, PS09/01286-León, PS09/01903-Valencia, PS09/02078-Huelva, PS09/01662-Granada, PI11/01403, PI11/01889-FEDER, PI11/00226, PI11/01810, PI11/02213, PI12/00488, PI12/00265, PI12/01270, PI12/00715, PI12/00150, PI14/01219, PI14/0613, PI15/00069, PI15/00914, PI15/01032, PI17CIII/00034, PI18/00181), by the Fundación Marqués de Valdecilla (API 10/09), by the ICGC International Cancer Genome Consortium CLL (The ICGC CLL-Genome Project is funded by Spanish Ministerio de Economía y Competitividad (MINECO) through the Instituto de Salud Carlos III (ISCIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036)), by the Junta de Castilla y León (LE22A10-2), by the Consejería de Salud of the Junta de Andalucía (PI-0571-2009, PI-0306-2011, salud201200057018tra), by the Conselleria de Sanitat of the Generalitat Valenciana (AP_061/10), by the Recercaixa (2010ACUP 00310), by the Regional Government of the Basque Country, by the Consejería de Sanidad de la Región de Murcia, by the European Commission grants FOOD-CT-2006-036224-HIWATE, by the Spanish Association Against Cancer (AECC) Scientific Foundation, by the Catalan Government- Agency for Management of University and Research Grants (AGAUR) grants 2017SGR723 and 2014SGR850, by the Fundación Caja de Ahorros de Asturias and by the University of Oviedo. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya.

SAMPLES:

Biological samples were stored at the biobanks supported by Instituto de Salud Carlos III- FEDER: Parc de Salut MAR Biobank (MARBiobanc) (RD09/0076/00036), "Biobanco La Fe" (RD 09 0076/00021) and FISABIO Biobank (RD09 0076/00058). Also at the Public Health Laboratory from Gipuzkoa, the Basque

Biobank, the ICOBIOBANC (sponsored by the Catalan Institute of Oncology), the IUOPA Biobank from the University of Oviedo and the ISCIII Biobank.

GENOTYPING:

SNP genotyping services were provided by the Spanish "Centro Nacional de Genotipado" (CEGEN-ISCIII)".

CONFLICT OF INTEREST:

The authors declare that they have no conflict of interest

REFERENCES

1 Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, *et al.* Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *GacSanit* 2015;**29**:308–15. doi:10.1016/j.gaceta.2014.12.003

2 Shukla N, Hagenbuchner M, Win KT, *et al.* Breast cancer data analysis for survivability studies and prediction. *Comput Methods Programs Biomed* 2018;**155**:199–208. doi:10.1016/j.cmpb.2017.12.011

3 Mirza AN, Mirza NQ, Vlastos G, *et al.* Prognostic Factors in Node-Negative Breast Cancer. *Ann Surg* 2002;**235**:10–26. doi:10.1097/00000658-200201000-00003

4 Zhang Z yu, Luo Q feng, Yin X wei, *et al.* Nomograms to predict survival after colorectal cancer resection without preoperative therapy. *BMC Cancer* 2016;**16**:1–21. doi:10.1186/s12885-016-2684-4

5 Merriel SWD, May MT, Martin RM. Predicting prostate cancer progression: Protocol for a retrospective cohort study to identify prognostic factors for prostate cancer outcomes using routine primary care data. *BMJ Open* 2018;**8**:1–5. doi:10.1109/TVLSI.2018.2801302

6 Navarro C. El Índice Nacional de Defunciones: Un avance en la accesibilidad de los datos de mortalidad largamente esperado. *Gac Sanit* 2006;**20**:421–3. doi:10.1157/13096513

7 Vilagut G, Valderas JM, Ferrer M, *et al.* [Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components]. *Med Clin (Barc)* 2008;**130**:726–35.http://www.ncbi.nlm.nih.gov/pubmed/18570798 (accessed 15 Jan 2019).

8 Colwell HH, Mathias SD, Solutions HO, *et al.* Psychometric Evaluation of the FACT Colorectal Cancer Symptom Index (FCSI-9): Reliability, Validity, Responsiveness, and Clinical Meaningfulness HILARY. *Oncologist* 2010;**15**:308–16. doi:10.1634/theoncologist.2009-0034

9 Garcia SF, Rosenbloom SK, Beaumont JL, *et al.* Priority symptoms in advanced breast cancer: Development and initial validation of the national comprehensive cancer Network-Functional Assessment of Cancer Therapy-Breast Cancer Symptom Index (NFBSI-16). *Value Heal* 2012;**15**:183–90. doi:10.1016/j.jval.2011.08.1739

10 Charlson M, Pompei P, Ales K, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.

11 Esper P, Mo F, Chodak G, *et al.* Measuring quality of life in men with prostate cancer using the functionale assessment of cancer therapy-prostate instrument. *Adult Urol* 1997;**30**:920–8.

12 Barry MJ, Fowler FJ, O'Leary MP, *et al.* The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol* 1992;**148**:1549–57.

13 Papantoniou K, Castaño-Vinyals G, Espinosa A, *et al.* Night shift work, chronotype and prostate cancer risk in the MCC-Spain case-control study. *Int J Cancer* 2015;**137**:1147–57. doi:10.1002/ijc.29400

14 Costas L, Benavente Y, Olmedo-Requena R, *et al.* Night shift work and chronic lymphocytic leukemia in the MCC-Spain case–control study. *Int J Cancer* 2016;**139**:1994–2000. doi:10.1002/ijc.30272

15 Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, *et al.* The use of antihypertensive medication and the risk of breast cancer in a case-control study in a Spanish population: The MCC-Spain study. *PLoS One* 2016;**11**:1–14. doi:10.1371/journal.pone.0159672

16 Castelló A, Boldo E, Pérez-Gómez B, *et al.* Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. *Maturitas* 2017;**103**:8–15. doi:10.1016/j.maturitas.2017.06.020

17 Huerta JM, Chirlaque MD, Molina AJ, *et al.* Physical activity domains and risk of gastric adenocarcinoma in the MCC-Spain case-control study. *PLoS One* 2017;**12**:1–16.

doi:10.1371/journal.pone.0179731

18 Ibáñez-Sanz G, Díez-Villanueva A, Alonso MH, *et al.* Risk Model for Colorectal Cancer in Spanish Population Using Environmental and Genetic Factors: Results from the MCC-Spain study. *Sci Rep* 2017;**7**:43263. doi:10.1038/srep43263

19 Romaguera D, Gracia-Lavedan E, Molinuevo A, *et al.* Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case–control study. *Int J Cancer* 2017;**141**:83–93. doi:10.1002/ijc.30722

FIGURE LEGENDS

Fig 1 Kaplan-Meier survival estimates for colorectal cancer (1a), colorectal cancer by sex (1b), breast cancer (1c) and prostate cancer (1d)

Fig 2 Kaplan-Meier estimates by stage at diagnosis for colorectal cancer (2a), breast cancer (2b) and prostate cancer (2c)

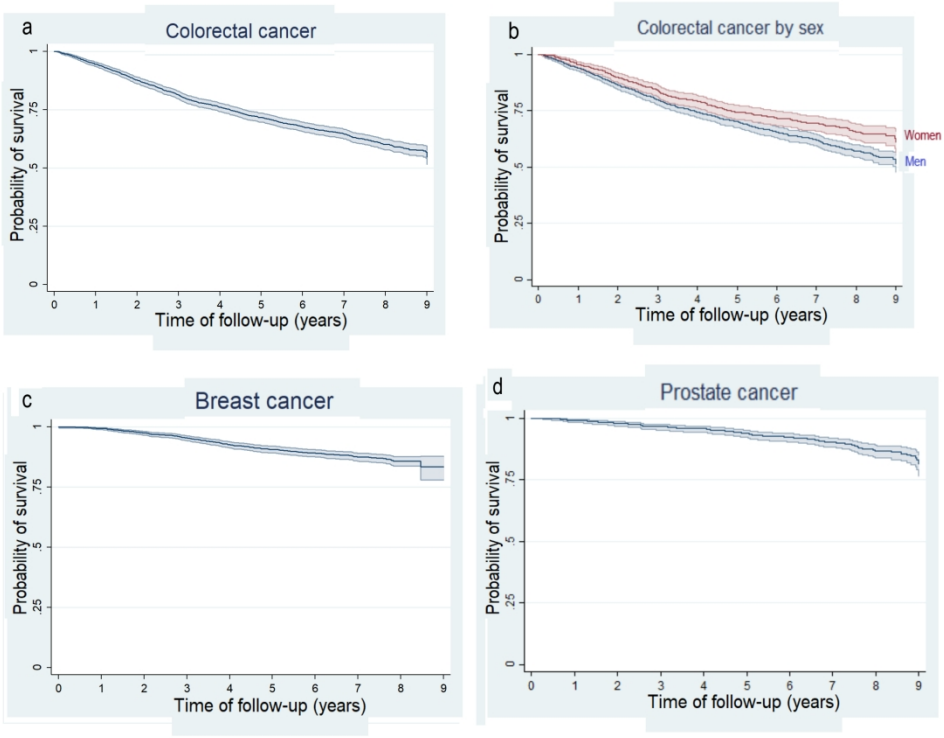

Figure 1. Kaplan-Meier survival estimates for colorectal cancer (1a), colorectal cancer by sex (1b), breast cancer (1c) and prostate cancer (1d)

727x602mm (96 x 96 DPI)

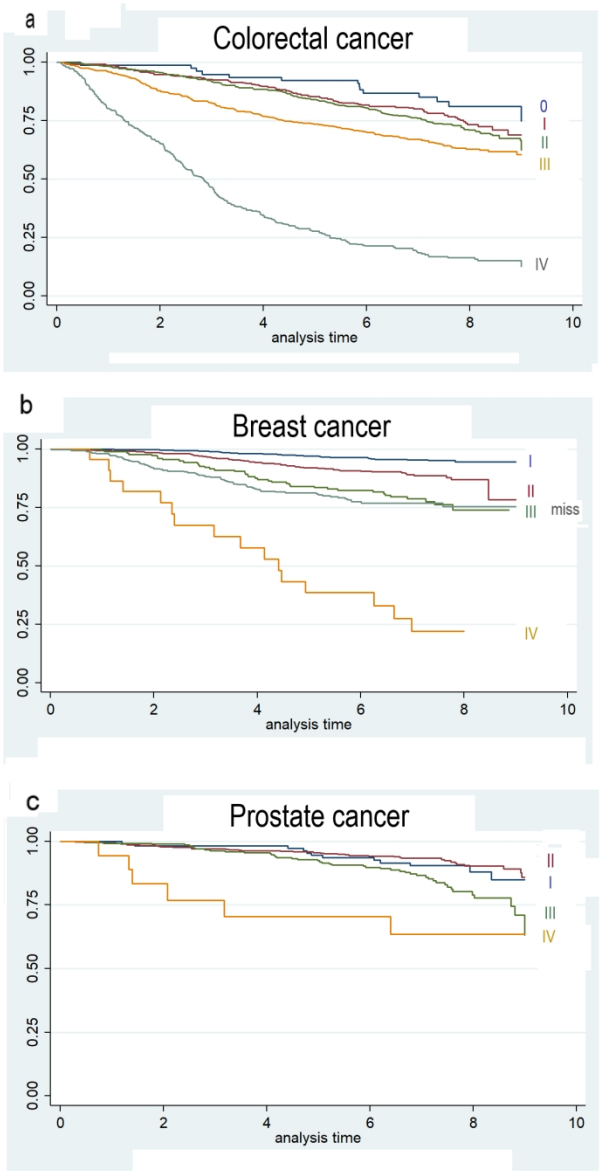

Figure 2. Kaplan-Meier estimates by stage at diagnosis for colorectal cancer (2a), breast cancer (2b) and prostate cancer (2c)

BMJ Open

Cohort profile: The MCC-Spain follow-up on colorectal, breast and prostate cancers. Study design and initial results

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-031904.R1
Article Type:	Cohort profile
Date Submitted by the Author:	28-Aug-2019
Complete List of Authors:	Alonso, Jessica ; University of Cantabria – IDIVAL, Santander, Spain Molina, Antonio J; Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain. Jiménez-Moleón, Jose Juan; University of Granada, Department of Preventive Medicine and Public Health, School of Medicine; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Pérez-Gómez, Beatriz ; National Center for Epidemiology, Carlos III Institute of Health; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Martin, Vicente; 2- Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain. Moreno, Victor; Catalan Institute of Oncology (ICO-IDIBELL), Cancer Prevention and Control Program; University of Barcelona, Clinical Sciences, Faculty of Medicine Amiano, Pilar; Public Health Division of Gipuzkoa, BioDonostia Research Institute; Consortium for Biomedical Research in Epidemiology and Public Health (CIBER Epidemiología y Salud Pública, CIBERESP) Ardanaz, Eva de Sanjose, Silvia ; L'Hospitalet de Llobregat, Barcelona SALCEDO, INMACULADA; Universidad de Granada, Preventive Medicine and Public Health+34 Fernandez-Tardon, Guillermo; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Alguacil, Juan; CIBER de Epidemiología y Salud Pública (CIBERESP), Salas, Dolores; Generalitat Valenciana Conselleria de Sanitat Marcos-Gragera, Rafael; Institut Catala d' Oncologia, Epidemiology Unit and Girona Cancer Registry; Universitat de Girona, Research Group on Statistics, Econometrics and Health (GRECS) Chirlaque, Maria Dolores; Murcia Cancer Registry Aragonés, Nuria; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Castaño-Vinyals, Gemma ; Instituto de Salud Global Barcelona Pollán, Marina; Centro Nacional de Epidemiologia. Instituto de Salud Carlos III, Area de Epidemiologia Ambiental y Cancer; CIBER Epidemiologia y Salud Publica (CIBERESP), Kogevinas, Manolis; Instituto de Salud Global de Barcelona, Llorca, Javier; University of Cantabria,
Primary Subject	Epidemiology

Heading:	
Secondary Subject Heading:	Oncology
Keywords:	Cohort, Epidemiology < TROPICAL MEDICINE, colorectal cancer, breast cancer, prostate cancer, MCC-Spain

Cohort profile: The MCC-Spain follow-up on colorectal, breast and prostate cancers. Study design and initial results

Authors: Jessica Alonso-Molero*[1], Antonio J Molina*[2], Jose J Jiménez-Moleón*[3,4], Beatriz Pérez-Gómez[3,5,6], Vicente Martín[2,3], Víctor Moreno[3,7,8], Pilar Amiano[3,9], Eva Ardanaz[3,10,11], Silvia de Sanjosé[3,7], Inmaculada Salcedo[3,4], Guillermo Fernández-Tardón[3,12], Juan Alguacil[3,13], Dolores Salas[3,14], Rafael Marcos-Gragera[3,15], M Dolores Chirlaque[3,16], Nuria Aragonés[3,17], Gemma Castaño-Vinyals[3,18,19,20], Marina Pollán[3,5,6], Manolis Kogevinas[3,18,19,20], Javier Llorca[1,3].

*Equal contribution

Affiliations:

1- University of Cantabria-IDIVAL, Santander, Spain.
2- Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain.
3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain.
4- Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain.
5- Cancer and Environmental Epidemiology Unit, National Center for Epidemiology, Carlos III Institute of Health, Madrid, Spain.
6- Cancer Epidemiology Research Group, Oncology and Hematology Area, IIS Puerta de Hierro (IDIPHIM), Madrid, Spain.
7- Cancer Prevention and Control Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Spain.
8- Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain.
9- Public Health Division of Gipuzkoa, BioDonostia Research Institute, San Sebastian, Spain.
10- Navarra Public Health Institute, Pamplona, Spain.
11- IdiSNA, Navarra Institute for Health Research, Pamplona, Spain.
12- Oncology Institute, University of Oviedo, Oviedo.
13- Centro de Investigación en Recursos Naturales, Salud y Medio Ambiente (RENSMA), Universidad de Huelva, 21004 Huelva, Spain.
14- Área de Cáncer y Salud Pública, FISABIO-Salud Pública, 46020 Valencia, Spain.
15- Unitat d'Epidemiologia i Registre de Càncer de Girona (UERCG), Pla Director d'Oncologia, Institut Català d'Oncologia, Institut d'Investigaciò Biomèdica de Girona (IdIBGi), Universitat de Girona, Girona, Spain.
16- Department of Epidemiology, Regional Health Authority, IMIB-Arrixaca, Murcia University, Murcia, Spain.
17- Department of Health, Epidemiology Section, Public Health Division, Madrid.
18- ISGlobal, Barcelona, Spain.

19- Universitat Pompeu Fabra (UPF), Barcelona, Spain.

20- IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain.

Correspondence:

Jéssica Alonso-Molero

Facultad de Medicina, Universidad de Cantabria

Avda. Herrera Oria s/n

39011 Santander (Cantabria)

Spain

ORCiD: https://orcid.org/0000-0002-1939-8798

e-mail: alonsomoleroj@gmail.com

Word count: 3216

ABSTRACT

Purpose: Since 2016, the Multicase-control study in Spain (MCC-Spain) has focused towards the identification of factors associated with cancer prognosis; inception cohorts of patients with colorectal, breast and prostate cancers were assembled using the incident cases originally recruited.

Participants: 2140 new cases of colorectal cancer, 1732 of breast cancer and 1112 of prostate cancer were initially recruited in 12 Spanish provinces; all cancers were incident and pathologically confirmed. Follow-up was obtained for 2097 (98%), 1685 (97%) and 1055 (94.9%) patients, respectively.

Findings to date: Information gathered at recruitment included sociodemographic factors, medical history, lifestyle and environmental exposures. Biological samples were obtained, and 80% patients were genotyped using a commercial exome array. The follow-up was performed by: (i) reviewing medical records; (ii) interviewing by phone the patients on quality-of-life and; (iii) verifying vital status and cause of death in the Spanish National Death Index. Ninety-seven percent of recruited patients were successfully followed-up in 2017 or 2018; patient-years of follow-up were 30914. Most colorectal cancers (52%) were at clinical stage II or less at recruitment; 819 patients died in the follow-up and five-year survival was better for women (74.4%) than men (70.0%). 71% breast cancers were diagnosed at stages I or II; 206 women with breast cancer died in the follow-up and five-year survival was 90.7%. 49% prostate cancers were diagnosed at stage II and 32% at stage III; 119 patients with prostate cancer died in the follow-up and five-year survival was 93.7%.

Future plans: MCC-Spain has built three prospective cohorts on highly frequent cancers across Spain, allowing to investigate socioeconomic, clinical, lifestyle, environmental and genetic variables as putative prognosis factors determining survival of patients of the three cancers and of the interrelathionship of these factors.

KEYWORDS:

Cohort, epidemiology, colorectal cancer, breast cancer, prostate cancer, MCC-Spain

STRENGTHS AND LIMITATIONS OF THIS STUDY:

- 4837 incident cases of cancer (2097 colorectal; 1685 breast; 1055 prostate) have been prospectively followed-up accounting for more than 30000 patients-year, and with only 153 patients (3%) lost to follow-up.
- The cohort a wide spectrum of the Spanish population including 23 hospitals across Spain.
- A major strength of this study is the amount of information gathered at diagnosis, including sociodemographic, lifestyle, nutrition, familial and personal medical history, reproductive history, use of drugs, sleep, genotyping, clinical and pathological characteristics of the tumour, first-line treatment, side effects, health-related quality of life and current vital status.
- Biological samples obtained at recruitment (tumour specimen, blood or saliva, toenail, hair and urine) will allow further investigations on metabolomics, epigenetics and exposure to chemicals such as metals.

- The multicentre characteristic of the study allows the evaluation of a wide geographical basis and increases the representativity of the recruited sample, but it also may introduce heterogeneity in the information gathered and in treatment.

INTRODUCTION

Tumour size, node infiltration, metastasis, histology, clinical stage, cancer subtype continue being main prognosis factors in patients with cancer in spite of the evolving first-line treatment [1–5]. Little effort, however, has been paid to examine the impact on survival of patient factors -such as lifestyle, genetics or environmental- together with tumour features and treatment.

Large prospective cohort studies on cancer focus on identifying risk factors [6] while clinical cohorts on cancer survival usually aim to analyse survival relationships with tumour properties, first-line treatment or patient characteristics. For instance, Lagendijk et al analysed data on 129,692 women with breast cancer from the Netherlands Cancer Registry to compare breast conserving therapy and mastectomy in subgroups according age at diagnosis, stage, systemic therapy, comorbidity, oestrogen/progesterone receptors and HER2 status [7]; Cardwell et al linked the National Cancer Data Repository to the United Kingdom Clinical Practice Research Datalink and mortality data from the Office of National Statistics to investigate if statin use after colorectal cancer diagnosis was associated with better prognosis [8]; Petterson et al studied survival after prostate cancer diagnosis in 121,392 Swedish men from the Prostate Cancer data Base Sweden 3.0, where data were available on age, stage, grade, prostate-specific antigen level, model of detection, comorbidity, educational level and primary treatment [9]. It is noteworthy that these cohorts were based on cancer registries were data availability is usually restricted to demographic variables (sometimes including educational level and deprivation), tumour characteristics and few data on comorbidities or healthy habits. A different approach has been the use of the Surveillance, Epidemiology and End Results (SEER) database to retrospectively analyse survivorship with breast cancer [10], colorectal cancer [11] or prostate cancer [12], but although the number of participants could be over 100,000, available data are restricted to those recorded for the general purposes of the SEER program, not specifically for studying survivorship with cancer.

The MCC-Spain includes three prospective cohorts of cancer patients (colorectal, female breast and prostate) with the aim of to investigate long-term survival factors including cancer characteristics and treatment, but also genetics and other omics, lifestyle (physical activity, nutrition, sleep, toxic habits), occupational exposures (including night shift work), environmental factors such as living area conditions and medical history, aiming to build integrative prognosis models. This multidisciplinary study will provide a complete evaluation of the biological, clinical, environmental, lifestyle and socio-economic factors determining survival of patients of the three cancers and of the interrelationship of these factors. Specific objectives for each cohort are: For the colorectal cancer cohort: (1) To study the accomplishment of primary treatment with ESMO and ASCO guidelines and factors associated with it, (2) to study factors associated with survivorship, response to treatment and toxicity due to chemotherapy using genetic, epidemiological and clinical-pathological variables, (3) to validate those models via comparison with Glasgow Prognostic Score predictions. For the breast cancer cohort: (1) To study whether first-line treatment accomplished St Gallen International Expert Consensus recommendations, (2) to study factors associated with survivorship, response to treatment and toxicity due to chemotherapy using genetic, epidemiological and clinical-pathological variables, (3) to validate those models via comparison with the Nottingham Prognostic Index and *Adjuvant!*. For prostate cancer cohort: (1) To analyse the adequacy of initial treatment to

recommendations by the European Association of Urology and the National Institute for Health and Care Excellence, (2) to elaborate models on survivorship, risk of biochemical relapse, quality of life, response to primary treatment, toxicity to chemotherapy/brachitherapy (3) to validate survivorship and risk of biochemical relapse models via comparison with Han and Kattan nomograms. In this article, we report the study design, the main description of all three cohorts and the preliminary results on survival.

COHORT DESCRIPTION AND METHODS

The MCC-Spain began as a case-control study in 2008, started by the Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), on both genetic and environmental exposures associated with colorectal, female breast, prostate and gastric cancers and chronic lymphocytic leukaemia. Its design has been published elsewhere[13]; it recruited 10,183 incident cases and controls between 2008 and 2013 in 12 Spanish provinces (Asturias, Barcelona, Cantabria, Girona, Granada, Gipuzkoa, Huelva, León, Madrid, Murcia, Navarra, and Valencia). From 2016, the MCC-Spain has turned towards the identification of factors associated with cancer prognosis; in this regard, inception cohorts on colorectal, breast and prostate cancers has been assembled using the incident cases originally recruited between 2008 and 2013, and their prospective follow-up has been carried out in 2017-2018. From here on, we only refer to the recruited cases of colorectal (2140 cases), breast (1738 cases) and prostate (1112 cases) cancers; their distribution by province and hospital appears in Supplementary Table 1 and the flow chart appears in Figure 1.

Patient recruitment and Public Involvement Statement

Patients recruited were between 20 and 85 years old, had resided in the catchment area for at least 6 months before the recruitment and were able to answer the epidemiological questionnaire and had incident colorectal, breast or prostate cancer. For the recruitment, study personnel contacted newly diagnosed cancer cases in the 21 collaborating hospitals. Cases were identified as soon as possible after the diagnosis; only histologically confirmed incident cases were included.

Participants are being informed on the project's main results via flyers. There is no other patient's involvement.

Information at recruitment and biological samples

The information obtained and its timing is summarized in Table 1.

Table 1. Information obtained in the MCC-Spain

Phase	Measurements
Phase I: Recruitment 2008-2013	**Contact with newly diagnosed cancer cases.** **Trained personnel perform a structured computerized epidemiological questionnaire in a face-to-face interview to obtain the follow information:** Sociodemographic; Personal and familial medical history; Use of drugs; Reproductive history; Physical activity; Environmental and occupational exposures **A validated semi-quantitative frequency-food questionnaire is self-completed to obtain diet information.** **Biological samples are obtained:** Peripheral blood or saliva; Toenail; Hair; Urine; Tumour biopsies **A genotype of exome is made using the Illumina® Infinium HumanExome.** **Medical Records review by trained personnel to obtain:** Pathology characteristics; Tumour extension; Clinical data; First-line treatment; Recurrence For colorectal cancer cases: First biopsy; Surgical piece dimensions; Histological type; Carcinoembryonic antigen levels For breast cancers cases: Differentiation's degree; Immuno-histochemical characteristics For prostate cancer cases: Gleason score; D'Amico classification; PSA levels
Phase II: Follow-up 2017-2018	**Medical Records review by trained personnel to obtain:** For colorectal cancer cases: TNM status at recruitment; First-line treatment; Surgical margins; Patient status after first-line treatment; Appearance of second primary tumour; Current patient's vital status For breast cancers cases: Histological grade at diagnosis; Nottingham index; Complete clinical/pathological remission; Grade of response to treatment; Relapse; Second primary tumour; Current patient's vital status For prostate cancer cases: PSA concentration; Gleason grade and biopsy characteristics at diagnosis; Pathological characteristics of the surgical specimen; First-line treatment; Clinical response to first-line treatment; Second primary tumour; Current patient's vital status **Consult in the IND to realize the vital status of patients.** **Contact by phone to complete specific quality of life questionnaires.** For colorectal cancer cases: SF-12; FACT-Colorectal Symptom Index For breast cancer cases: SF-12; FACT/NCCN Breast Symptom Index For prostate cancer cases: SF-12; Charlson Comorbidity Index; FACT-P questionnaire; International Prostate Symptom Score (I-PSS).

Information about sociodemographic, personal and familial medical history, use of drugs, reproductive history, physical activity, environmental and occupational exposures was gathered using a standardized questionnaire administered by trained personnel in a face-to-face interview. Diet information was obtained using a validated semi-quantitative frequency-food questionnaire filled by the participants. Biological samples were obtained, including peripheral blood or saliva (from 92% breast cancer cases, 95% colorectal cancer cases and 97% prostate cancer cases), toenail, hair (from 77% and 81% participants, respectively), urine or tumour biopsies. Regarding peripheral blood, 27ml were aliquoted in whole blood, plasma, serum and cellular fraction for DNA extraction and stored at -80ºC. Saliva was collected from people unable to donate a blood sample.

Genotyping

From 80% participants, a genotype of exome was made using the Illumina® Infinium HumanExome. In addition to the about 250,000 exome variants included in the original beadcheap, 6000 SNPs previously found in GWAS or localized in metabolic pathways of interest were added upon MCC-Spain researchers' request. MCC-Spain has recently obtained funding for carrying out a GWAS with all the participants and to launch an analysis on circulant miRNA in breast cancer patients.

Initial clinical information

Trained personnel reviewed the medical records in order to collect information on pathology characteristics, tumour extension, clinical data, first-line treatment and recurrence. For colorectal cancer cases, we documented the first biopsy, the tumour location, surgical piece dimensions, histological type according to the ICD-O-3 version, TNM status, carcinoembryonic antigen levels and first-line treatment (surgery extension -if done-; neoadjuvant, adjuvant or palliative chemotherapy or radiotherapy). For breast cancers, we obtained information on tumour location, differentiation's degree, immuno-histochemical characteristics (hormonal receptors, Erb-B2), TNM status and first-line treatment (mastectomy / conservative surgery; neoadjuvant, adjuvant or palliative hormonotherapy, chemotherapy or radiotherapy; target-directed therapy such as transtuzumab). For prostate cancer cases, we gathered information on tumour location, Gleason score, D'Amico classification, TNM status, PSA levels and first-line treatment (none, surgery, hormonotherapy, chemotherapy or radiotherapy; including, when appropriate, the therapy purpose -neoadjuvant, adjuvant or palliative.) TNM status for all three tumours was classified according to the TNM-6th edition.

Follow-up information

Follow-up was carried-out between 2017 and 2018 by reviewing medical records. For colorectal cancer patients, we collected data on TNM status at recruitment, first-line treatment, surgical margins, patient status after first-line treatment (free of disease, partial response, progression, relapse or stable disease), appearance of second primary tumour, and current patient's vital status. For breast cancer patients, we gathered information on histological grade at diagnosis, Nottingham index, complete clinical/pathological remission, grade of response to treatment (according to the Miller and Payne system or similar classifications), relapse, second primary tumour, and current patient's vital status. For prostate cancer patients, the information assembled included PSA concentration, Gleason grade and biopsy characteristics

at diagnosis; pathological characteristics of the surgical specimen, first-line treatment, clinical response to first-line treatment (stable disease / progression or relapse / unknown), chemical relapses, relapse clinical characteristics (local / metastatic and its location), second primary tumour, and current patient's vital status. Some of these data were obtained in order to double check the clinical information collected at recruitment.

The National Death Index (Índice Nacional de Defunciones -IND-) was consulted to realize the vital status of patients whose last contact with the hospital had occurred 3 or more months before our revision of his/her medical record. The IND is a nation-wide data-base supported by the Spanish Ministry of Health; it is intended to allow the researchers to establish the vital status of patients under study[14]

Patients alive at the follow-up were contacted by phone and asked to complete specific quality of life questionnaires: SF-12[15] (colorectal, breast and prostate cancers), FACT-Colorectal Symptom Index (FCSI)[16] (colorectal cancer), FACT/NCCN Breast Symptom Index [17] (breast cancer) and -for prostate cancer- the Charlson Comorbidity Index [18], the FACT-P questionnaire [19] and the International Prostate Symptom Score (I-PSS)[20].

Statistical analysis

For preliminary results shown in this paper, data are described using absolute frequencies with percentages and means with standard deviations. Patients died by any cause before the end of follow-up were classified as events and censored otherwise. Time of follow-up was the difference between date of diagnosis and date of death or date of last contact with the hospital or the researchers. Survival probabilities were obtained using unadjusted Kaplan-Meier estimators.

Ethics

The protocol of MCC-Spain was approved by the Ethics committees of the participating institutions [13] At recruitment, all participants were informed about the study objectives and signed an informed consent, which also included the authorization for following-up the patient via medical records or phone calls; only participants agreeing in being followed-up were included in the inception cohorts. Confidentiality of data is secured by removing personal identifiers in the datasets. The database was registered in the Spanish Agency for Data Protection, number 2102672171.

FINDINGS TO DATE

The MCC-Spain has provided results on the effects of different risk factors. For instance, night shift work increased the risk of more aggressive prostate cancers [21], although this excess risk almost disappeared 20 years after last exposure [22] o; long-term consumption of calcium channel blockers was associated with higher breast cancer risk in overweight women [23]; adherence to the Western dietary patterns increased breast cancer risk in both pre- and post-menopausal women [24]; first validation in a European population of a risk model for breast cancer developed in American women using both modifiable and non-modifiable risk factors as well as 92 genetic variants [25]; use of environmental and genetic factors to elaborate a model to stratify the risk of colorectal cancer [26]; adherence to the World Cancer Research Fund/American Institute for Cancer Research nutrition-based guidelines was associated with lower risk of colorectal and

breast cancers, but not of prostate cancer [27]. A complete list of published results from MCC-Spain appears in Supplementary Table 2 and Supplementary reference list.

Initial results of the follow-up are showed in this work. Table 2 displays the main characteristics of the patients; Table 3 details specific information of each tumour; Table 4 describes first-line treatment.

Table 2. Main characteristics of the followed patients

Variable	Category	Colorectal cancer (n = 2097)	Breast cancer (n = 1685)	Prostate cancer (n =1055)
Age (mean±sd)		66.98 (±10.85)	56.5 (±12.6)	65.86 (±7.38)
Gender	**Women**	763 (36.39%)	1685 (100%)	-
	Men	1334 (63.61%)	-	1055 (100%)
Postmenopausal	**Yes**	-	1095 (65.0%)	-
	No	-	589 (35.0%)	-
	Missing	-	1 (0.1%)	-
Histology (specific types in each tumour)		Adenocarcinoma 1882 (89.75%)	Ductal 1276 (75.7%)	Adenocarcinoma (acinar) 1053 (99.91%)
		Mucinous adenocarcinoma 125 (5.96%)	Lobular 110 (6.5%)	Others 2 (0.09%)
		Signet ring cells adenocarcinoma 12 (0.57%)	Paget disease 19 (1.1%)	-
		Others 4 (0.19%)	Others 280 (16.6%)	-
		Unknow 74 (3.53%)	-	-
Tumour size	**T0**	98 (4.67%)	23 (1.4%)	-
	T1	125 (5.96%)	861 (51.1%)	227 (21.52%)
	T2	283 (13.49%)	424 (25.2%)	521 (49.38%)
	T3	1172 (55.89%)	73 (4.3%)	98 (9.29%)
	T4	319 (15.21%)	39 (2.3%)	8 (0.76%)
	Tis	-	109 (6.5%)	-
	Missing	100 (4.77%)	156 (9.3%)	196 (18.58%)
	Not evaluable	-	-	5 (0.47%)
Node infiltration	**N0**	1193 (56.89%)	877 (52.0%)	271 (25.69%)
	N1	515 (24.56%)	441 (26.2%)	9 (0.85%)
	N2	286 (13.64%)	186 (11.0%)	-
	N3	-	5 (0.3%)	-
	Missing	103 (4.91%)	176 (10.4%)	224 (21.23%)
	Not evaluable	-	-	551 (52.23%)
Metastasis	**No**	1721 (82.07%)	1376 (81.7%)	532 (50.43%)
	Yes	330 (15.74%)	41 (2.4%)	17 (1.61%)

	Missing	46 (2.19%)	268 (15.9%)	215 (20.38%)
	Not evaluable	-	-	291 (27.58%)
Clinical stage	**0**	77 (3.67%)	-	-
	I	338 (16.12%)	702 (41.7%)	367 (34.79%)
	II	673 (32.09%)	479 (28.4%)	496 (47.01%)
	III	569 (27.13%)	179 (10.6%)	132 (12.51%)
	IV	330 (15.74%)	22 (1.3%)	17 (1.61%)
	Missing	110 (5.25%)	303 (18.0%)	43 (4.08%)

Table 3. Specific information for each cancer

Specific information for colorectal cancer			Specific information for breast cancer			Specific information for prostate cancer		
			Oestrogen receptor	**Positive**	1398 (83.0%)		**1 (Gleason score = 6)**	449 (42.56%)
				Negative	244 (14.5%)			
				Missing	43 (2.6%)		**2 (Gleason score = 3+4)**	299 (28.34%)
Location	**Right colon**	566 (26.99%)	**Progesterone receptor**	**Positive**	1237 (73.4%)			
				Negative	401 (23.8%)		**3 (Gleason score =4+3)**	120 (11.37%)
	Left colon	719 (34.29%)		**Missing**	47 (2.8%)	**Gleason grade**		
	Rectum-sigma	791 (37.72%)	**Her2**	**Positive**	294 (17.4%)		**4 (Gleason score = 8)**	83 (7.87%)
				Negative	1250 (74.2%)			
	Unknown	21 (1%)		**Missing**	141 (8.4%)		**5 (Gleason score 9 or 10)**	65 (6.16%)
			Intrinsic subtype	**Luminal A**	997 (59.2%)			
				Luminal B	331 (19.6%)			
				Her2	81 (4.8%)		**Missing**	39 (3.70%)
				Basal-like	130 (7.7%)			
Differentiation's degree	**I**	520 (24.8%)		**Luminal ONI***	91 (5.4%)	**PSA** (ng/ml)**		11.51 (±16.28)
	II	1100 (52.46%)		**Non-luminal ONI***	13 (0.8%)			
	III	247 (11.78%)		**Missing**	42 (2.5%)	**D'Amico**	**Low risk**	325 (30.81%)
	Not evaluable	230 (10.97%)	**Grade**	**I**	329 (19.5%)		**Intermediate risk**	425 (40.28%)
				II	520 (30.9%)			
				III	355 (21.1%)		**High risk**	284 (26.92%)
				Missing	481 (28.5%)		**Missing**	21 (1.99%)

*ONI: Otherwise Not Identified. **PSA: Prostate Specific Antigen

Table 4. First-line treatment

Treatment	Category	Colorectal cancer	Breast cancer	Prostate cancer
None (active surveillance)		-	-	38 (3.6%)
Surgery		Total: 1999 (95.3%)	Conservative: 1231(73.1%)	Prostatectomy: 639 (61.4%)
		Resection: 1800 (85.8%)		
		Palliative: 127 (6.1%)	Mastectomy: 454 (26.9%)	
		No resection: 61 (2.9%)		
		Others: 11 (0.5%)		
Chemotherapy	**Neoadjuvant**	427 (20.4%)	200 (11.9%)	1 (0.1%)
	Adjuvant	1024 (48.8%)	664 (39.4%)	1 (0.1%)
	Palliative	67 (3.2%)	25 (1.5%)	7 (0.7%)
Radiotherapy	**Neoadjuvant**	401 (19.1%)	5 (0.3%)	227 (21.5%)
	Adjuvant	82 (3.9%)	1132 (67.2%)	36 (3.4%)
	Palliative	5(0.2%)	21 (1.2%)	2 (0.2%)
Endocrine therapy	**Yes**	-	1023 (60.7%)	Adjuvant to surgery: 19 (1.8%)
				Adjuvant to radiotherapy: 99 (9.4%)
				Neoadjuvant: 102 (9.7%)
				Palliative: 69 (6.5%)
	No	-	662 (39.3%)	689 (65.3%)
Others (specify for each tumour)	**Endoscopy**	Complete resection: 107 (5.1%)	-	-
		Non-complete resection: 62 (3.0%)		
	Her2-targeted therapy	-	152(9.0%)	-
	Cryotherapy	-	-	21 (2.0%)
	Transurethral resection	-	-	4 (0.4%)

Colorectal Cancer

Out of 2140 patients with colorectal cancer, 2097 (98%) have been followed. They were 67±10.9 years-old on average at recruitment; 1334 (63.4%) were men. The first case was recruited on 18th of March 2007 and the follow-up was closed on 23rd of August 2018, accounting for 12813.8 person-years of follow-up. 819 (39.1%) cases died in this period; linearized mortality rate was 6.4 per 100 patient-years (95% CI: 6.0 – 6.8) (Table 2).

Most cases (1882, 90%) were adenocarcinoma; the most frequent location was rectum-sigma (37.7%) and the less frequent right colon (27%). 52% patients were at clinical stage II or less; in 110 patients (5.3%) we could not establish the clinical stage. 52.5% cancers were moderately differentiated (grade II) and 24.8% well differentiated (grade I) (Table 3).

Surgery was carried out in 1999 colorectal cancer patients; it was for palliative purposes in 127 patients (6.1%). 169 patients were treated via endoscopy, reaching complete resection in 107 of them. 1518 (72.4%) patients received chemotherapy; most of them (1451) for adjuvant or neoadjuvant purposes; 488 (23.2%) received radiotherapy (401 neoadjuvant, 82 adjuvant and only 5 palliative) (Table 4).

Five-year survival probability estimated via Kaplan-Meier was 71.6% (95% CI: 69.6 – 73.5) (Figure 2a). Survival was higher in women (74.4%, 95% CI: 71.0 – 77.2) than in men (70.0%, 95% CI: 67.5 – 72.4) ($p<0.001$) (Figure 2b). Five-year survival probability was 85.2% (81.0 – 88.6) in patients diagnosed in stage I, 84.0% (81.0 – 86.6) in stage II, 73.4% (69.6 – 76.9) in stage III and 27.6% (22.9 – 32.5) in stage IV (Figure 3a).

Breast Cancer

The maximum span for breast cancer follow-up was nine and a half years (from 13th July 2007 to 22nd March 2017). Follow-up was obtained for 1685 out of 1738 breast cancer patients (97%), adding 10931 person-years; 206 patients died in the follow-up; the linearized mortality rate was 1.9 per 100 patient-years (95% CI: 1.6 – 2.2).

Women with breast cancer were 56.5±12.6 year-old on average at recruitment; 65% were postmenopausal. The most usual type of tumour was ductal (75.7%), followed by lobular (6.5%). Most breast cancers were diagnosed at early stages (71% at stages I or II) and only 41 (2.4%) had metastasized at the time of diagnosis (Table 2). 83% cancers were oestrogen receptor positive, 73.4% progesterone receptor positive and 17.4% Her2 positive. Regarding intrinsic subtypes, 997 (59.2%) could be classified as luminal A, 331 (19.6%) as luminal B, 81 (4.8%) as Her2 and 130 (7.7%) as basal-like. According to grade of differentiation, moderately differentiated accounted for 30.9% breast cancers; well differentiated and bad differentiated accounted for about 20% cancers each. Grade could not be obtained from medical records in 481 patients (28.5%) (Table 3).

Conservative surgery was performed in 1231 (73.1%) patients and mastectomy in the remaining 454 (26.9%). Adjuvant or neoadjuvant chemotherapy was administered to 50.3% patients, while radiotherapy was used in 1158 women (68.7%), endocrine therapy was used in 1023 women (60.7%) and Her2-targeted therapy in 152 patients (9.0%) (Table 4). Kaplan-Meier 5-year survival with breast cancer was 90.7% (95%

CI: 89.2 – 92.0) (Figure 2c). Women diagnosed in stage I had 97% (95.5 – 98.1) 5-year survival probability, 91.9% (89.1 – 94.1) in stage II, 84.1% (77.8 – 88.7) in stage III and 38.5% (18.6 – 58.2) in stage IV (Figure 3b).

Prostate Cancer

A total of 1112 men with prostate cancer were recruited and 1055 (94.9%) have been followed-up; the first patient was included on 26th January, 2008 and the end of follow-up was on 13th July, 2018, adding 7169.6 person-year of follow-up. Patients were 65.9 years-old on average at recruitment. 119 patients died in the follow-up, making the linearized mortality rate 1.7 per 100 patient-years (95% CI: 1.4 – 2.0).

Almost all prostate cancers (99.9%) were adenocarcinoma; 496 (47%) were diagnosed at stage II and 132 (12.5%) at stage III (Table 2). The level of PSA gives an average of 11.5±16.3 ng/ml. Considering the Gleason score, 42.6% prostate cancers were well differentiated (Gleason grade = 1, i.e. Gleason score = 6); 28.3% were at Gleason grade 2 (Gleason score = 3+4), and only 14.0% were bad differentiated (Gleason grade 4 or 5; Gleason score ≥8); Gleason grade could not be established in 17.4% patients. D'Amico classification system results in 31.4% patients with low-risk cancer, 41.1% intermediate and 27.4% high-risk cancer (Table 3).

Thirty-eight prostate cancer patients were not treated medically at the beginning, being followed by active surveillance; prostatectomy was performed in 61.4% cases; radiotherapy in 265 patients (25.1%) and endocrine therapy in 289 patients (27.4%). A small number of patients were treated via trans-urethral resection, cryotherapy or chemotherapy (Table 4). Five-year survival probability by Kaplan-Meier was 93.7% (95% CI: 92.0 – 95.1) (Figure 2d). Survival probability 5 years after being diagnosed was 94.5% (88.1 – 97.5) for patients in stage I, 95.6% (93.3 – 97.2) in stage II, 92.4% (88.5 – 95.0) in stage III and 70.5 (42.8 – 88.6) in stage IV (Figure 3c).

STRENGTHS AND LIMITATIONS

In this article, we have described how three prospective cohorts on colorectal, breast and prostate cancers have been assembled from patients originally recruited for a case-control study, which makes 97% patients followed-up and accounts for more than 30,000 person-years. This is a main achievement of a network settled within the CIBERESP in 12 Spanish provinces. The study is population based and included only incident cancers; the amount of detailed information recorded as well as the availability of biological samples at recruitment will allow the identification of genetics, environmental, lifestyle and clinical prognosis factors in three frequent cancers in Spain. In this regard, a remarkable feature of the study is the feasibility of studying cancer risk factors as putative prognosis factors; for example, risk factors already analysed in the case-control phase have been diet, circadian cycle disruption, some drugs, endocrine disruptors, artificial light or proximity to green spaces; information regarding these risk factors was recorded at recruitment and is available for a prognosis factor analysis in the follow-up (see Supplementary material for a complete reference list of MCC-Spain articles).

Obtaining information on personal history, occupational exposures, diet, physical exercise or other lifestyle components is somewhat subjective as both patients and interviewers could be prone to be influenced by

their feelings or beliefs about the hypotheses under study, eventually leading to differential misclassification bias. This could hardly have occurred in this study: Firstly, patients were not aware of the hypotheses. Secondly, interviewers were familiar with the case-control study, not with the cohort design as it was decided later; therefore, if interviewers or patients have introduced some misclassification, it could probably have been non-differential, eventually leading to bias towards the null, which would make more robust the positive findings in this cohort study.

This study has also some weaknesses: Firstly, multicentre studies are double edged; they are needed in order to include many patients, but they could introduce heterogeneity in both the information gathered and the way patients are treated. In this regard, the analysis of prognosis factors should be adjusted for the hospital of recruitment. Secondly, 113 participating patients have been lost (43 with colorectal cancer, 53 with breast cancer and 57 with prostate cancer); we have tried to minimise it by searching information in three ways: medical records, phone calls and IND; however, we cannot rule out that some patients without follow-up could have died. It is noteworthy that -due to the small number of patients without follow-up- the maximum bias it could introduce in our survival estimates is 2% for colorectal cancer, 3% for breast cancer and 5% for prostate cancer. Thirdly, we have not obtained information on lifestyle changes after diagnosis, which limits lifestyle analysis to habits before cancer appearance. Fourthly, the number of patients included in our cohorts is small compared with those based on cancer registries, limiting the analysis of subgroups.

Summarizing, the MCC-Spain study has assembled three cohorts with about 4,700 cancer patients accounting for 30,000 patient-years of follow-up, with only 3% patient withdrawals. The information gathered at recruitment will allow to prospectively investigate clinical, lifestyle, environmental and genetic variables as prognosis factors in colorectal, breast and prostate cancers in Spain.

COLLABORATION

MCC-Spain already participates in international consortiums as Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO; https://www.fredhutch.org/en/labs/phs/projects/cancer-prevention/projects/gecco.html), Breast Cancer Association Consortium (BCAC; http://bcac.ccge.medschl.cam.ac.uk/) and Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL; http://practical.icr.ac.uk/blog/), where MCC-Spain would contribute to study interactions among the putative prognosis factors in vast population samples.

DATA STATEMENT

Permission to use the study database (individual-level deidentified patient data) will be granted to researchers outside the study group, after revision and approval of each request by the Steering Committee. Any kind of collaboration are encouraged.

ACKNOWLEDGEMENTS

We thank all the subjects who participated in the study and all MCC-Spain collaborators.

INFORMED CONSENT:

Informed consent was obtained from all individual participants included in the study.

AUTHOR STATEMENT:

Jessica Alonso-Molero, Antonio J Molina, Jose J Jiménez-Moleón have contributed to the conception and design of the study, considering the same contribution. JAM, AJM, JJJM, BPG, VM, VM, PA, EA, SS, IS, GFT, JA, DS, RMG, MDC, NA, GCV, MP, MK, JL have acquired data and have been involved in drafting the manuscript. All of them read and approved the final manuscript.

FUNDING:

The study was partially funded by the "Accion Transversal del Cancer", approved on the Spanish Ministry Council on the 11th October 2007, by the Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PS09/00773-Cantabria, PS09/01286-León, PS09/01903-Valencia, PS09/02078-Huelva, PS09/01662-Granada, PI11/01403, PI11/01889-FEDER, PI11/00226, PI11/01810, PI11/02213, PI12/00488, PI12/00265, PI12/01270, PI12/00715, PI12/00150, PI14/01219, PI14/0613, PI15/00069, PI15/00914, PI15/01032, PI17CIII/00034, PI18/00181), by the Fundación Marqués de Valdecilla (API 10/09), by the ICGC International Cancer Genome Consortium CLL (The ICGC CLL-Genome Project is funded by Spanish Ministerio de Economía y Competitividad (MINECO) through the Instituto de Salud Carlos III (ISCIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036)), by the Junta de Castilla y León (LE22A10-2), by the Consejería de Salud of the Junta de Andalucía (PI-0571-2009, PI-0306-2011, salud201200057018tra), by the Conselleria de Sanitat of the Generalitat Valenciana (AP_061/10), by the Recercaixa (2010ACUP 00310), by the Regional Government of the Basque Country, by the Consejería de Sanidad de la Región de Murcia, by the European Commission grants FOOD-CT-2006-036224-HIWATE, by the Spanish Association Against Cancer (AECC) Scientific

Foundation (GCTRA18022MORE), by the Catalan Government- Agency for Management of University and Research Grants (AGAUR) grants 2017SGR723 and 2014SGR850, by the Fundación Caja de Ahorros de Asturias and by the University of Oviedo. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya.

SAMPLES:

Biological samples were stored at the biobanks supported by Instituto de Salud Carlos III- FEDER: Parc de Salut MAR Biobank (MARBiobanc) (RD09/0076/00036), “Biobanco La Fe” (RD 09 0076/00021) and FISABIO Biobank (RD09 0076/00058). Also at the Public Health Laboratory from Gipuzkoa, the Basque Biobank, the ICOBIOBANC (sponsored by the Catalan Institute of Oncology), the IUOPA Biobank from the University of Oviedo and the ISCIII Biobank.

GENOTYPING:

SNP genotyping services were provided by the Spanish "Centro Nacional de Genotipado” (CEGEN-ISCIII)".

CONFLICT OF INTEREST:

The authors declare that they have no conflict of interest

REFERENCES

1 Saadatmand S, Bretveld R, Siesling S, *et al.* Influence of tumour stage at breast cancer detection on survival in modern times: Population based study in 173 797 patients. *BMJ* 2015;**351**:h4901. doi:10.1136/bmj.h4901

2 Shukla N, Hagenbuchner M, Win KT, *et al.* Breast cancer data analysis for survivability studies and prediction. *Comput Methods Programs Biomed* 2018;**155**:199–208. doi:10.1016/j.cmpb.2017.12.011

3 Mirza AN, Mirza NQ, Vlastos G, *et al.* Prognostic Factors in Node-Negative Breast Cancer. *Ann Surg* 2002;**235**:10–26. doi:10.1097/00000658-200201000-00003

4 Zhang Z yu, Luo Q feng, Yin X wei, *et al.* Nomograms to predict survival after colorectal cancer resection without preoperative therapy. *BMC Cancer* 2016;**16**:1–21. doi:10.1186/s12885-016-2684-4

5 Merriel SWD, May MT, Martin RM. Predicting prostate cancer progression: Protocol for a retrospective cohort study to identify prognostic factors for prostate cancer outcomes using routine primary care data. *BMJ Open* 2018;**8**:1–5. doi:10.1109/TVLSI.2018.2801302

6 Riboli E, Hunt K, Slimani N, *et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002;**5**:1113–24. doi:10.1079/phn2002394

7 Lagendijk M, van Maaren MC, Saadatmand S, *et al.* Breast conserving therapy and mastectomy revisited: Breast cancer-specific survival and the influence of prognostic factors in 129,692 patients. *Int J Cancer* 2018;**142**:165–75. doi:10.1002/ijc.31034

8 Cardwell CR, Hicks BM, Hughes C, *et al.* Statin Use after colorectal cancer diagnosis and survival: A population-based cohort study. *J Clin Oncol* 2014;**32**:3177–83. doi:10.1200/JCO.2013.54.4569

9 Pettersson A, Robinson D, Garmo H, *et al.* Age at diagnosis and prostate cancer treatment and prognosis: A population-based cohort study. *Ann Oncol* 2018;**29**:377–85. doi:10.1093/annonc/mdx742

10 Leone JP, Leone J, Zwenger AO, *et al.* Prognostic Significance of Tumor Subtypes in Women With Breast Cancer According to Stage: A Population-based Study. *Am J Clin Oncol* 2019;**42**:588–95. doi:10.1097/COC.0000000000000563

11 Li Y, Feng Y, Dai W, *et al.* Prognostic Effect of Tumor Sidedness in Colorectal Cancer: A SEER-Based Analysis. *Clin Colorectal Cancer* 2019;**18**:e104–16. doi:10.1016/j.clcc.2018.10.005

12 Roy S, Morgan SC. Who Dies From Prostate Cancer? An Analysis of the Surveillance, Epidemiology and End Results Database. *Clin Oncol* 2019;**31**:630–6. doi:10.1016/j.clon.2019.04.012

13 Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, *et al.* Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *GacSanit* 2015;**29**:308–15. doi:10.1016/j.gaceta.2014.12.003

14 Navarro C. El Índice Nacional de Defunciones: Un avance en la accesibilidad de los datos de mortalidad largamente esperado. *Gac Sanit* 2006;**20**:421–3. doi:10.1157/13096513

15 Ware J, Kosinki M, Turner-Bowker D, *et al.* How to score Version 2 of the SF-12 Health survey. *Lincoln, RJ Qual Inc* 2004.

16 Colwell HH, Mathias SD, Solutions HO, *et al.* Psychometric Evaluation of the FACT Colorectal Cancer Symptom Index (FCSI-9): Reliability, Validity, Responsiveness, and Clinical Meaningfulness HILARY. *Oncologist* 2010;**15**:308–16. doi:10.1634/theoncologist.2009-0034

17 Garcia SF, Rosenbloom SK, Beaumont JL, *et al.* Priority symptoms in advanced breast cancer: Development and initial validation of the national comprehensive cancer Network-Functional Assessment of Cancer Therapy-Breast Cancer Symptom Index (NFBSI-16). *Value Heal*

2012;**15**:183–90. doi:10.1016/j.jval.2011.08.1739

18 Charlson M, Pompei P, Ales K, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.

19 Esper P, Mo F, Chodak G, *et al.* Measuring quality of life in men with prostate cancer using the functionale assessment of cancer therapy-prostate instrument. *Adult Urol* 1997;**30**:920–8.

20 Barry MJ, Fowler FJ, O'Leary MP, *et al.* The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol* 1992;**148**:1549–57.

21 Papantoniou K, Castaño-Vinyals G, Espinosa A, *et al.* Night shift work, chronotype and prostate cancer risk in the MCC-Spain case-control study. *Int J Cancer* 2015;**137**:1147–57. doi:10.1002/ijc.29400

22 Kogevinas M, Espinosa A, Papantoniou K, *et al.* Prostate cancer risk decreases following cessation of night shift work. *Int J cancer* Published Online First: 24 June 2019. doi:10.1002/ijc.32528

23 Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, *et al.* The use of antihypertensive medication and the risk of breast cancer in a case-control study in a Spanish population: The MCC-Spain study. *PLoS One* 2016;**11**:1–14. doi:10.1371/journal.pone.0159672

24 Castelló A, Boldo E, Pérez-Gómez B, *et al.* Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. *Maturitas* 2017;**103**:8–15. doi:10.1016/j.maturitas.2017.06.020

25 Dierssen-sotos T, Gómez-acebo I, Palazuelos C, *et al.* Validating a breast cancer score in Spanish women . The MCC-Spain study. 2018;:1–8. doi:10.1038/s41598-018-20832-0

26 Ibáñez-Sanz G, Díez-Villanueva A, Alonso MH, *et al.* Risk Model for Colorectal Cancer in Spanish Population Using Environmental and Genetic Factors: Results from the MCC-Spain study. *Sci Rep* 2017;**7**:43263. doi:10.1038/srep43263

27 Romaguera D, Gracia-Lavedan E, Molinuevo A, *et al.* Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case–control study. *Int J Cancer* 2017;**141**:83–93. doi:10.1002/ijc.30722

FIGURE LEGENDS

Figure 1. Flow chart of the participants in the MCC-Spain study

Figure 2. Kaplan-Meier survival estimates for colorectal cancer (2a), colorectal cancer by sex (2b), breast cancer (2c) and prostate cancer (2d)

Figure 3. Kaplan-Meier estimates by stage at diagnosis for colorectal cancer (3a), breast cancer (3b) and prostate cancer (3c)

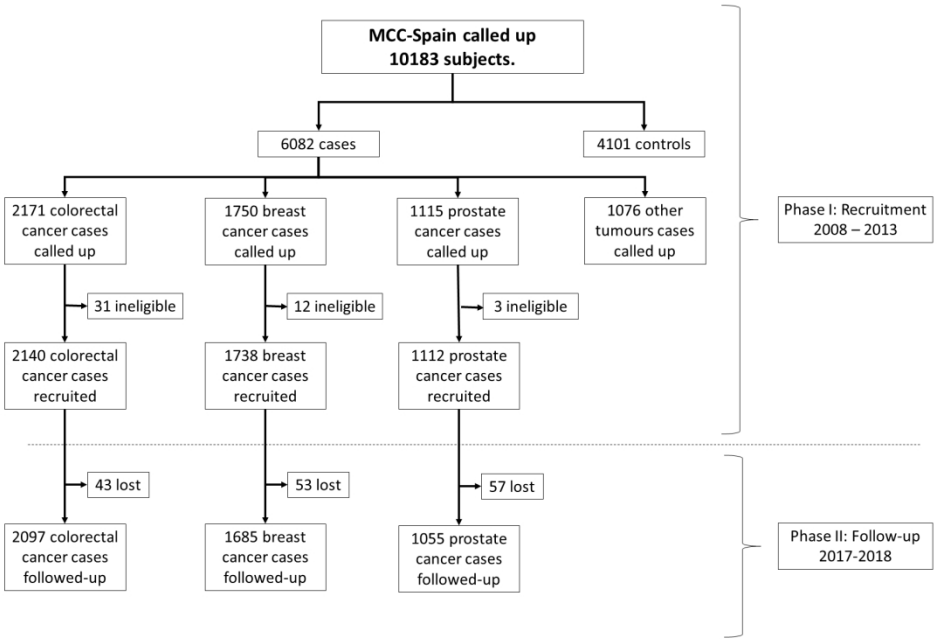

Figure 1. Flow chart of the participants in the MCC-Spain study

500x338mm (96 x 96 DPI)

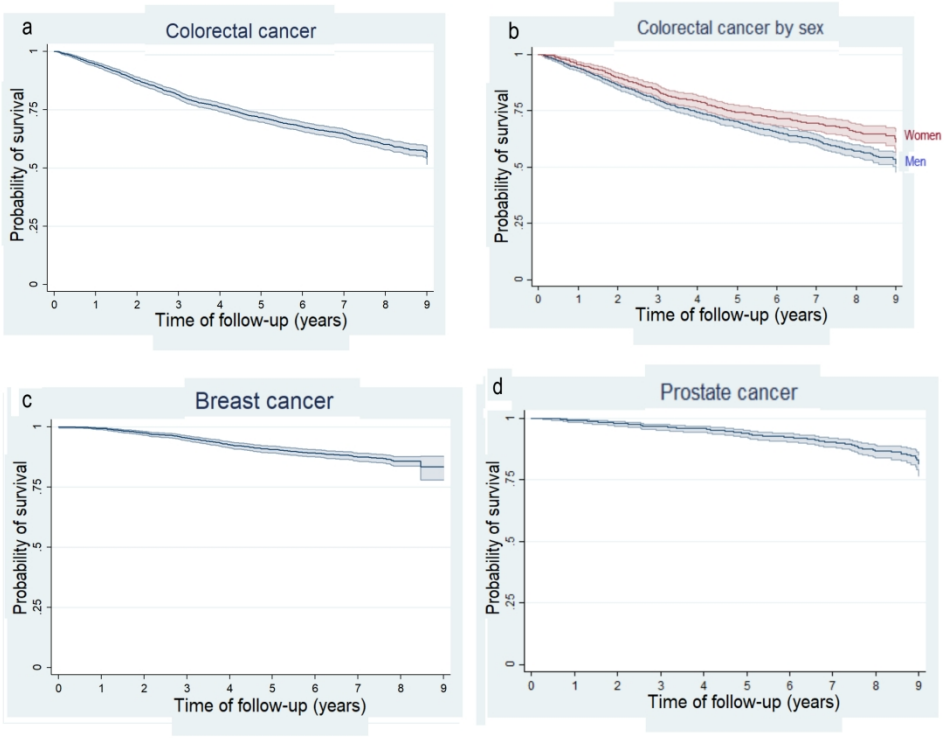

Figure 2. Kaplan-Meier survival estimates for colorectal cancer (2a), colorectal cancer by sex (2b), breast cancer (2c) and prostate cancer (2d)

727x602mm (96 x 96 DPI)

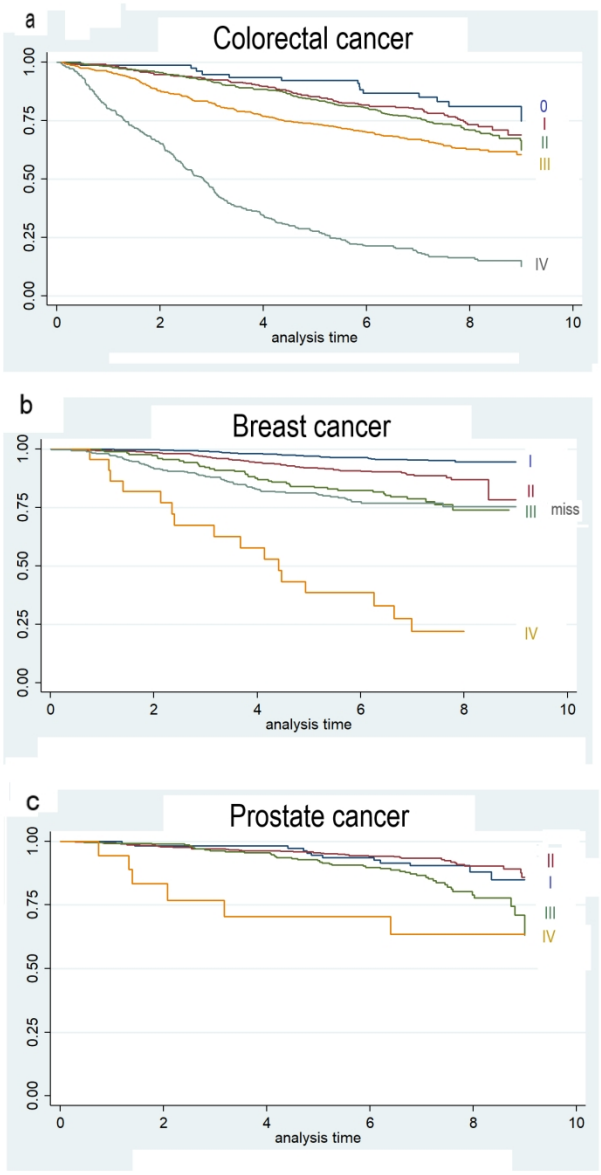

Figure 3. Kaplan-Meier estimates by stage at diagnosis for colorectal cancer (3a), breast cancer (3b) and prostate cancer (3c)

Supplementary Table 1. Provinces and hospital of recruitment

Province	Hospital	Colorectal cancer	Breast cancer	Prostate cancer
Asturias	Hospital de Cabueñes	77	70	16
Barcelona	Hospital Clinic	69	47	53
Barcelona	Hospital de Bellvitge – ICO	375	109	-
Barcelona	Hospital del Mar	222	136	152
Barcelona	Hospital Germans Trias i Pujol	30	-	199
Cantabria	Hospital Universitario Marqués de Valdecilla	151	141	175
Gipuzkoa	Hospital Donostia	119	126	-
Gipuzkoa	Instituto Oncológico	-	100	-
Girona	Hospital Dr. Josep Trueta	-	21	-
Girona	Hospital Santa Caterina	-	26	-
Granada	Hospital San Cecilio	164	-	64
Huelva	Hospital Infanta Elena	16	24	16
Huelva	Hospital Juan Ramón Jiménez	55	84	36
León	Hospital de León	390	226	-
Madrid	Hospital La Paz	110	164	155
Madrid	Hospital Ramón y Cajal	122	177	160
Murcia	Hospital Morales Messeguer	34	-	-
Navarra	Complejo Hospitalario de Navarra A (Hospital de Navarra)	76	112	-
Navarra	Complejo Hospitalario de Navarra B (Virgen del Camino)	49	114	-
Valencia	Hospital Dr. Peset	25	4	-
Valencia	Hospital La Fe	56	57	86

Supplementary Table 2. Previous results in the MCC-Spain study

Supp. Reference	Journal	Year of publication	Cancer	Exposure
1	Environ Res	2012	NA	Disinfection by-products in municipal drinking water
2	Gac Sanit	2012	Breast, prostate	Screening practices and lifestyles
3	BJU Int	2012	Prostate	Anogenital distance
4	Gac Sanit	2013	NA	Nitrate and trace elements in municipal and bottled water
5	Int J Cancer	2015	Prostate	Night shift work and chronotype
6	Gac Sanit	2015	Colorectal, breast, prostate, gastric, CLL	Rational and study design for case-control
7	J Gen Virol	2015	CLL	Polyomaviruses
8	Infect Agent Cancer	2015	CLL	Aberrant Epstein-Barr virus
9	Menopause	2015	NA	Hormonal contraception and postmenopausal hormone therapy
10	Sci Total Environ	2015	NA	Persistent organic pollutants in adult population
11	Acta Diabetol	2016	Breast	Diabetes and diabetes treatment
12	Eur J Epidemiol	2016	Breast	Night shift work
13	Int J Cancer	2016	Colorectal	Streptococcus gallolyticus
14	Cancer Epidemiol	2016	Breast	Perinatal and childhood factors
15	Environ Health Perspect	2016	Breast	Ingested nitrate
16	Int J Cancer	2016	Colorectal	Ingested nitrate
17	Environ Health Perspect	2016	Breast	Xenoestrogen burden
18	Occup Environ Med	2016	Gastric	Night shift work
19	Cancer Epidemiol	2016	Prostate	Perinatal and childhood factors
20	Int J Cancer	2016	CLL	Night shift work
21	PLoS One	2016	Breast	Antihypertensive medication
22	BMC Cancer	2016	Breast	Non-steroidal anti-inflammatory drugs
23	PLoS One	2016	Colorectal, gastric	Menstrual and reproductive factors
24	Eur J Nutr	2017	CLL	Fruit and vegetable intake and vit C transporter gene
25	Environ Health Perspect	2017	Colorectal	Trihalomethanes in drinking water

26	Prev Med	2017	Colorectal	Drugs affecting renin-angiotensin system
27	Sci Rep	2017	Colorectal	Environmental and genetic factors
28	Scand J Work Environ Health	2017	Colorectal	Shift work
29	Int J Cancer	2017	Colorectal, breast, prostate	Nutrition-based cancer prevention guidelines
30	Front Microbiol	2017	Colorectal	Helicobacter pylori
31	PLoS One	2017	Gastric	Physical activity
32	Helicobacter	2017	NA	Helicobacter pylori in adult population
33	Maturitas	2017	Breast	Dietary patterns
34	Sci Rep	2017	Prostate	Environmental and genetic factors
35	Cancer Epidemiol	2017	Gastric	Helicobacter pylori
36	BMC Med Genet	2017	Colorectal	SMAD7 gene and Mediterranean diet
37	Gastric Cancer	2017	Gastric	Dietary patterns
38	Eur J Nutr	2018	Colorectal	Meat intake, cooking methods and doneness
39	J Urol	2018	Prostate	Dietary patterns
40	Environ Int	2018	Breast	Trihalomethanes in drinking water
41	Cancer Epidemiol	2018	CLL	CLL etiology (review)
42	Sci Rep	2018	Colorectal	Chondroitin sulphate and glucosamine
43	Sci Rep	2018	Breast	Risk score
44	Environ Pollut	2018	Breast	Residential proximity to industrial installations
45	BMC Cancer	2018	Breast	Reproductive factors and genetic hormonal pathways
46	Maturitas	2018	Breast	Meat intake, methods of cooking
47	J Steroid Biochem Mol Biol	2018	Breast	Vitamin D
48	Environ Health Perspect	2018	Breast, prostate	Artificial light-at-night
49	Haematologica	2018	CLL	Dietary patterns
50	Int J Cancer	2018	Breast, prostate	Mistimed eating patterns
51	Stat Methods Med Res	2018	NA	Compositional analysis of dietary patterns
52	Int J Hyg Environ Health	2018	Breast	Residential proximity to green spaces
53	PLos One	2018	Breast, prostate	Pigmentation phototype
54	BMC Public Health	2018	NA	Non-steroidal anti-inflammatory drugs consumption
55	Environ Int	2018	Colorectal	Sun exposure and vit D

56	Nutrients	2018	Prostate	Dietary zinc
57	Eur J Nutr	2019	Colorectal	Dietary patterns
58	Eur J Nutr	2019	Colorectal	Dietary non-enzymatic antioxidant capacity
59	Br J Haematol	2019	CLL	Insulin-like growth factor
60	Eur J Cancer Prev	2019	NA	Helicobacter pylori seroprevalence
61	Environ Int	2019	Breast, prostate	Alkylphenolic compounds
62	Nutrients	2019	NA	Mediterranean diet
63	Nutrients	2019	Gastric	Flavonoids
64	Eur J Nutr	2019	Breast	Fatty acid intake
65	Cancer Epidemiol	2019	Gastric	Epstein-Barr virus
66	Int J Cancer	2019	Prostate	Cessation of night shift work
67	Nutrients	2019	Colorectal, breast	Dietary inflammatory index and dietary non-enzymatic antioxidant capacity
68	Breast Cancer Res Treat	2019	Breast	Physical activity
69	Sci Rep	2019	Colorectal	Flagelin C and Streptococcus gallolyticus proteins

CLL: Chronic lymphocytic leukaemia. NA: Not applicable

Supplementary reference list

1. Villanueva CM, Castaño-Vinyals G, Moreno V, et al. Concentrations and correlations of disinfection by-products in municipal drinking water from an exposure assessment perspective. Environ Res. 2012;114:1-11. doi: 10.1016/j.envres.2012.02.002.
2. Perea MD, Castaño-Vinyals G, Altzibar JM, et al. [Cancer screening practices and associated lifestyles in population controls of the Spanish multi-case control study]. Gac Sanit. 2012;26:301-10. doi: 10.1016/j.gaceta.2012.01.020.
3. Castaño-Vinyals G, Carrasco E, Lorente JA, et al. Anogenital distance and the risk of prostate cancer. BJU Int. 2012;110(11 Pt B):E707-10. doi: 10.1111/j.1464-410X.2012.11516.x.
4. Espejo-Herrera N, Kogevinas M, Castaño-Vinyals G, et al and Multicase Control Study of Cancer (MCC)-Spain Water Working Group. Nitrate and trace elements in municipal and bottled water in Spain. Gac Sanit. 2013 Mar-Apr;27(2):156-60. doi:10.1016/j.gaceta.2012.02.002.
5. Papantoniou K, Castaño-Vinyals G, Espinosa A, et al. Night shift work, chronotype and prostate cancer risk in the MCC-Spain case-control study. Int J Cancer. 2015;137:1147-57. doi: 10.1002/ijc.29400.
6. Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, et al and MCC-Spain Study Group. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. Gac Sanit. 2015;29:308-15. doi: 10.1016/j.gaceta.2014.12.003.
7. Robles C, Casabonne D, Benavente Y, et al. Seroreactivity against Merkel cell polyomavirus and other polyomaviruses in chronic lymphocytic leukaemia, the MCC-Spain study. J Gen Virol. 2015;96:2286-92. doi: 10.1099/vir.0.000167.
8. Casabonne D, Benavente Y, Robles C, et al. Aberrant Epstein-Barr virus antibody patterns and chronic lymphocytic leukemia in a Spanish multicentric case-control study. Infect Agent Cancer. 2015;10:5. doi: 10.1186/1750-9378-10-5.
9. Costas L, Sequera VG, Quesada P, et al. Hormonal contraception and postmenopausal hormone therapy in Spain: time trends and patterns of use. Menopause. 2015;22:1138-46. doi: 10.1097/GME.0000000000000487.
10. Fernández-Rodríguez M, Arrebola JP, Artacho-Cordón F, et al. Levels and predictors of persistent organic pollutants in an adult population from four Spanish regions. Sci Total Environ. 2015;538:152-61. doi: 10.1016/j.scitotenv.2015.07.162.
11. García-Esquinas E, Guinó E, Castaño-Vinyals G, et al. Association of diabetes and diabetes treatment with incidence of breast cancer. Acta Diabetol. 2016;53:99-107. doi: 10.1007/s00592-015-0756-6.
12. Papantoniou K, Castaño-Vinyals G, Espinosa A, et al. Breast cancer risk and night shift work in a case-control study in a Spanish population. Eur J Epidemiol. 2016;31:867-78. doi: 10.1007/s10654-015-0073-y.

13. Butt J, Romero-Hernández B, Pérez-Gómez B, et al. Association of Streptococcus gallolyticus subspecies gallolyticus with colorectal cancer: Serological evidence. Int J Cancer. 2016;138:1670-9. doi: 10.1002/ijc.29914.

14. Lope V, García-Esquinas E, Pérez-Gómez B, et al. Perinatal and childhood factors and risk of breast cancer subtypes in adulthood. Cancer Epidemiol. 2016;40:22-30. doi: 10.1016/j.canep.2015.11.004.

15. Espejo-Herrera N, Gracia-Lavedan E, Pollan M, et al. Ingested Nitrate and Breast Cancer in the Spanish Multicase-Control Study on Cancer (MCC-Spain). Environ Health Perspect. 2016;124:1042-9. doi: 10.1289/ehp.1510334.

16. Espejo-Herrera N, Gràcia-Lavedan E, Boldo E, et al.Colorectal cancer risk and nitrate exposure through drinking water and diet. Int J Cancer. 2016;139:334-46. doi: 10.1002/ijc.30083.

17. Pastor-Barriuso R, Fernández MF, Castaño-Vinyals G, et al. Total Effective Xenoestrogen Burden in Serum Samples and Risk for Breast Cancer in a Population-Based Multicase-Control Study in Spain. Environ Health Perspect. 2016;124:1575-1582.

18. Gyarmati G, Turner MC, Castaño-Vinyals G, et al. Night shift work and stomach cancer risk in the MCC-Spain study. Occup Environ Med. 2016;73:520-7. doi: 10.1136/oemed-2016-103597.

19. Lope V, García-Esquinas E, Ruiz-Dominguez JM, et al. Perinatal and childhood factors and risk of prostate cancer in adulthood: MCC-Spain case-control study. Cancer Epidemiol. 2016;43:49-55. doi: 10.1016/j.canep.2016.06.012.

20. Costas L, Benavente Y, Olmedo-Requena R, et al. Night shift work and chronic lymphocytic leukemia in the MCC-Spain case-control study. Int J Cancer. 2016;139:1994-2000. doi: 10.1002/ijc.30272.

21. Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, et al. The Use of Antihypertensive Medication and the Risk of Breast Cancer in a Case-Control Study in a Spanish Population: The MCC-Spain Study. PLoS One. 2016;11:e0159672. doi: 10.1371/journal.pone.0159672.

22. Dierssen-Sotos T, Gómez-Acebo I, de Pedro M, et al. Use of non-steroidal anti-inflammatory drugs and risk of breast cancer: The Spanish Multi-Case-control (MCC) study. BMC Cancer. 2016;16:660. doi: 10.1186/s12885-016-2692-4.

23. Lope V, Fernández de Larrea N, Pérez-Gómez B, et al. Menstrual and Reproductive Factors and Risk of Gastric and Colorectal Cancer in Spain. PloS One. 2016;11:e0164620. doi: 10.1371/journal.pone.0164620.

24. Casabonne D, Gracia E, Espinosa A, et al. Fruit and vegetable intake and vitamin C transporter gene (SLC23A2) polymorphisms in chronic lymphocytic leukaemia. Eur J Nutr. 2017;56:1123-1133. doi: 10.1007/s00394-016-1162-8.

25. Villanueva CM, Gracia-Lavedan E, Bosetti C, et al. Colorectal Cancer and Long-Term Exposure to Trihalomethanes in Drinking Water: A Multicenter Case-Control Study in Spain and Italy. Environ Health Perspect. 2017;125:56-65. doi: 10.1289/EHP155.

26. Dierssen-Sotos T, Gómez-Acebo I, Palazuelos C, et al. Relationship between drugs affecting the renin-angiotensin system and colorectal cancer: The MCC-Spain study. Prev Med. 2017;99:178-184. doi: 10.1016/j.ypmed.2017.01.011.

27. Ibáñez-Sanz G, Díez-Villanueva A, Alonso MH, et al. Risk Model for Colorectal Cancer in Spanish Population Using Environmental and Genetic Factors: Results from the MCC-Spain study. Sci Rep. 2017;7:43263. doi: 10.1038/srep43263.

28. Papantoniou K, Castaño-Vinyals G, Espinosa A, et al. Shift work and colorectal cancer risk in the MCC-Spain case-control study. Scand J Work Environ Health. 2017;43:250-259. doi: 10.5271/sjweh.3626.

29. Romaguera D, Gracia-Lavedan E, Molinuevo A, et al. Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case-control study. Int J Cancer. 2017;141:83-93. doi: 10.1002/ijc.30722.

30. Fernández de Larrea-Baz N, Michel A, Romero B, et al. Helicobacter pylori Antibody Reactivities and Colorectal Cancer Risk in a Case-control Study in Spain. Front Microbiol. 2017 29;8:888. doi: 10.3389/fmicb.2017.00888. Huerta JM, Chirlaque MD, Molina AJ, et al. Physical activity domains and risk of gastric adenocarcinoma in the MCC-Spain case-control study. PloS One. 2017;12:e0179731. doi: 10.1371/journal.pone.0179731.

31. Fernández-de-Larrea N, Michel A, Romero B, et al. Antibody reactivity against Helicobacter pylori proteins in a sample of the Spanish adult population in 2008-2013. Helicobacter. 2017;22. doi: 10.1111/hel.12401.

32. Castelló A, Boldo E, Pérez-Gómez B, et al. Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. Maturitas. 2017;103:8-15. doi: 10.1016/j.maturitas.2017.06.020.

33. Gómez-Acebo I, Dierssen-Sotos T, Fernandez-Navarro P, et al. Risk Model for Prostate Cancer Using Environmental and Genetic Factors in the Spanish Multi-Case-Control (MCC) Study. Sci Rep. 2017;7:8994. doi: 10.1038/s41598-017-09386-9.

34. Fernández de Larrea-Baz N, Pérez-Gómez B, Michel A, et al. Helicobacter pylori serological biomarkers of gastric cancer risk in the MCC-Spain case-control Study. Cancer Epidemiol. 2017;50(Pt A):76-84. doi: 10.1016/j.canep.2017.08.002.

35. Alonso-Molero J, González-Donquiles C, Palazuelos C, et al. The RS4939827 polymorphism in the SMAD7 GENE and its association with Mediterranean diet in colorectal carcinogenesis. BMC Med Genet. 2017;18:122. doi: 10.1186/s12881-017-0485-5.

36. Castelló A, Fernández de Larrea N, Martín V, et al and MCC-Spain researchers. High adherence to the Western, Prudent, and Mediterranean dietary patterns and risk of gastric adenocarcinoma: MCC-Spain study. Gastric Cancer. 2018;21:372-382. doi: 10.1007/s10120-017-0774-x.

37. de Batlle J, Gracia-Lavedan E, Romaguera D, et al. Meat intake, cooking methods and doneness and risk of colorectal tumours in the Spanish multicase-control study (MCC-Spain). Eur J Nutr. 2018;57:643-653. doi: 10.1007/s00394-016-1350-6.

38. Castelló A, Boldo E, Amiano P, et al and MCC-Spain Researchers. Mediterranean Dietary Pattern is Associated with Low Risk of Aggressive Prostate Cancer: MCC-Spain Study. J Urol. 2018;199:430-437. doi: 10.1016/j.juro.2017.08.087.

39. Font-Ribera L, Gràcia-Lavedan E, Aragonés N, et al. Long-term exposure to trihalomethanes in drinking water and breast cancer in the Spanish multicase-control study on cancer (MCC-SPAIN). Environ Int. 2018;112:227-234. doi: 10.1016/j.envint.2017.12.031.

40. Benavente Y, Casabonne D, Costas L, et al. Established and suggested exposures on CLL/SLL etiology: Results from the CLL-MCC-Spain study. Cancer Epidemiol. 2018;52:106-111. doi: 10.1016/j.canep.2017.12.012.

41. Ibáñez-Sanz G, Díez-Villanueva A, Vilorio-Marqués L, et al. Possible role of chondroitin sulphate and glucosamine for primary prevention of colorectal cancer. Results from the MCC-Spain study. Sci Rep. 2018;8:2040. doi: 10.1038/s41598-018-20349-6.

42. Dierssen-Sotos T, Gómez-Acebo I, Palazuelos C, et al. Validating a breast cancer score in Spanish women. The MCC-Spain study. Sci Rep. 2018;8:3036. doi: 10.1038/s41598-018-20832-0.

43. García-Pérez J, Lope V, Pérez-Gómez B, et al. Risk of breast cancer and residential proximity to industrial installations: New findings from a multicase-control study (MCC-Spain). Environ Pollut. 2018;237:559-568. doi: 10.1016/j.envpol.2018.02.065.

44. Dierssen-Sotos T, Palazuelos-Calderón C, Jiménez-Moleón JJ, et al. Reproductive risk factors in breast cancer and genetic hormonal pathways: a gene-environment interaction in the MCC-Spain project. BMC Cancer. 2018;18:280. doi: 10.1186/s12885-018-4182-3.

45. Boldo E, Castelló A, Aragonés N, et al and MCC-Spain researchers. Meat intake, methods and degrees of cooking and breast cancer risk in the MCC-Spain study. Maturitas. 2018;110:62-70. doi: 10.1016/j.maturitas.2018.01.020.

46. Lope V, Castelló A, Mena-Bravo A, et al. Serum 25-hydroxyvitamin D and breast cancer risk by pathological subtype (MCC-Spain). J Steroid Biochem Mol Biol. 2018;182:4-13. doi: 10.1016/j.jsbmb.2018.04.005.

47. Garcia-Saenz A, Sánchez de Miguel A, Espinosa A, et al. Evaluating the Association between Artificial Light-at-Night Exposure and Breast and Prostate Cancer Risk in Spain (MCC-Spain Study). Environ Health Perspect. 2018;126:047011. doi: 10.1289/EHP1837.

48. Solans M, Castelló A, Benavente Y, et al. Adherence to the Western, Prudent, and Mediterranean dietary patterns and chronic lymphocytic leukemia in the MCC-Spain study. Haematologica. 2018;103:1881-1888. doi: 10.3324/haematol.2018.192526.

49. Kogevinas M, Espinosa A, Castelló A, et al. Effect of mistimed eating patterns on breast and prostate cancer risk (MCC-Spain Study). Int J Cancer. 2018;143:2380-2389. doi: 10.1002/ijc.31649.

50. Solans M, Coenders G, Marcos-Gragera R, et al. Compositional analysis of dietary patterns. Stat Methods Med Res. 2018 (In press). doi: 10.1177/0962280218790110.

51. O'Callaghan-Gordo C, Kogevinas M, Cirach M, et al. Residential proximity to green spaces and breast cancer risk: The multicase-control study in Spain (MCC-Spain). Int J Hyg Environ Health. 2018;221:1097-1106. doi: 10.1016/j.ijheh.2018.07.014.

52. Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, et al. Pigmentation phototype and prostate and breast cancer in a select Spanish population-A Mendelian randomization analysis in the MCC-Spain study. PloS One. 2018;13:e0201750. doi: 10.1371/journal.pone.0201750.

53. Gómez-Acebo I, Dierssen-Sotos T, de Pedro M, et al. The MCC-Spain study. BMC Public Health. 2018;18:1134. doi: 10.1186/s12889-018-6019-z.

54. Vallès X, Alonso MH, López-Caleya JF, et al. Colorectal cancer, sun exposure and dietary vitamin D and calcium intake in the MCC-Spain study. Environ Int. 2018 Dec;121(Pt 1):428-434. doi: 10.1016/j.envint.2018.09.030.

55. Gutiérrez-González E, Castelló A, Fernández-Navarro P, et al. Dietary Zinc and Risk of Prostate Cancer in Spain: MCC-Spain Study. Nutrients. 2018 Dec 20;11. pii: E18. doi: 10.3390/nu11010018.

56. Castelló A, Amiano P, Fernández de Larrea N, et al and MCC-Spain researchers. Low adherence to the western and high adherence to the mediterranean dietary patterns could prevent colorectal cancer. Eur J Nutr. 2019;58:1495-1505. doi: 10.1007/s00394-018-1674-5.

57. Amiano P, Molina-Montes E, Molinuevo A, et al. Association study of dietary non-enzymatic antioxidant capacity (NEAC) and colorectal cancer risk in the Spanish Multicase-Control Cancer (MCC-Spain) study. Eur J Nutr. 2019;58:2229-2242. doi: 10.1007/s00394-018-1773-3.

58. Casabonne D, Benavente Y, Costas L, et al. Insulin-like growth factor levels and chronic lymphocytic leukaemia: results from the MCC-Spain and EpiLymph-Spain studies. Br J Haematol. 2019;185:608-612. doi: 10.1111/bjh.15583.

59. Lorenzo I, Fernández-de-Larrea N, Michel A, et al. Helicobacter pylori seroprevalence in Spain: influence of adult and childhood sociodemographic factors. Eur J Cancer Prev. 2019;28:294-303. doi: 10.1097/CEJ.0000000000000483.

60. Peremiquel-Trillas P, Benavente Y, Martín-Bustamante M, et al. Alkylphenolic compounds and risk of breast and prostate cancer in the MCC-Spain study. Environ Int. 2019;122:389-399. doi: 10.1016/j.envint.2018.12.007.

61. Olmedo-Requena R, González-Donquiles C, et al. Agreement among Mediterranean Diet Pattern Adherence Indexes: MCC-Spain Study. Nutrients. 2019;11(3). pii: E488. doi: 10.3390/nu11030488.

62. Vitelli Storelli F, Molina AJ, Zamora-Ros R, et al. Flavonoids and the Risk of Gastric Cancer: An Exploratory Case-Control Study in the MCC-Spain Study. Nutrients. 2019;11(5). pii: E967. doi: 10.3390/nu11050967.

63. Dierssen-Sotos T, Gómez-Acebo I, Palazuelos C, et al. Fatty acid intake and breast cancer in the Spanish multicase-control study on cancer (MCC-Spain). Eur J Nutr. 2019 (In press). doi: 10.1007/s00394-019-01977-8.

64. Aragonés N, Fernández de Larrea N, Pastor-Barriuso R, et al. Epstein Barr virus antibody reactivity and gastric cancer: A population-based case-control study. Cancer Epidemiol. 2019;61:79-88. doi: 10.1016/j.canep.2019.05.008.

65. Kogevinas M, Espinosa A, Papantoniou K, et al. Prostate cancer risk decreases following cessation of night shift work. Int J Cancer. 2019 (In press). doi: 10.1002/ijc.32528.

66. Obón-Santacana M, Romaguera D, Gracia-Lavedan E, et al. Dietary Inflammatory Index, Dietary Non-Enzymatic Antioxidant Capacity, and Colorectal and Breast Cancer Risk (MCC-Spain Study). Nutrients. 2019;11. pii: E1406. doi: 10.3390/nu11061406.

67. Huerta JM, Molina AJ, Chirlaque MD, et al. Domain-specific patterns of physical activity and risk of breast cancer sub-types in the MCC-Spain study. Breast Cancer Res Treat. 2019 (In press). doi: 10.1007/s10549-019-05358-x.

68. Butt J, Fernández de Larrea N, et al. Antibody responses to flagellin C and Streptococcus gallolyticus pilus proteins in colorectal cancer. Sci Rep. 2019;9:10847. doi: 10.1038/s41598-019-47347-6.

BMJ Open

Cohort profile: The MCC-Spain follow-up on colorectal, breast and prostate cancers. Study design and initial results

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2019-031904.R2
Article Type:	Cohort profile
Date Submitted by the Author:	10-Sep-2019
Complete List of Authors:	Alonso, Jessica ; University of Cantabria – IDIVAL, Santander, Spain Molina, Antonio J; Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain. Jiménez-Moleón, Jose Juan; University of Granada, Department of Preventive Medicine and Public Health, School of Medicine; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Pérez-Gómez, Beatriz ; National Center for Epidemiology, Carlos III Institute of Health; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Martin, Vicente; 2- Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain. Moreno, Victor; Catalan Institute of Oncology (ICO-IDIBELL), Cancer Prevention and Control Program; University of Barcelona, Clinical Sciences, Faculty of Medicine Amiano, Pilar; Public Health Division of Gipuzkoa, BioDonostia Research Institute; Consortium for Biomedical Research in Epidemiology and Public Health (CIBER Epidemiología y Salud Pública, CIBERESP) Ardanaz, Eva de Sanjose, Silvia ; L'Hospitalet de Llobregat, Barcelona SALCEDO, INMACULADA; Universidad de Granada, Preventive Medicine and Public Health+34 Fernandez-Tardon, Guillermo; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Alguacil, Juan; CIBER de Epidemiología y Salud Pública (CIBERESP), Salas, Dolores; Generalitat Valenciana Conselleria de Sanitat Marcos-Gragera, Rafael; Institut Catala d' Oncologia, Epidemiology Unit and Girona Cancer Registry; Universitat de Girona, Research Group on Statistics, Econometrics and Health (GRECS) Chirlaque, Maria Dolores; Murcia Cancer Registry Aragonés, Nuria; 3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain. Castaño-Vinyals, Gemma ; Instituto de Salud Global Barcelona Pollán, Marina; Centro Nacional de Epidemiologia. Instituto de Salud Carlos III, Area de Epidemiologia Ambiental y Cancer; CIBER Epidemiologia y Salud Publica (CIBERESP), Kogevinas, Manolis; Instituto de Salud Global de Barcelona, Llorca, Javier; University of Cantabria,
Primary Subject	Epidemiology

Heading:	
Secondary Subject Heading:	Oncology
Keywords:	Cohort, Epidemiology < TROPICAL MEDICINE, colorectal cancer, breast cancer, prostate cancer, MCC-Spain

Cohort profile: The MCC-Spain follow-up on colorectal, breast and prostate cancers. Study design and initial results

Authors: Jessica Alonso-Molero*[1], Antonio J Molina*[2], Jose J Jiménez-Moleón*[3,4], Beatriz Pérez-Gómez[3,5,6], Vicente Martín[2,3], Víctor Moreno[3,7,8], Pilar Amiano[3,9], Eva Ardanaz[3,10,11], Silvia de Sanjosé[3,7], Inmaculada Salcedo[3,4], Guillermo Fernández-Tardón[3,12], Juan Alguacil[3,13], Dolores Salas[3,14], Rafael Marcos-Gragera[3,15], M Dolores Chirlaque[3,16], Nuria Aragonés[3,17], Gemma Castaño-Vinyals[3,18,19,20], Marina Pollán[3,5,6], Manolis Kogevinas[3,18,19,20], Javier Llorca[1,3].

*Equal contribution

Affiliations:

1- University of Cantabria-IDIVAL, Santander, Spain.
2- Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain.
3- CIBER Epidemiología y Salud Pública (CIBERESP),Spain.
4- Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain.
5- Cancer and Environmental Epidemiology Unit, National Center for Epidemiology, Carlos III Institute of Health, Madrid, Spain.
6- Cancer Epidemiology Research Group, Oncology and Hematology Area, IIS Puerta de Hierro (IDIPHIM), Madrid, Spain.
7- Cancer Prevention and Control Program, Catalan Institute of Oncology-IDIBELL, L'Hospitalet de Llobregat, Spain.
8- Department of Clinical Sciences, Faculty of Medicine, University of Barcelona, Barcelona, Spain.
9- Public Health Division of Gipuzkoa, BioDonostia Research Institute, San Sebastian, Spain.
10- Navarra Public Health Institute, Pamplona, Spain.
11- IdiSNA, Navarra Institute for Health Research, Pamplona, Spain.
12- Oncology Institute, University of Oviedo, Oviedo.
13- Centro de Investigación en Recursos Naturales, Salud y Medio Ambiente (RENSMA), Universidad de Huelva, 21004 Huelva, Spain.
14- Área de Cáncer y Salud Pública, FISABIO-Salud Pública, 46020 Valencia, Spain.
15- Unitat d'Epidemiologia i Registre de Càncer de Girona (UERCG), Pla Director d'Oncologia, Institut Català d'Oncologia, Institut d'Investigaciò Biomèdica de Girona (IdIBGi), Universitat de Girona, Girona, Spain.
16- Department of Epidemiology, Regional Health Authority, IMIB-Arrixaca, Murcia University, Murcia, Spain.
17- Department of Health, Epidemiology Section, Public Health Division, Madrid.
18- ISGlobal, Barcelona, Spain.

19- Universitat Pompeu Fabra (UPF), Barcelona, Spain.

20- IMIM (Hospital Del Mar Medical Research Institute), Barcelona, Spain.

Correspondence:

Jéssica Alonso-Molero

Facultad de Medicina, Universidad de Cantabria

Avda. Herrera Oria s/n

39011 Santander (Cantabria)

Spain

ORCiD: https://orcid.org/0000-0002-1939-8798

e-mail: alonsomoleroj@gmail.com

Word count: 3347

ABSTRACT

Purpose: Since 2016, the Multicase-control study in Spain (MCC-Spain) has focused towards the identification of factors associated with cancer prognosis; inception cohorts of patients with colorectal, breast and prostate cancers were assembled using the incident cases originally recruited.

Participants: 2140 new cases of colorectal cancer, 1732 of breast cancer and 1112 of prostate cancer were initially recruited in 12 Spanish provinces; all cancers were incident and pathologically confirmed. Follow-up was obtained for 2097 (98%), 1685 (97%) and 1055 (94.9%) patients, respectively.

Findings to date: Information gathered at recruitment included sociodemographic factors, medical history, lifestyle and environmental exposures. Biological samples were obtained, and 80% patients were genotyped using a commercial exome array. The follow-up was performed by: (i) reviewing medical records; (ii) interviewing by phone the patients on quality-of-life and; (iii) verifying vital status and cause of death in the Spanish National Death Index. Ninety-seven percent of recruited patients were successfully followed-up in 2017 or 2018; patient-years of follow-up were 30914. Most colorectal cancers (52%) were at clinical stage II or less at recruitment; 819 patients died in the follow-up and five-year survival was better for women (74.4%) than men (70.0%). 71% breast cancers were diagnosed at stages I or II; 206 women with breast cancer died in the follow-up and five-year survival was 90.7%. 49% prostate cancers were diagnosed at stage II and 32% at stage III; 119 patients with prostate cancer died in the follow-up and five-year survival was 93.7%.

Future plans: MCC-Spain has built three prospective cohorts on highly frequent cancers across Spain, allowing to investigate socioeconomic, clinical, lifestyle, environmental and genetic variables as putative prognosis factors determining survival of patients of the three cancers and the interrelationship of these factors.

KEYWORDS:

Cohort, epidemiology, colorectal cancer, breast cancer, prostate cancer, MCC-Spain

STRENGTHS AND LIMITATIONS OF THIS STUDY:

- 4837 incident cases of cancer (2097 colorectal; 1685 breast; 1055 prostate) have been prospectively followed-up accounting for more than 30000 patients-year, and with only 153 patients (3%) lost to follow-up.
- The cohort covers a wide spectrum of the Spanish population including 23 hospitals across Spain.
- A major strength of this study is the amount of information gathered at diagnosis, including sociodemographic, lifestyle, nutrition, familial and personal medical history, reproductive history, use of drugs, sleep, genotyping, clinical and pathological characteristics of the tumour, first-line treatment, side effects, health-related quality of life and current vital status.
- Biological samples obtained at recruitment (tumour specimen, blood or saliva, toenail, hair and urine) will allow further investigations on metabolomics, epigenetics and exposure to chemicals such as metals.

- The multicentre characteristic of the study allows the evaluation of a wide geographical basis and increases the representativity of the recruited sample, but it also may introduce heterogeneity in the information gathered and in treatment.

INTRODUCTION

Tumour size, node infiltration, metastasis, histology, clinical stage, cancer subtype continue being main prognosis factors in patients with cancer in spite of the evolving first-line treatment [1–5]. Little effort, however, has been paid to examine the impact on survival of patient factors -such as lifestyle, genetics or environmental- together with tumour features and treatment.

Large prospective cohort studies on cancer focus on identifying risk factors [6] while clinical cohorts on cancer survival usually aim to analyse survival relationships with tumour properties, first-line treatment or patient characteristics. For instance, Lagendijk et al analysed data on 129,692 women with breast cancer from the Netherlands Cancer Registry to compare breast conserving therapy and mastectomy in subgroups according age at diagnosis, stage, systemic therapy, comorbidity, oestrogen/progesterone receptors and HER2 status [7]; Cardwell et al linked the National Cancer Data Repository to the United Kingdom Clinical Practice Research Datalink and mortality data from the Office of National Statistics to investigate if statin use after colorectal cancer diagnosis was associated with better prognosis [8]; Petterson et al studied survival after prostate cancer diagnosis in 121,392 Swedish men from the Prostate Cancer data Base Sweden 3.0, where data were available on age, stage, grade, prostate-specific antigen level, model of detection, comorbidity, educational level and primary treatment [9]. It is noteworthy that these cohorts were based on cancer registries were data availability is usually restricted to demographic variables (sometimes including educational level and deprivation), tumour characteristics and few data on comorbidities or healthy habits. A different approach has been the use of the Surveillance, Epidemiology and End Results (SEER) database to retrospectively analyse survivorship with breast cancer [10], colorectal cancer [11] or prostate cancer [12], but although the number of participants could be over 100,000, available data are restricted to those recorded for the general purposes of the SEER program, not specifically for studying survivorship with cancer.

The MCC-Spain includes three prospective cohorts of cancer patients (colorectal, female breast and prostate) with the aim of to investigate long-term survival factors including cancer characteristics and treatment, but also genetics and other omics, lifestyle (physical activity, nutrition, sleep, toxic habits), occupational exposures (including night shift work), environmental factors such as living area conditions and medical history, aiming to build integrative prognosis models. This multidisciplinary study will provide a complete evaluation of the biological, clinical, environmental, lifestyle and socio-economic factors determining survival of patients of the three cancers and of the interrelationship of these factors. Specific objectives for each cohort are: For the colorectal cancer cohort: (1) To study the accomplishment of primary treatment with ESMO and ASCO guidelines and factors associated with it, (2) to study factors associated with survivorship, response to treatment and toxicity due to chemotherapy using genetic, epidemiological and clinical-pathological variables, (3) to validate those models via comparison with Glasgow Prognostic Score predictions. For the breast cancer cohort: (1) To study whether first-line treatment accomplished St Gallen International Expert Consensus recommendations, (2) to study factors associated with survivorship, response to treatment and toxicity due to chemotherapy using genetic, epidemiological and clinical-pathological variables, (3) to validate those models via comparison with the Nottingham Prognostic Index and *Adjuvant!*. For prostate cancer cohort: (1) To analyse the adequacy of initial treatment to

recommendations by the European Association of Urology and the National Institute for Health and Care Excellence, (2) to elaborate models on survivorship, risk of biochemical relapse, quality of life, response to primary treatment, toxicity to chemotherapy/brachytherapy (3) to validate survivorship and risk of biochemical relapse models via comparison with Han and Kattan nomograms. In this article, we report the study design, the main description of all three cohorts and the preliminary results on survival.

COHORT DESCRIPTION AND METHODS

The MCC-Spain began as a case-control study in 2008, started by the Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), on both genetic and environmental exposures associated with colorectal, female breast, prostate and gastric cancers and chronic lymphocytic leukaemia. Its design has been published elsewhere[13]; it recruited 10,183 incident cases and controls between 2008 and 2013 in 12 Spanish provinces (Asturias, Barcelona, Cantabria, Girona, Granada, Gipuzkoa, Huelva, León, Madrid, Murcia, Navarra, and Valencia). Using the incident cases originally recruited between 2008 and 2013, and given that in 2016 the MCC-Spain has turned towards the identification of factors associated with cancer prognosis; inception cohorts on colorectal, breast and prostate cancers has been assembled, enrolling the patients for a prospective follow-up carried out in 2017-2018.From here on, we only refer to the recruited cases of colorectal (2140 cases), breast (1738 cases) and prostate (1112 cases) cancers; their distribution by province and hospital appears in Supplementary Table 1 and the flow chart appears in Figure 1.

Patient recruitment and Public Involvement Statement

Patients recruited were between 20 and 85 years old, had resided in the catchment area for at least 6 months before the recruitment and were able to answer the epidemiological questionnaire and had incident colorectal, breast or prostate cancer. For the recruitment, study personnel contacted newly diagnosed cancer cases in the 21 collaborating hospitals. Cases were identified as soon as possible after the diagnosis; only histologically confirmed incident cases were included.

Participants are being informed on the project's main results via flyers. There is no other patient's involvement.

Information at recruitment and biological samples

The information obtained and its timing is summarized in Table 1.

Table 1. Information obtained in the MCC-Spain

Phase	Measurements
Phase I: Recruitment 2008-2013	**Contact with newly diagnosed cancer cases.** **Trained personnel perform a structured computerized epidemiological questionnaire in a face-to-face interview to obtain the follow information:** Sociodemographic; Personal and familial medical history; Use of drugs; Reproductive history; Physical activity; Environmental and occupational exposures **A validated semi-quantitative frequency-food questionnaire is self-completed to obtain diet information.** **Biological samples are obtained:** Peripheral blood or saliva; Toenail; Hair; Urine; Tumour biopsies **A genotype of exome is made using the Illumina® Infinium HumanExome.** **Medical Records review by trained personnel to obtain:** Pathology characteristics; Tumour extension; Clinical data; First-line treatment; Recurrence For colorectal cancer cases: First biopsy; Surgical piece dimensions; Histological type; Carcinoembryonic antigen levels For breast cancers cases: Differentiation's degree; Immuno-histochemical characteristics For prostate cancer cases: Gleason score; D'Amico classification; PSA levels
Phase II: Follow-up 2017-2018	**Medical Records review by trained personnel to obtain:** For colorectal cancer cases: TNM status at recruitment; First-line treatment; Surgical margins; Patient status after first-line treatment; Appearance of second primary tumour; Current patient's vital status For breast cancers cases: Histological grade at diagnosis; Nottingham index; Complete clinical/pathological remission; Grade of response to treatment; Relapse; Second primary tumour; Current patient's vital status For prostate cancer cases: PSA concentration; Gleason grade and biopsy characteristics at diagnosis; Pathological characteristics of the surgical specimen; First-line treatment; Clinical response to first-line treatment; Second primary tumour; Current patient's vital status **Consult in the IND to realize the vital status of patients.** **Contact by phone to complete specific quality of life questionnaires.** For colorectal cancer cases: SF-12; FACT-Colorectal Symptom Index For breast cancer cases: SF-12; FACT/NCCN Breast Symptom Index For prostate cancer cases: SF-12; Charlson Comorbidity Index; FACT-P questionnaire; International Prostate Symptom Score (I-PSS).

Information about sociodemographic, personal and familial medical history, use of drugs, reproductive history, physical activity, environmental and occupational exposures was gathered using a standardized questionnaire [14] administered by trained personnel in a face-to-face interview. Diet information in the year before diagnosis was obtained using a validated semi-quantitative frequency-food questionnaire [15] filled by the participants. Both questionnaires can be found in http://www.mccspain.org. Biological samples were obtained, including peripheral blood or saliva (from 92% breast cancer cases, 95% colorectal cancer cases and 97% prostate cancer cases), toenail, hair (from 77% and 81% participants, respectively), urine or tumour biopsies. Regarding peripheral blood, 27ml were aliquoted in whole blood, plasma, serum and cellular fraction for DNA extraction and stored at -80ºC. Saliva was collected from people unable to donate a blood sample.

Genotyping

From 80% participants, a genotype of exome was made using the Illumina® Infinium HumanExome. In addition to the about 250,000 exome variants included in the original beadcheap, 6000 SNPs previously found in GWAS or localized in metabolic pathways of interest were added upon MCC-Spain researchers' request. MCC-Spain has recently obtained funding for carrying out a GWAS with all the participants and to launch an analysis on circulant miRNA in breast cancer patients.

Initial clinical information

Trained personnel reviewed the medical records in order to collect information on pathology characteristics, tumour extension, clinical data, first-line treatment and recurrence. For colorectal cancer cases, we documented the first biopsy, the tumour location, surgical piece dimensions, histological type according to the ICD-O-3 version, TNM status, carcinoembryonic antigen levels and first-line treatment (surgery extension -if done-; neoadjuvant, adjuvant or palliative chemotherapy or radiotherapy). For breast cancers, we obtained information on tumour location, differentiation's degree, immuno-histochemical characteristics (hormonal receptors, Erb-B2), TNM status and first-line treatment (mastectomy / conservative surgery; neoadjuvant, adjuvant or palliative hormonotherapy, chemotherapy or radiotherapy; target-directed therapy such as transtuzumab). For prostate cancer cases, we gathered information on tumour location, Gleason score, D'Amico classification, TNM status, PSA levels and first-line treatment (none, surgery, hormonotherapy, chemotherapy or radiotherapy; including, when appropriate, the therapy purpose -neoadjuvant, adjuvant or palliative.) TNM status for all three tumours was classified according to the TNM-6th edition.

Follow-up information

Follow-up was carried-out between 2017 and 2018 by reviewing medical records. For colorectal cancer patients, we collected data on TNM status at recruitment, first-line treatment, surgical margins, patient status after first-line treatment (free of disease, partial response, progression, relapse or stable disease), appearance of second primary tumour, and current patient's vital status. For breast cancer patients, we gathered information on histological grade at diagnosis, Nottingham index, complete clinical/pathological remission, grade of response to treatment (according to the Miller and Payne system or similar classifications), relapse, second primary tumour, and current patient's vital status. For prostate cancer

patients, the information assembled included PSA concentration, Gleason grade and biopsy characteristics at diagnosis; pathological characteristics of the surgical specimen, first-line treatment, clinical response to first-line treatment (stable disease / progression or relapse / unknown), chemical relapses, relapse clinical characteristics (local / metastatic and its location), second primary tumour, and current patient's vital status. Some of these data were obtained in order to double check the clinical information collected at recruitment.

The National Death Index (Índice Nacional de Defunciones -IND-) was consulted to realize the vital status of patients whose last contact with the hospital had occurred 3 or more months before our revision of his/her medical record. The IND is a nation-wide data-base supported by the Spanish Ministry of Health; it is intended to allow the researchers to establish the vital status of patients under study[16]

Patients alive at the follow-up were contacted by phone and asked to complete specific quality of life questionnaires: SF-12[17] (colorectal, breast and prostate cancers), FACT-Colorectal Symptom Index (FCSI)[18] (colorectal cancer), FACT/NCCN Breast Symptom Index [19] (breast cancer) and -for prostate cancer- the Charlson Comorbidity Index [20], the FACT-P questionnaire [21] and the International Prostate Symptom Score (I-PSS)[22].

The number of patients with follow-up is 2097 for colorectal, 1685 for breast, and 1055 for prostate cancer cohorts. This gives a 91% statistical power for colorectal cancer to detect hazard ratio ≥1.2; an 83% statistical power for breast cancer to detect the same hazard ratio; and an 80% statistical power for prostate cancer to detect hazard ratio ≥ 1.25 (assuming 20% exposed patients and 75, 90 and, 85% survival probability in the non-exposed group, respectively).

Statistical analysis

For preliminary results shown in this paper, data are described using absolute frequencies with percentages and means with standard deviations. Patients died by any cause before the end of follow-up were classified as events and censored otherwise. Time of follow-up was the difference between date of diagnosis and date of death or date of last contact with the hospital or the researchers. Survival probabilities were obtained using unadjusted Kaplan-Meier estimators. Further analyses should deal with confounding and modifiers using multivariate regression models (e.g.: Cox or Weibull regression). Initial treatment could be related with both basal factors and survivorship, eventually leading to confounding by indication; it would be controlled using propensity scores.

Ethics

The protocol of MCC-Spain was approved by the Ethics committees of the participating institutions [13] At recruitment, all participants were informed about the study objectives and signed an informed consent, which also included the authorization for following-up the patient via medical records or phone calls; only participants agreeing in being followed-up were included in the inception cohorts. Confidentiality of data is secured by removing personal identifiers in the datasets. The database was registered in the Spanish Agency for Data Protection, number 2102672171.

FINDINGS TO DATE

The MCC-Spain has provided results on the effects of different risk factors. For instance, night shift work increased the risk of more aggressive prostate cancers [23], although this excess risk almost disappeared 20 years after last exposure [24] o; long-term consumption of calcium channel blockers was associated with higher breast cancer risk in overweight women [25]; adherence to the Western dietary patterns increased breast cancer risk in both pre- and post-menopausal women [26]; first validation in a European population of a risk model for breast cancer developed in American women using both modifiable and non-modifiable risk factors as well as 92 genetic variants [27]; use of environmental and genetic factors to elaborate a model to stratify the risk of colorectal cancer [28]; adherence to the World Cancer Research Fund/American Institute for Cancer Research nutrition-based guidelines was associated with lower risk of colorectal and breast cancers, but not of prostate cancer [29]. A complete list of published results from MCC-Spain appears in Supplementary Table 2 and Supplementary reference list.

Initial results of the follow-up are showed in this work. Table 2 displays the main characteristics of the patients; Table 3 details specific information of each tumour; Table 4 describes first-line treatment.

Table 2. Main characteristics of the followed patients

Variable	Category	Colorectal cancer (n = 2097)	Breast cancer (n = 1685)	Prostate cancer (n =1055)
Age (mean±sd)		66.98 (±10.85)	56.5 (±12.6)	65.86 (±7.38)
Gender	**Women**	763 (36.39%)	1685 (100%)	-
	Men	1334 (63.61%)	-	1055 (100%)
Postmenopausal	**Yes**	-	1095 (65.0%)	-
	No	-	589 (35.0%)	-
	Missing	-	1 (0.1%)	-
Histology (specific types in each tumour)		Adenocarcinoma 1882 (89.75%)	Ductal 1276 (75.7%)	Adenocarcinoma (acinar) 1053 (99.91%)
		Mucinous adenocarcinoma 125 (5.96%)	Lobular 110 (6.5%)	Others 2 (0.09%)
		Signet ring cells adenocarcinoma 12 (0.57%)	Paget disease 19 (1.1%)	-
		Others 4 (0.19%)	Others 280 (16.6%)	-
		Unknow 74 (3.53%)	-	-
Tumour size	**T0**	98 (4.67%)	23 (1.4%)	-
	T1	125 (5.96%)	861 (51.1%)	227 (21.52%)
	T2	283 (13.49%)	424 (25.2%)	521 (49.38%)
	T3	1172 (55.89%)	73 (4.3%)	98 (9.29%)
	T4	319 (15.21%)	39 (2.3%)	8 (0.76%)
	Tis	-	109 (6.5%)	-
	Missing	100 (4.77%)	156 (9.3%)	196 (18.58%)
	Not evaluable	-	-	5 (0.47%)
Node infiltration	**N0**	1193 (56.89%)	877 (52.0%)	271 (25.69%)
	N1	515 (24.56%)	441 (26.2%)	9 (0.85%)
	N2	286 (13.64%)	186 (11.0%)	-
	N3	-	5 (0.3%)	-
	Missing	103 (4.91%)	176 (10.4%)	224 (21.23%)
	Not evaluable	-	-	551 (52.23%)
Metastasis	**No**	1721 (82.07%)	1376 (81.7%)	532 (50.43%)
	Yes	330 (15.74%)	41 (2.4%)	17 (1.61%)

	Missing	46 (2.19%)	268 (15.9%)	215 (20.38%)
	Not evaluable	-	-	291 (27.58%)
Clinical stage	**0**	77 (3.67%)	-	-
	I	338 (16.12%)	702 (41.7%)	367 (34.79%)
	II	673 (32.09%)	479 (28.4%)	496 (47.01%)
	III	569 (27.13%)	179 (10.6%)	132 (12.51%)
	IV	330 (15.74%)	41 (2.4%)	17 (1.61%)
	Missing	110 (5.25%)	284 (16.9%)	43 (4.08%)

Table 3. Specific information for each cancer

Specific information for colorectal cancer			Specific information for breast cancer			Specific information for prostate cancer		
			Oestrogen receptor	**Positive**	1398 (83.0%)		**1 (Gleason score = 6)**	449 (42.56%)
				Negative	244 (14.5%)			
				Missing	43 (2.6%)		**2 (Gleason score = 3+4)**	299 (28.34%)
Location	**Right colon**	566 (26.99%)	**Progesterone receptor**	**Positive**	1237 (73.4%)			
				Negative	401 (23.8%)		**3 (Gleason score =4+3)**	120 (11.37%)
	Left colon	719 (34.29%)		**Missing**	47 (2.8%)	**Gleason grade**		
	Rectum-sigma	791 (37.72%)	**Her2**	**Positive**	294 (17.4%)		**4 (Gleason score = 8)**	83 (7.87%)
				Negative	1250 (74.2%)			
	Unknown	21 (1%)		**Missing**	141 (8.4%)		**5 (Gleason score 9 or 10)**	65 (6.16%)
			Intrinsic subtype	**Luminal A**	997 (59.2%)			
				Luminal B	331 (19.6%)			
				Her2	81 (4.8%)		**Missing**	39 (3.70%)
				Basal-like	130 (7.7%)			
Differentiation's degree	**I**	520 (24.8%)		**Luminal ONI***	91 (5.4%)	**PSA** (ng/ml)**		11.51 (±16.28)
	II	1100 (52.46%)		**Non-luminal ONI***	13 (0.8%)			
	III	247 (11.78%)		**Missing**	42 (2.5%)	**D'Amico**	**Low risk**	325 (30.81%)
	Not evaluable	230 (10.97%)	**Grade**	**I**	329 (19.5%)		**Intermediate risk**	425 (40.28%)
				II	520 (30.9%)			
				III	355 (21.1%)		**High risk**	284 (26.92%)
				Missing	481 (28.5%)		**Missing**	21 (1.99%)

*ONI: Otherwise Not Identified. **PSA: Prostate Specific Antigen

Table 4. First-line treatment

Treatment	Category	Colorectal cancer	Breast cancer	Prostate cancer
None (active surveillance)		-	-	38 (3.6%)
Surgery		Total: 1999 (95.3%)	Conservative: 1231(73.1%)	Prostatectomy: 639 (61.4%)
		Resection: 1800 (85.8%)		
		Palliative: 127 (6.1%)	Mastectomy: 454 (26.9%)	
		No resection: 61 (2.9%)		
		Others: 11 (0.5%)		
Chemotherapy	**Neoadjuvant**	427 (20.4%)	200 (11.9%)	1 (0.1%)
	Adjuvant	1024 (48.8%)	664 (39.4%)	1 (0.1%)
	Palliative	67 (3.2%)	25 (1.5%)	7 (0.7%)
Radiotherapy	**Neoadjuvant**	401 (19.1%)	5 (0.3%)	227 (21.5%)
	Adjuvant	82 (3.9%)	1132 (67.2%)	36 (3.4%)
	Palliative	5(0.2%)	21 (1.2%)	2 (0.2%)
Endocrine therapy	**Yes**	-	1023 (60.7%)	Adjuvant to surgery: 19 (1.8%)
				Adjuvant to radiotherapy: 99 (9.4%)
				Neoadjuvant: 102 (9.7%)
				Palliative: 69 (6.5%)
	No	-	662 (39.3%)	689 (65.3%)
Others (specify for each tumour)	**Endoscopy**	Complete resection: 107 (5.1%)	-	-
		Non-complete resection: 62 (3.0%)		
	Her2-targeted therapy	-	152(9.0%)	-
	Cryotherapy	-	-	21 (2.0%)
	Transurethral resection	-	-	4 (0.4%)

Colorectal Cancer

Out of 2140 patients with colorectal cancer, 2097 (98%) have been followed. They were 67±10.9 years-old on average at recruitment; 1334 (63.4%) were men. The first case was recruited on 18th of March 2007 and the follow-up was closed on 23rd of August 2018, accounting for 12813.8 person-years of follow-up. 819 (39.1%) cases died in this period; linearized mortality rate was 6.4 per 100 patient-years (95% CI: 6.0 – 6.8) (Table 2).

Most cases (1882, 90%) were adenocarcinoma; the most frequent location was rectum-sigma (37.7%) and the less frequent right colon (27%). 52% patients were at clinical stage II or less; in 110 patients (5.3%) we could not establish the clinical stage. 52.5% cancers were moderately differentiated (grade II) and 24.8% well differentiated (grade I) (Table 3).

Surgery was carried out in 1999 colorectal cancer patients; it was for palliative purposes in 127 patients (6.1%). 169 patients were treated via endoscopy, reaching complete resection in 107 of them. 1518 (72.4%) patients received chemotherapy; most of them (1451) for adjuvant or neoadjuvant purposes; 488 (23.2%) received radiotherapy (401 neoadjuvant, 82 adjuvant and only 5 palliative) (Table 4).

Five-year survival probability estimated via Kaplan-Meier was 71.6% (95% CI: 69.6 – 73.5) (Figure 2a). Survival was higher in women (74.4%, 95% CI: 71.0 – 77.2) than in men (70.0%, 95% CI: 67.5 – 72.4) ($p<0.001$) (Figure 2b). Five-year survival probability was 85.2% (81.0 – 88.6) in patients diagnosed in stage I, 84.0% (81.0 – 86.6) in stage II, 73.4% (69.6 – 76.9) in stage III and 27.6% (22.9 – 32.5) in stage IV (Figure 3a).

Breast Cancer

The maximum span for breast cancer follow-up was nine and a half years (from 13th July 2007 to 22nd March 2017). Follow-up was obtained for 1685 out of 1738 breast cancer patients (97%), adding 10931 person-years; 206 patients died in the follow-up; the linearized mortality rate was 1.9 per 100 patient-years (95% CI: 1.6 – 2.2).

Women with breast cancer were 56.5±12.6 year-old on average at recruitment; 65% were postmenopausal. The most usual type of tumour was ductal (75.7%), followed by lobular (6.5%). Most breast cancers were diagnosed at early stages (71% at stages I or II) and only 41 (2.4%) had metastasized at the time of diagnosis (Table 2). 83% cancers were oestrogen receptor positive, 73.4% progesterone receptor positive and 17.4% Her2 positive. Regarding intrinsic subtypes, 997 (59.2%) could be classified as luminal A, 331 (19.6%) as luminal B, 81 (4.8%) as Her2 and 130 (7.7%) as basal-like. According to grade of differentiation, moderately differentiated accounted for 30.9% breast cancers; well differentiated and bad differentiated accounted for about 20% cancers each. Grade could not be obtained from medical records in 481 patients (28.5%) (Table 3).

Conservative surgery was performed in 1231 (73.1%) patients and mastectomy in the remaining 454 (26.9%). Adjuvant or neoadjuvant chemotherapy was administered to 50.3% patients, while radiotherapy was used in 1158 women (68.7%), endocrine therapy was used in 1023 women (60.7%) and Her2-targeted therapy in 152 patients (9.0%) (Table 4). Kaplan-Meier 5-year survival with breast cancer was 90.7% (95%

CI: 89.2 – 92.0) (Figure 2c). Women diagnosed in stage I had 97% (95.5 – 98.1) 5-year survival probability, 91.9% (89.1 – 94.1) in stage II, 84.1% (77.8 – 88.7) in stage III and 38.5% (18.6 – 58.2) in stage IV (Figure 3b).

Prostate Cancer

A total of 1112 men with prostate cancer were recruited and 1055 (94.9%) have been followed-up; the first patient was included on 26th January, 2008 and the end of follow-up was on 13th July, 2018, adding 7169.6 person-year of follow-up. Patients were 65.9 years-old on average at recruitment. 119 patients died in the follow-up, making the linearized mortality rate 1.7 per 100 patient-years (95% CI: 1.4 – 2.0).

Almost all prostate cancers (99.9%) were adenocarcinoma; 496 (47%) were diagnosed at stage II and 132 (12.5%) at stage III (Table 2). The level of PSA gives an average of 11.5±16.3 ng/ml. Considering the Gleason score, 42.6% prostate cancers were well differentiated (Gleason grade = 1, i.e. Gleason score = 6); 28.3% were at Gleason grade 2 (Gleason score = 3+4), and only 14.0% were bad differentiated (Gleason grade 4 or 5; Gleason score ≥8); Gleason grade could not be established in 17.4% patients. D'Amico classification system results in 31.4% patients with low-risk cancer, 41.1% intermediate and 27.4% high-risk cancer (Table 3).

Thirty-eight prostate cancer patients were not treated medically at the beginning, being followed by active surveillance; prostatectomy was performed in 61.4% cases; radiotherapy in 265 patients (25.1%) and endocrine therapy in 289 patients (27.4%). A small number of patients were treated via trans-urethral resection, cryotherapy or chemotherapy (Table 4). Five-year survival probability by Kaplan-Meier was 93.7% (95% CI: 92.0 – 95.1) (Figure 2d). Survival probability 5 years after being diagnosed was 94.5% (88.1 – 97.5) for patients in stage I, 95.6% (93.3 – 97.2) in stage II, 92.4% (88.5 – 95.0) in stage III and 70.5 (42.8 – 88.6) in stage IV (Figure 3c).

STRENGTHS AND LIMITATIONS

In this article, we have described how three prospective cohorts on colorectal, breast and prostate cancers have been assembled from patients originally recruited for a case-control study, which makes 97% patients followed-up and accounts for more than 30,000 person-years. This is a main achievement of a network settled within the CIBERESP in 12 Spanish provinces. The study is population based and included only incident cancers; the amount of detailed information recorded as well as the availability of biological samples at recruitment will allow the identification of genetics, environmental, lifestyle and clinical prognosis factors in three frequent cancers in Spain. In this regard, a remarkable feature of the study is the feasibility of studying cancer risk factors as putative prognosis factors; for example, risk factors already analysed in the case-control phase have been diet, circadian cycle disruption, some drugs, endocrine disruptors, artificial light or proximity to green spaces; information regarding these risk factors was recorded at recruitment and is available for a prognosis factor analysis in the follow-up (see Supplementary material for a complete reference list of MCC-Spain articles).

Obtaining information on personal history, occupational exposures, diet, physical exercise or other lifestyle components is somewhat subjective as both patients and interviewers could be prone to be influenced by

their feelings or beliefs about the hypotheses under study, eventually leading to differential misclassification bias. This could hardly have occurred in this study: Firstly, patients were not aware of the hypotheses. Secondly, interviewers were familiar with the case-control study, not with the cohort design as it was decided later; therefore, if interviewers or patients have introduced some misclassification, it could probably have been non-differential, eventually leading to bias towards the null [30], which would make more robust the positive findings in this cohort study.

This study has also some weaknesses: Firstly, multicentre studies are double edged; they are needed in order to include many patients, but they could introduce heterogeneity in both the information gathered and the way patients are treated. In this regard, the analysis of prognosis factors should be adjusted for the hospital of recruitment. Secondly, 113 participating patients have been lost (43 with colorectal cancer, 53 with breast cancer and 57 with prostate cancer); we have tried to minimise it by searching information in three ways: medical records, phone calls and IND; however, we cannot rule out that some patients without follow-up could have died. It is noteworthy that -due to the small number of patients without follow-up- the maximum bias it could introduce in our survival estimates is 2% for colorectal cancer, 3% for breast cancer and 5% for prostate cancer. Thirdly, we have not obtained information on lifestyle changes after diagnosis, which limits lifestyle analysis to habits before cancer appearance. Fourthly, the number of patients included in our cohorts is small compared with those based on cancer registries, limiting the analysis of subgroups.

Summarizing, the MCC-Spain study has assembled three cohorts with about 4,700 cancer patients accounting for 30,000 patient-years of follow-up, with only 3% patient withdrawals. The information gathered at recruitment will allow to prospectively investigate clinical, lifestyle, environmental and genetic variables as prognosis factors in colorectal, breast and prostate cancers in Spain.

COLLABORATION

MCC-Spain already participates in international consortiums as Genetics and Epidemiology of Colorectal Cancer Consortium (GECCO; https://www.fredhutch.org/en/labs/phs/projects/cancer-prevention/projects/gecco.html), Breast Cancer Association Consortium (BCAC; http://bcac.ccge.medschl.cam.ac.uk/) and Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL; http://practical.icr.ac.uk/blog/), where MCC-Spain would contribute to study interactions among the putative prognosis factors in vast population samples.

DATA STATEMENT

Permission to use the study database (individual-level deidentified patient data) will be granted to researchers outside the study group, after revision and approval of each request by the Steering Committee. Any kind of collaboration are encouraged.

ACKNOWLEDGEMENTS

We thank all the subjects who participated in the study and all MCC-Spain collaborators.

INFORMED CONSENT:

Informed consent was obtained from all individual participants included in the study.

AUTHOR STATEMENT:

Jessica Alonso-Molero, Antonio J Molina, Jose J Jiménez-Moleón have contributed to the conception and design of the study, considering the same contribution. JAM, AJM, JJJM, BPG, VM, VM, PA, EA, SS, IS, GFT, JA, DS, RMG, MDC, NA, GCV, MP, MK, JL have acquired data and have been involved in drafting the manuscript. All of them read and approved the final manuscript.

FUNDING:

The study was partially funded by the "Accion Transversal del Cancer", approved on the Spanish Ministry Council on the 11th October 2007, by the Instituto de Salud Carlos III-FEDER (PI08/1770, PI08/0533, PI08/1359, PS09/00773-Cantabria, PS09/01286-León, PS09/01903-Valencia, PS09/02078-Huelva, PS09/01662-Granada, PI11/01403, PI11/01889-FEDER, PI11/00226, PI11/01810, PI11/02213, PI12/00488, PI12/00265, PI12/01270, PI12/00715, PI12/00150, PI14/01219, PI14/0613, PI15/00069, PI15/00914, PI15/01032, PI17CIII/00034, PI18/00181), by the Fundación Marqués de Valdecilla (API 10/09), by the ICGC International Cancer Genome Consortium CLL (The ICGC CLL-Genome Project is funded by Spanish Ministerio de Economía y Competitividad (MINECO) through the Instituto de Salud Carlos III (ISCIII) and Red Temática de Investigación del Cáncer (RTICC) del ISCIII (RD12/0036/0036)), by the Junta de Castilla y León (LE22A10-2), by the Consejería de Salud of the Junta de Andalucía (PI-0571-2009, PI-0306-2011, salud201200057018tra), by the Conselleria de Sanitat of the Generalitat Valenciana (AP_061/10), by the Recercaixa (2010ACUP 00310), by the Regional Government of the Basque Country, by the Consejería de Sanidad de la Región de Murcia, by the European Commission grants FOOD-CT-2006-036224-HIWATE, by the Spanish Association Against Cancer (AECC) Scientific

Foundation (GCTRA18022MORE), by the Catalan Government- Agency for Management of University and Research Grants (AGAUR) grants 2017SGR723 and 2014SGR850, by the Fundación Caja de Ahorros de Asturias and by the University of Oviedo. ISGlobal is a member of the CERCA Programme, Generalitat de Catalunya.

SAMPLES:

Biological samples were stored at the biobanks supported by Instituto de Salud Carlos III- FEDER: Parc de Salut MAR Biobank (MARBiobanc) (RD09/0076/00036), “Biobanco La Fe” (RD 09 0076/00021) and FISABIO Biobank (RD09 0076/00058). Also at the Public Health Laboratory from Gipuzkoa, the Basque Biobank, the ICOBIOBANC (sponsored by the Catalan Institute of Oncology), the IUOPA Biobank from the University of Oviedo and the ISCIII Biobank.

GENOTYPING:

SNP genotyping services were provided by the Spanish "Centro Nacional de Genotipado” (CEGEN-ISCIII)".

CONFLICT OF INTEREST:

The authors declare that they have no conflict of interest

REFERENCES

1 Saadatmand S, Bretveld R, Siesling S, *et al.* Influence of tumour stage at breast cancer detection on survival in modern times: Population based study in 173 797 patients. *BMJ* 2015;**351**:h4901. doi:10.1136/bmj.h4901

2 Shukla N, Hagenbuchner M, Win KT, *et al.* Breast cancer data analysis for survivability studies and prediction. *Comput Methods Programs Biomed* 2018;**155**:199–208. doi:10.1016/j.cmpb.2017.12.011

3 Mirza AN, Mirza NQ, Vlastos G, *et al.* Prognostic Factors in Node-Negative Breast Cancer. *Ann Surg* 2002;**235**:10–26. doi:10.1097/00000658-200201000-00003

4 Zhang Z yu, Luo Q feng, Yin X wei, *et al.* Nomograms to predict survival after colorectal cancer resection without preoperative therapy. *BMC Cancer* 2016;**16**:1–21. doi:10.1186/s12885-016-2684-4

5 Merriel SWD, May MT, Martin RM. Predicting prostate cancer progression: Protocol for a retrospective cohort study to identify prognostic factors for prostate cancer outcomes using routine primary care data. *BMJ Open* 2018;**8**:1–5. doi:10.1109/TVLSI.2018.2801302

6 Riboli E, Hunt K, Slimani N, *et al.* European Prospective Investigation into Cancer and Nutrition (EPIC): study populations and data collection. *Public Health Nutr* 2002;**5**:1113–24. doi:10.1079/phn2002394

7 Lagendijk M, van Maaren MC, Saadatmand S, *et al.* Breast conserving therapy and mastectomy revisited: Breast cancer-specific survival and the influence of prognostic factors in 129,692 patients. *Int J Cancer* 2018;**142**:165–75. doi:10.1002/ijc.31034

8 Cardwell CR, Hicks BM, Hughes C, *et al.* Statin Use after colorectal cancer diagnosis and survival: A population-based cohort study. *J Clin Oncol* 2014;**32**:3177–83. doi:10.1200/JCO.2013.54.4569

9 Pettersson A, Robinson D, Garmo H, *et al.* Age at diagnosis and prostate cancer treatment and prognosis: A population-based cohort study. *Ann Oncol* 2018;**29**:377–85. doi:10.1093/annonc/mdx742

10 Leone JP, Leone J, Zwenger AO, *et al.* Prognostic Significance of Tumor Subtypes in Women With Breast Cancer According to Stage: A Population-based Study. *Am J Clin Oncol* 2019;**42**:588–95. doi:10.1097/COC.0000000000000563

11 Li Y, Feng Y, Dai W, *et al.* Prognostic Effect of Tumor Sidedness in Colorectal Cancer: A SEER-Based Analysis. *Clin Colorectal Cancer* 2019;**18**:e104–16. doi:10.1016/j.clcc.2018.10.005

12 Roy S, Morgan SC. Who Dies From Prostate Cancer? An Analysis of the Surveillance, Epidemiology and End Results Database. *Clin Oncol* 2019;**31**:630–6. doi:10.1016/j.clon.2019.04.012

13 Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, *et al.* Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. *GacSanit* 2015;**29**:308–15. doi:10.1016/j.gaceta.2014.12.003

14 Estudio MCC-Spain. Epidemiological Questionnaire. 2010;:1–52.http://www.mccspain.org/wp-content/uploads/2016/07/Quest_MCCSpain.pdf (accessed 6 Sep 2019).

15 Estudio MCC-Spain. Semi-quantitative frequency-food questionnaire. 2010;:1–36.http://www.mccspain.org/wp-content/uploads/2016/04/03_Cuestionario-alimentario_09Nov09.pdf (accessed 6 Sep 2019).

16 Navarro C. El Índice Nacional de Defunciones: Un avance en la accesibilidad de los datos de mortalidad largamente esperado. *Gac Sanit* 2006;**20**:421–3. doi:10.1157/13096513

17 Ware J, Kosinki M, Turner-Bowker D, *et al.* How to score Version 2 of the SF-12 Health survey. *Lincoln, RJ Qual Inc* 2004.

18 Colwell HH, Mathias SD, Solutions HO, *et al.* Psychometric Evaluation of the FACT Colorectal

Cancer Symptom Index (FCSI-9): Reliability, Validity, Responsiveness, and Clinical Meaningfulness HILARY. *Oncologist* 2010;**15**:308–16. doi:10.1634/theoncologist.2009-0034

19 Garcia SF, Rosenbloom SK, Beaumont JL, *et al.* Priority symptoms in advanced breast cancer: Development and initial validation of the national comprehensive cancer Network-Functional Assessment of Cancer Therapy-Breast Cancer Symptom Index (NFBSI-16). *Value Heal* 2012;**15**:183–90. doi:10.1016/j.jval.2011.08.1739

20 Charlson M, Pompei P, Ales K, *et al.* A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 1987;**40**:373–83.

21 Esper P, Mo F, Chodak G, *et al.* Measuring quality of life in men with prostate cancer using the functionale assessment of cancer therapy-prostate instrument. *Adult Urol* 1997;**30**:920–8.

22 Barry MJ, Fowler FJ, O'Leary MP, *et al.* The American Urological Association symptom index for benign prostatic hyperplasia. The Measurement Committee of the American Urological Association. *J Urol* 1992;**148**:1549–57.

23 Papantoniou K, Castaño-Vinyals G, Espinosa A, *et al.* Night shift work, chronotype and prostate cancer risk in the MCC-Spain case-control study. *Int J Cancer* 2015;**137**:1147–57. doi:10.1002/ijc.29400

24 Kogevinas M, Espinosa A, Papantoniou K, *et al.* Prostate cancer risk decreases following cessation of night shift work. *Int J cancer* Published Online First: 24 June 2019. doi:10.1002/ijc.32528

25 Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, *et al.* The use of antihypertensive medication and the risk of breast cancer in a case-control study in a Spanish population: The MCC-Spain study. *PLoS One* 2016;**11**:1–14. doi:10.1371/journal.pone.0159672

26 Castelló A, Boldo E, Pérez-Gómez B, *et al.* Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. *Maturitas* 2017;**103**:8–15. doi:10.1016/j.maturitas.2017.06.020

27 Dierssen-sotos T, Gómez-acebo I, Palazuelos C, *et al.* Validating a breast cancer score in Spanish women . The MCC-Spain study. 2018;:1–8. doi:10.1038/s41598-018-20832-0

28 Ibáñez-Sanz G, Díez-Villanueva A, Alonso MH, *et al.* Risk Model for Colorectal Cancer in Spanish Population Using Environmental and Genetic Factors: Results from the MCC-Spain study. *Sci Rep* 2017;**7**:43263. doi:10.1038/srep43263

29 Romaguera D, Gracia-Lavedan E, Molinuevo A, *et al.* Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case–control study. *Int J Cancer* 2017;**141**:83–93. doi:10.1002/ijc.30722

30 Hill HA, Kleinbaum DG. Encyclopedia of epidemiologic methods. In: Gail MH, Benichou J, eds. *Encyclopedia of epidemiologic methods*. Wiley 2000. 92–3.

FIGURE LEGENDS

Figure 1. Flow chart of the participants in the MCC-Spain study

Figure 2. Kaplan-Meier survival estimates for colorectal cancer (2a), colorectal cancer by sex (2b), breast cancer (2c) and prostate cancer (2d)

Figure 3. Kaplan-Meier estimates by stage at diagnosis for colorectal cancer (3a), breast cancer (3b) and prostate cancer (3c)

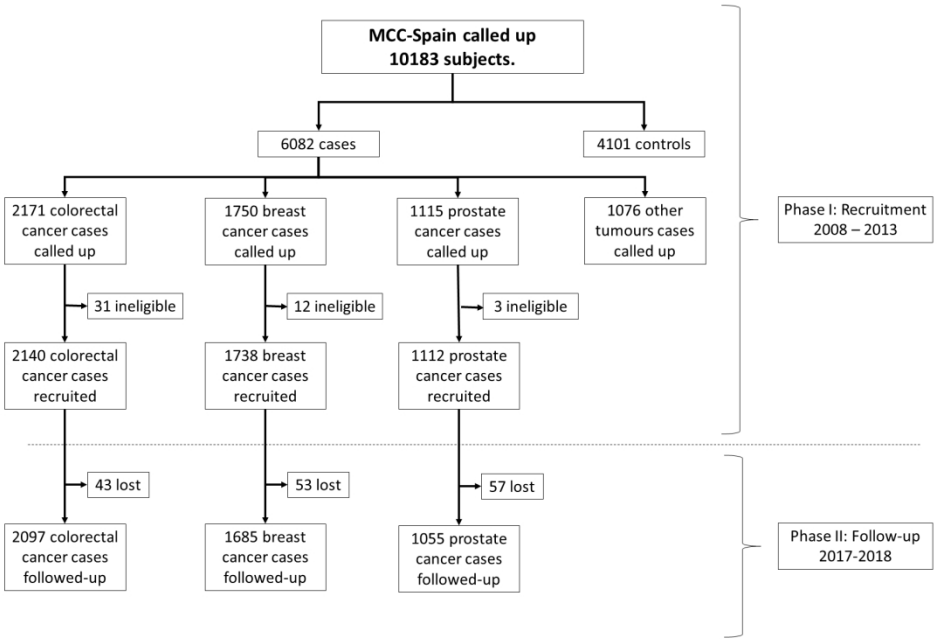

Figure 1. Flow chart of the participants in the MCC-Spain study

500x338mm (96 x 96 DPI)

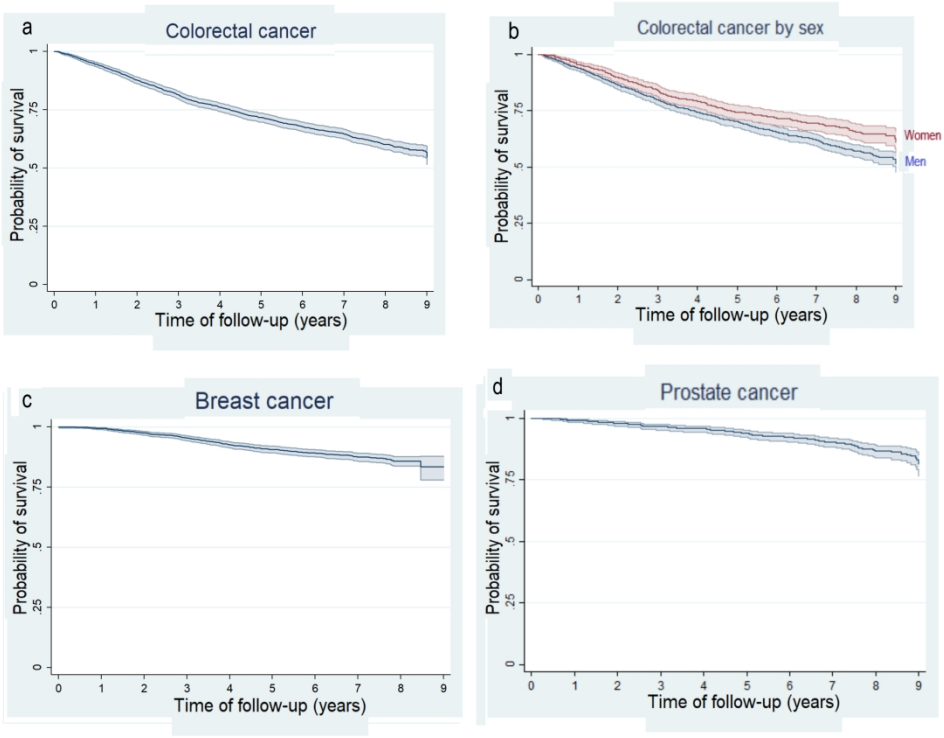

Figure 2. Kaplan-Meier survival estimates for colorectal cancer (2a), colorectal cancer by sex (2b), breast cancer (2c) and prostate cancer (2d)

727x602mm (96 x 96 DPI)

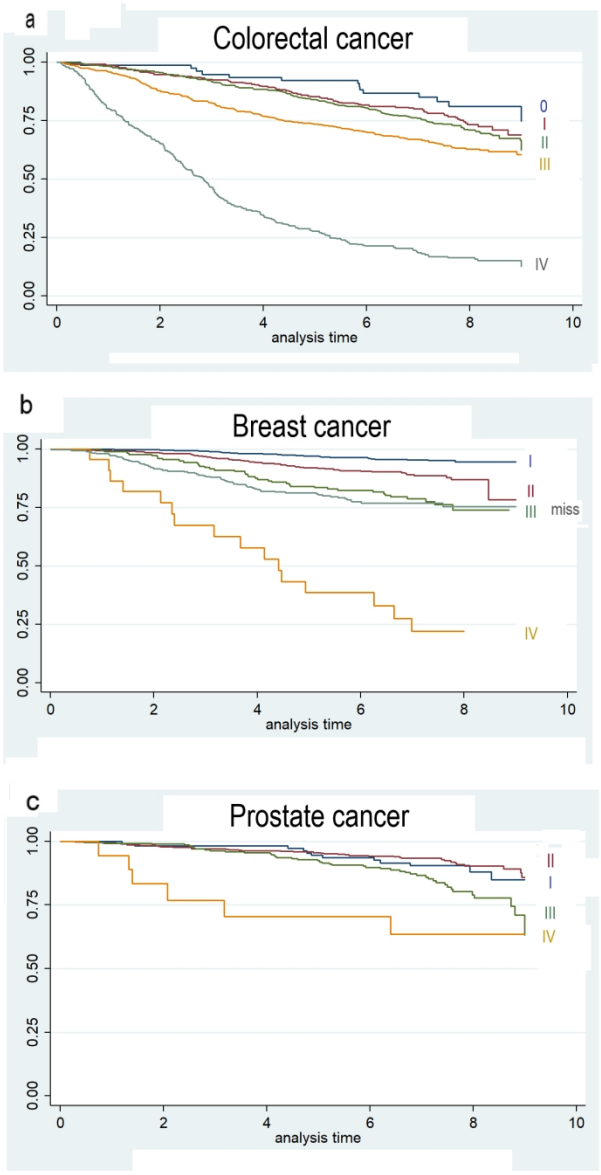

Figure 3. Kaplan-Meier estimates by stage at diagnosis for colorectal cancer (3a), breast cancer (3b) and prostate cancer (3c)

Supplementary Table 1. Provinces and hospital of recruitment

Province	Hospital	Colorectal cancer	Breast cancer	Prostate cancer
Asturias	Hospital de Cabueñes	77	70	16
Barcelona	Hospital Clinic	69	47	53
Barcelona	Hospital de Bellvitge – ICO	375	109	-
Barcelona	Hospital del Mar	222	136	152
Barcelona	Hospital Germans Trias i Pujol	30	-	199
Cantabria	Hospital Universitario Marqués de Valdecilla	151	141	175
Gipuzkoa	Hospital Donostia	119	126	-
Gipuzkoa	Instituto Oncológico	-	100	-
Girona	Hospital Dr. Josep Trueta	-	21	-
Girona	Hospital Santa Caterina	-	26	-
Granada	Hospital San Cecilio	164	-	64
Huelva	Hospital Infanta Elena	16	24	16
Huelva	Hospital Juan Ramón Jiménez	55	84	36
León	Hospital de León	390	226	-
Madrid	Hospital La Paz	110	164	155
Madrid	Hospital Ramón y Cajal	122	177	160
Murcia	Hospital Morales Messeguer	34	-	-
Navarra	Complejo Hospitalario de Navarra A (Hospital de Navarra)	76	112	-
Navarra	Complejo Hospitalario de Navarra B (Virgen del Camino)	49	114	-
Valencia	Hospital Dr. Peset	25	4	-
Valencia	Hospital La Fe	56	57	86

Supplementary Table 2. Previous results in the MCC-Spain study

Supp. Reference	Journal	Year of publication	Cancer	Exposure
1	Environ Res	2012	NA	Disinfection by-products in municipal drinking water
2	Gac Sanit	2012	Breast, prostate	Screening practices and lifestyles
3	BJU Int	2012	Prostate	Anogenital distance
4	Gac Sanit	2013	NA	Nitrate and trace elements in municipal and bottled water
5	Int J Cancer	2015	Prostate	Night shift work and chronotype
6	Gac Sanit	2015	Colorectal, breast, prostate, gastric, CLL	Rational and study design for case-control
7	J Gen Virol	2015	CLL	Polyomaviruses
8	Infect Agent Cancer	2015	CLL	Aberrant Epstein-Barr virus
9	Menopause	2015	NA	Hormonal contraception and postmenopausal hormone therapy
10	Sci Total Environ	2015	NA	Persistent organic pollutants in adult population
11	Acta Diabetol	2016	Breast	Diabetes and diabetes treatment
12	Eur J Epidemiol	2016	Breast	Night shift work
13	Int J Cancer	2016	Colorectal	Streptococcus gallolyticus
14	Cancer Epidemiol	2016	Breast	Perinatal and childhood factors
15	Environ Health Perspect	2016	Breast	Ingested nitrate
16	Int J Cancer	2016	Colorectal	Ingested nitrate
17	Environ Health Perspect	2016	Breast	Xenoestrogen burden
18	Occup Environ Med	2016	Gastric	Night shift work
19	Cancer Epidemiol	2016	Prostate	Perinatal and childhood factors
20	Int J Cancer	2016	CLL	Night shift work
21	PLoS One	2016	Breast	Antihypertensive medication
22	BMC Cancer	2016	Breast	Non-steroidal anti-inflammatory drugs
23	PLoS One	2016	Colorectal, gastric	Menstrual and reproductive factors
24	Eur J Nutr	2017	CLL	Fruit and vegetable intake and vit C transporter gene
25	Environ Health Perspect	2017	Colorectal	Trihalomethanes in drinking water

26	Prev Med	2017	Colorectal	Drugs affecting renin-angiotensin system
27	Sci Rep	2017	Colorectal	Environmental and genetic factors
28	Scand J Work Environ Health	2017	Colorectal	Shift work
29	Int J Cancer	2017	Colorectal, breast, prostate	Nutrition-based cancer prevention guidelines
30	Front Microbiol	2017	Colorectal	Helicobacter pylori
31	PLoS One	2017	Gastric	Physical activity
32	Helicobacter	2017	NA	Helicobacter pylori in adult population
33	Maturitas	2017	Breast	Dietary patterns
34	Sci Rep	2017	Prostate	Environmental and genetic factors
35	Cancer Epidemiol	2017	Gastric	Helicobacter pylori
36	BMC Med Genet	2017	Colorectal	SMAD7 gene and Mediterranean diet
37	Gastric Cancer	2017	Gastric	Dietary patterns
38	Eur J Nutr	2018	Colorectal	Meat intake, cooking methods and doneness
39	J Urol	2018	Prostate	Dietary patterns
40	Environ Int	2018	Breast	Trihalomethanes in drinking water
41	Cancer Epidemiol	2018	CLL	CLL etiology (review)
42	Sci Rep	2018	Colorectal	Chondroitin sulphate and glucosamine
43	Sci Rep	2018	Breast	Risk score
44	Environ Pollut	2018	Breast	Residential proximity to industrial installations
45	BMC Cancer	2018	Breast	Reproductive factors and genetic hormonal pathways
46	Maturitas	2018	Breast	Meat intake, methods of cooking
47	J Steroid Biochem Mol Biol	2018	Breast	Vitamin D
48	Environ Health Perspect	2018	Breast, prostate	Artificial light-at-night
49	Haematologica	2018	CLL	Dietary patterns
50	Int J Cancer	2018	Breast, prostate	Mistimed eating patterns
51	Stat Methods Med Res	2018	NA	Compositional analysis of dietary patterns
52	Int J Hyg Environ Health	2018	Breast	Residential proximity to green spaces
53	PLos One	2018	Breast, prostate	Pigmentation phototype
54	BMC Public Health	2018	NA	Non-steroidal anti-inflammatory drugs consumption
55	Environ Int	2018	Colorectal	Sun exposure and vit D

56	Nutrients	2018	Prostate	Dietary zinc
57	Eur J Nutr	2019	Colorectal	Dietary patterns
58	Eur J Nutr	2019	Colorectal	Dietary non-enzymatic antioxidant capacity
59	Br J Haematol	2019	CLL	Insulin-like growth factor
60	Eur J Cancer Prev	2019	NA	Helicobacter pylori seroprevalence
61	Environ Int	2019	Breast, prostate	Alkylphenolic compounds
62	Nutrients	2019	NA	Mediterranean diet
63	Nutrients	2019	Gastric	Flavonoids
64	Eur J Nutr	2019	Breast	Fatty acid intake
65	Cancer Epidemiol	2019	Gastric	Epstein-Barr virus
66	Int J Cancer	2019	Prostate	Cessation of night shift work
67	Nutrients	2019	Colorectal, breast	Dietary inflammatory index and dietary non-enzymatic antioxidant capacity
68	Breast Cancer Res Treat	2019	Breast	Physical activity
69	Sci Rep	2019	Colorectal	Flagelin C and Streptococcus gallolyticus proteins

CLL: Chronic lymphocytic leukaemia. NA: Not applicable

Supplementary reference list

1. Villanueva CM, Castaño-Vinyals G, Moreno V, et al. Concentrations and correlations of disinfection by-products in municipal drinking water from an exposure assessment perspective. Environ Res. 2012;114:1-11. doi: 10.1016/j.envres.2012.02.002.
2. Perea MD, Castaño-Vinyals G, Altzibar JM, et al. [Cancer screening practices and associated lifestyles in population controls of the Spanish multi-case control study]. Gac Sanit. 2012;26:301-10. doi: 10.1016/j.gaceta.2012.01.020.
3. Castaño-Vinyals G, Carrasco E, Lorente JA, et al. Anogenital distance and the risk of prostate cancer. BJU Int. 2012;110(11 Pt B):E707-10. doi: 10.1111/j.1464-410X.2012.11516.x.
4. Espejo-Herrera N, Kogevinas M, Castaño-Vinyals G, et al and Multicase Control Study of Cancer (MCC)-Spain Water Working Group. Nitrate and trace elements in municipal and bottled water in Spain. Gac Sanit. 2013 Mar-Apr;27(2):156-60. doi:10.1016/j.gaceta.2012.02.002.
5. Papantoniou K, Castaño-Vinyals G, Espinosa A, et al. Night shift work, chronotype and prostate cancer risk in the MCC-Spain case-control study. Int J Cancer. 2015;137:1147-57. doi: 10.1002/ijc.29400.
6. Castaño-Vinyals G, Aragonés N, Pérez-Gómez B, et al and MCC-Spain Study Group. Population-based multicase-control study in common tumors in Spain (MCC-Spain): rationale and study design. Gac Sanit. 2015;29:308-15. doi: 10.1016/j.gaceta.2014.12.003.
7. Robles C, Casabonne D, Benavente Y, et al. Seroreactivity against Merkel cell polyomavirus and other polyomaviruses in chronic lymphocytic leukaemia, the MCC-Spain study. J Gen Virol. 2015;96:2286-92. doi: 10.1099/vir.0.000167.
8. Casabonne D, Benavente Y, Robles C, et al. Aberrant Epstein-Barr virus antibody patterns and chronic lymphocytic leukemia in a Spanish multicentric case-control study. Infect Agent Cancer. 2015;10:5. doi: 10.1186/1750-9378-10-5.
9. Costas L, Sequera VG, Quesada P, et al. Hormonal contraception and postmenopausal hormone therapy in Spain: time trends and patterns of use. Menopause. 2015;22:1138-46. doi: 10.1097/GME.0000000000000487.
10. Fernández-Rodríguez M, Arrebola JP, Artacho-Cordón F, et al. Levels and predictors of persistent organic pollutants in an adult population from four Spanish regions. Sci Total Environ. 2015;538:152-61. doi: 10.1016/j.scitotenv.2015.07.162.
11. García-Esquinas E, Guinó E, Castaño-Vinyals G, et al. Association of diabetes and diabetes treatment with incidence of breast cancer. Acta Diabetol. 2016;53:99-107. doi: 10.1007/s00592-015-0756-6.
12. Papantoniou K, Castaño-Vinyals G, Espinosa A, et al. Breast cancer risk and night shift work in a case-control study in a Spanish population. Eur J Epidemiol. 2016;31:867-78. doi: 10.1007/s10654-015-0073-y.

13. Butt J, Romero-Hernández B, Pérez-Gómez B, et al. Association of Streptococcus gallolyticus subspecies gallolyticus with colorectal cancer: Serological evidence. Int J Cancer. 2016;138:1670-9. doi: 10.1002/ijc.29914.

14. Lope V, García-Esquinas E, Pérez-Gómez B, et al. Perinatal and childhood factors and risk of breast cancer subtypes in adulthood. Cancer Epidemiol. 2016;40:22-30. doi: 10.1016/j.canep.2015.11.004.

15. Espejo-Herrera N, Gracia-Lavedan E, Pollan M, et al. Ingested Nitrate and Breast Cancer in the Spanish Multicase-Control Study on Cancer (MCC-Spain). Environ Health Perspect. 2016;124:1042-9. doi: 10.1289/ehp.1510334.

16. Espejo-Herrera N, Gràcia-Lavedan E, Boldo E, et al.Colorectal cancer risk and nitrate exposure through drinking water and diet. Int J Cancer. 2016;139:334-46. doi: 10.1002/ijc.30083.

17. Pastor-Barriuso R, Fernández MF, Castaño-Vinyals G, et al. Total Effective Xenoestrogen Burden in Serum Samples and Risk for Breast Cancer in a Population-Based Multicase-Control Study in Spain. Environ Health Perspect. 2016;124:1575-1582.

18. Gyarmati G, Turner MC, Castaño-Vinyals G, et al. Night shift work and stomach cancer risk in the MCC-Spain study. Occup Environ Med. 2016;73:520-7. doi: 10.1136/oemed-2016-103597.

19. Lope V, García-Esquinas E, Ruiz-Dominguez JM, et al. Perinatal and childhood factors and risk of prostate cancer in adulthood: MCC-Spain case-control study. Cancer Epidemiol. 2016;43:49-55. doi: 10.1016/j.canep.2016.06.012.

20. Costas L, Benavente Y, Olmedo-Requena R, et al. Night shift work and chronic lymphocytic leukemia in the MCC-Spain case-control study. Int J Cancer. 2016;139:1994-2000. doi: 10.1002/ijc.30272.

21. Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, et al. The Use of Antihypertensive Medication and the Risk of Breast Cancer in a Case-Control Study in a Spanish Population: The MCC-Spain Study. PLoS One. 2016;11:e0159672. doi: 10.1371/journal.pone.0159672.

22. Dierssen-Sotos T, Gómez-Acebo I, de Pedro M, et al. Use of non-steroidal anti-inflammatory drugs and risk of breast cancer: The Spanish Multi-Case-control (MCC) study. BMC Cancer. 2016;16:660. doi: 10.1186/s12885-016-2692-4.

23. Lope V, Fernández de Larrea N, Pérez-Gómez B, et al. Menstrual and Reproductive Factors and Risk of Gastric and Colorectal Cancer in Spain. PloS One. 2016;11:e0164620. doi: 10.1371/journal.pone.0164620.

24. Casabonne D, Gracia E, Espinosa A, et al. Fruit and vegetable intake and vitamin C transporter gene (SLC23A2) polymorphisms in chronic lymphocytic leukaemia. Eur J Nutr. 2017;56:1123-1133. doi: 10.1007/s00394-016-1162-8.

25. Villanueva CM, Gracia-Lavedan E, Bosetti C, et al. Colorectal Cancer and Long-Term Exposure to Trihalomethanes in Drinking Water: A Multicenter Case-Control Study in Spain and Italy. Environ Health Perspect. 2017;125:56-65. doi: 10.1289/EHP155.

26. Dierssen-Sotos T, Gómez-Acebo I, Palazuelos C, et al. Relationship between drugs affecting the renin-angiotensin system and colorectal cancer: The MCC-Spain study. Prev Med. 2017;99:178-184. doi: 10.1016/j.ypmed.2017.01.011.

27. Ibáñez-Sanz G, Díez-Villanueva A, Alonso MH, et al. Risk Model for Colorectal Cancer in Spanish Population Using Environmental and Genetic Factors: Results from the MCC-Spain study. Sci Rep. 2017;7:43263. doi: 10.1038/srep43263.

28. Papantoniou K, Castaño-Vinyals G, Espinosa A, et al. Shift work and colorectal cancer risk in the MCC-Spain case-control study. Scand J Work Environ Health. 2017;43:250-259. doi: 10.5271/sjweh.3626.

29. Romaguera D, Gracia-Lavedan E, Molinuevo A, et al. Adherence to nutrition-based cancer prevention guidelines and breast, prostate and colorectal cancer risk in the MCC-Spain case-control study. Int J Cancer. 2017;141:83-93. doi: 10.1002/ijc.30722.

30. Fernández de Larrea-Baz N, Michel A, Romero B, et al. Helicobacter pylori Antibody Reactivities and Colorectal Cancer Risk in a Case-control Study in Spain. Front Microbiol. 2017 29;8:888. doi: 10.3389/fmicb.2017.00888. Huerta JM, Chirlaque MD, Molina AJ, et al. Physical activity domains and risk of gastric adenocarcinoma in the MCC-Spain case-control study. PloS One. 2017;12:e0179731. doi: 10.1371/journal.pone.0179731.

31. Fernández-de-Larrea N, Michel A, Romero B, et al. Antibody reactivity against Helicobacter pylori proteins in a sample of the Spanish adult population in 2008-2013. Helicobacter. 2017;22. doi: 10.1111/hel.12401.

32. Castelló A, Boldo E, Pérez-Gómez B, et al. Adherence to the Western, Prudent and Mediterranean dietary patterns and breast cancer risk: MCC-Spain study. Maturitas. 2017;103:8-15. doi: 10.1016/j.maturitas.2017.06.020.

33. Gómez-Acebo I, Dierssen-Sotos T, Fernandez-Navarro P, et al. Risk Model for Prostate Cancer Using Environmental and Genetic Factors in the Spanish Multi-Case-Control (MCC) Study. Sci Rep. 2017;7:8994. doi: 10.1038/s41598-017-09386-9.

34. Fernández de Larrea-Baz N, Pérez-Gómez B, Michel A, et al. Helicobacter pylori serological biomarkers of gastric cancer risk in the MCC-Spain case-control Study. Cancer Epidemiol. 2017;50(Pt A):76-84. doi: 10.1016/j.canep.2017.08.002.

35. Alonso-Molero J, González-Donquiles C, Palazuelos C, et al. The RS4939827 polymorphism in the SMAD7 GENE and its association with Mediterranean diet in colorectal carcinogenesis. BMC Med Genet. 2017;18:122. doi: 10.1186/s12881-017-0485-5.

36. Castelló A, Fernández de Larrea N, Martín V, et al and MCC-Spain researchers. High adherence to the Western, Prudent, and Mediterranean dietary patterns and risk of gastric adenocarcinoma: MCC-Spain study. Gastric Cancer. 2018;21:372-382. doi: 10.1007/s10120-017-0774-x.

37. de Batlle J, Gracia-Lavedan E, Romaguera D, et al. Meat intake, cooking methods and doneness and risk of colorectal tumours in the Spanish multicase-control study (MCC-Spain). Eur J Nutr. 2018;57:643-653. doi: 10.1007/s00394-016-1350-6.

38. Castelló A, Boldo E, Amiano P, et al and MCC-Spain Researchers. Mediterranean Dietary Pattern is Associated with Low Risk of Aggressive Prostate Cancer: MCC-Spain Study. J Urol. 2018;199:430-437. doi: 10.1016/j.juro.2017.08.087.

39. Font-Ribera L, Gràcia-Lavedan E, Aragonés N, et al. Long-term exposure to trihalomethanes in drinking water and breast cancer in the Spanish multicase-control study on cancer (MCC-SPAIN). Environ Int. 2018;112:227-234. doi: 10.1016/j.envint.2017.12.031.

40. Benavente Y, Casabonne D, Costas L, et al. Established and suggested exposures on CLL/SLL etiology: Results from the CLL-MCC-Spain study. Cancer Epidemiol. 2018;52:106-111. doi: 10.1016/j.canep.2017.12.012.

41. Ibáñez-Sanz G, Díez-Villanueva A, Vilorio-Marqués L, et al. Possible role of chondroitin sulphate and glucosamine for primary prevention of colorectal cancer. Results from the MCC-Spain study. Sci Rep. 2018;8:2040. doi: 10.1038/s41598-018-20349-6.

42. Dierssen-Sotos T, Gómez-Acebo I, Palazuelos C, et al. Validating a breast cancer score in Spanish women. The MCC-Spain study. Sci Rep. 2018;8:3036. doi: 10.1038/s41598-018-20832-0.

43. García-Pérez J, Lope V, Pérez-Gómez B, et al. Risk of breast cancer and residential proximity to industrial installations: New findings from a multicase-control study (MCC-Spain). Environ Pollut. 2018;237:559-568. doi: 10.1016/j.envpol.2018.02.065.

44. Dierssen-Sotos T, Palazuelos-Calderón C, Jiménez-Moleón JJ, et al. Reproductive risk factors in breast cancer and genetic hormonal pathways: a gene-environment interaction in the MCC-Spain project. BMC Cancer. 2018;18:280. doi: 10.1186/s12885-018-4182-3.

45. Boldo E, Castelló A, Aragonés N, et al and MCC-Spain researchers. Meat intake, methods and degrees of cooking and breast cancer risk in the MCC-Spain study. Maturitas. 2018;110:62-70. doi: 10.1016/j.maturitas.2018.01.020.

46. Lope V, Castelló A, Mena-Bravo A, et al. Serum 25-hydroxyvitamin D and breast cancer risk by pathological subtype (MCC-Spain). J Steroid Biochem Mol Biol. 2018;182:4-13. doi: 10.1016/j.jsbmb.2018.04.005.

47. Garcia-Saenz A, Sánchez de Miguel A, Espinosa A, et al. Evaluating the Association between Artificial Light-at-Night Exposure and Breast and Prostate Cancer Risk in Spain (MCC-Spain Study). Environ Health Perspect. 2018;126:047011. doi: 10.1289/EHP1837.

48. Solans M, Castelló A, Benavente Y, et al. Adherence to the Western, Prudent, and Mediterranean dietary patterns and chronic lymphocytic leukemia in the MCC-Spain study. Haematologica. 2018;103:1881-1888. doi: 10.3324/haematol.2018.192526.

49. Kogevinas M, Espinosa A, Castelló A, et al. Effect of mistimed eating patterns on breast and prostate cancer risk (MCC-Spain Study). Int J Cancer. 2018;143:2380-2389. doi: 10.1002/ijc.31649.

50. Solans M, Coenders G, Marcos-Gragera R, et al. Compositional analysis of dietary patterns. Stat Methods Med Res. 2018 (In press). doi: 10.1177/0962280218790110.

51. O'Callaghan-Gordo C, Kogevinas M, Cirach M, et al. Residential proximity to green spaces and breast cancer risk: The multicase-control study in Spain (MCC-Spain). Int J Hyg Environ Health. 2018;221:1097-1106. doi: 10.1016/j.ijheh.2018.07.014.

52. Gómez-Acebo I, Dierssen-Sotos T, Palazuelos C, et al. Pigmentation phototype and prostate and breast cancer in a select Spanish population-A Mendelian randomization analysis in the MCC-Spain study. PloS One. 2018;13:e0201750. doi: 10.1371/journal.pone.0201750.

53. Gómez-Acebo I, Dierssen-Sotos T, de Pedro M, et al. The MCC-Spain study. BMC Public Health. 2018;18:1134. doi: 10.1186/s12889-018-6019-z.

54. Vallès X, Alonso MH, López-Caleya JF, et al. Colorectal cancer, sun exposure and dietary vitamin D and calcium intake in the MCC-Spain study. Environ Int. 2018 Dec;121(Pt 1):428-434. doi: 10.1016/j.envint.2018.09.030.

55. Gutiérrez-González E, Castelló A, Fernández-Navarro P, et al. Dietary Zinc and Risk of Prostate Cancer in Spain: MCC-Spain Study. Nutrients. 2018 Dec 20;11. pii: E18. doi: 10.3390/nu11010018.

56. Castelló A, Amiano P, Fernández de Larrea N, et al and MCC-Spain researchers. Low adherence to the western and high adherence to the mediterranean dietary patterns could prevent colorectal cancer. Eur J Nutr. 2019;58:1495-1505. doi: 10.1007/s00394-018-1674-5.

57. Amiano P, Molina-Montes E, Molinuevo A, et al. Association study of dietary non-enzymatic antioxidant capacity (NEAC) and colorectal cancer risk in the Spanish Multicase-Control Cancer (MCC-Spain) study. Eur J Nutr. 2019;58:2229-2242. doi: 10.1007/s00394-018-1773-3.

58. Casabonne D, Benavente Y, Costas L, et al. Insulin-like growth factor levels and chronic lymphocytic leukaemia: results from the MCC-Spain and EpiLymph-Spain studies. Br J Haematol. 2019;185:608-612. doi: 10.1111/bjh.15583.

59. Lorenzo I, Fernández-de-Larrea N, Michel A, et al. Helicobacter pylori seroprevalence in Spain: influence of adult and childhood sociodemographic factors. Eur J Cancer Prev. 2019;28:294-303. doi: 10.1097/CEJ.0000000000000483.

60. Peremiquel-Trillas P, Benavente Y, Martín-Bustamante M, et al. Alkylphenolic compounds and risk of breast and prostate cancer in the MCC-Spain study. Environ Int. 2019;122:389-399. doi: 10.1016/j.envint.2018.12.007.

61. Olmedo-Requena R, González-Donquiles C, et al. Agreement among Mediterranean Diet Pattern Adherence Indexes: MCC-Spain Study. Nutrients. 2019;11(3). pii: E488. doi: 10.3390/nu11030488.

62. Vitelli Storelli F, Molina AJ, Zamora-Ros R, et al. Flavonoids and the Risk of Gastric Cancer: An Exploratory Case-Control Study in the MCC-Spain Study. Nutrients. 2019;11(5). pii: E967. doi: 10.3390/nu11050967.

63. Dierssen-Sotos T, Gómez-Acebo I, Palazuelos C, et al. Fatty acid intake and breast cancer in the Spanish multicase-control study on cancer (MCC-Spain). Eur J Nutr. 2019 (In press). doi: 10.1007/s00394-019-01977-8.

64. Aragonés N, Fernández de Larrea N, Pastor-Barriuso R, et al. Epstein Barr virus antibody reactivity and gastric cancer: A population-based case-control study. Cancer Epidemiol. 2019;61:79-88. doi: 10.1016/j.canep.2019.05.008.

65. Kogevinas M, Espinosa A, Papantoniou K, et al. Prostate cancer risk decreases following cessation of night shift work. Int J Cancer. 2019 (In press). doi: 10.1002/ijc.32528.

66. Obón-Santacana M, Romaguera D, Gracia-Lavedan E, et al. Dietary Inflammatory Index, Dietary Non-Enzymatic Antioxidant Capacity, and Colorectal and Breast Cancer Risk (MCC-Spain Study). Nutrients. 2019;11. pii: E1406. doi: 10.3390/nu11061406.

67. Huerta JM, Molina AJ, Chirlaque MD, et al. Domain-specific patterns of physical activity and risk of breast cancer sub-types in the MCC-Spain study. Breast Cancer Res Treat. 2019 (In press). doi: 10.1007/s10549-019-05358-x.

68. Butt J, Fernández de Larrea N, et al. Antibody responses to flagellin C and Streptococcus gallolyticus pilus proteins in colorectal cancer. Sci Rep. 2019;9:10847. doi: 10.1038/s41598-019-47347-6.