BMJ Open

Metabolic syndrome and risk of cancer: a study protocol of case–control study using data from the Information System for the Development of Research in Primary Care (SIDIAP) in Catalonia

Diana Puente,[1,5] Tomàs López-Jiménez,[1,5] Xavier Cos-Claramunt,[2,3] Yolanda Ortega,[4] Talita Duarte-Salles[1]

ABSTRACT

Background Metabolic syndrome (MS) is defined by the clustering of specific metabolic disorders in one subject. MS is highly prevalent globally and currently considered a growing public health concern. MS comprises obesity, hypertension, dyslipidaemia and insulin resistance. Mechanisms linking MS with cancer are poorly understood, and it is as yet unknown if MS confers a greater risk than the risk entailed by each of its separate components. The main objective of this study is to compare the association between MS and 14 site-specific cancer against the association between one or two individual components of MS and cancer. The secondary objective is to evaluate the time elapsed since the diagnosis of MS and the subsequent onset of cancer within the 2006–2017 period by sex.

Methods and analysis A case–control study will be conducted for the main objective and a cohort of patients with MS will be followed for the evaluation of the second objective. Incident cases of fourteen types of cancer in patients ≥40 years of age diagnosed prospectively will be selected from electronic primary care records in the Information System for Research in Primary Care (SIDIAP database; www.sidiap.org). The SIDIAP database includes anonymous data from 6 million people (80% of the Catalan population) registered in 286 primary healthcare centres. Each matched control (four controls for each case) will have the same inclusion date, the same sex and age (±1 year) than the paired case. Logistic regression and a descriptive analysis and Kaplan-Meier analysis will be performed, in accordance with the objectives.

Ethics and dissemination The protocol of the study was approved by the IDIAP Jordi Gol Clinical Research Ethics Committee (protocol P17/212). The study's findings will be published in a peer-reviewed journal and disseminated at national and international conferences and oral presentations to researchers, clinicians and policy makers.

Strengths and limitations of this study

► Data on cancer diagnoses in the Information System for Research in Primary Care database are reliable.
► Long follow-up times.
► Underestimation of the prevalence of some diseases and left truncation.
► Waist circumference is only available for a small proportion of individuals.
► Under-recording and quality of variables in the primary care setting.

INTRODUCTION

Metabolic syndrome (MS) consists in the clustering of various cardiometabolic risk factors such as obesity, specifically central obesity as measured by waist circumference, hypertension, dyslipidaemia and insulin resistance.[1] It is a growing public health problem due to its high global prevalence. For instance, a study from the USA indicated a prevalence of MS around 24% in the general population and of 50% in patients with ischaemic cardiopathy and other vascular conditions and also a significant increase in prevalence with age.[2] In Spain, previous studies have shown that the prevalence of MS ranges between 23% and 31%; it increases with age, and until 65 years of age, it is higher in men than in women; conversely, the prevalence of MS after 65 years of age is higher in women.[3 4]

While MS was initially considered a risk factor for diabetes and cardiovascular disease,[5] emerging studies suggest that it is also a risk factor for different types of cancer. However, to our knowledge, no evidence exists to date to indicate whether MS confers a higher risk of cancer than its individual components.[6]

According to various studies,[7–10] MS is associated with a higher risk of liver, colorectal and bladder cancer in men and endometrial, pancreatic, colorectal, ovarian and postmenopausal breast cancer in women. Evidence concerning the possible association

with prostate cancer is still conflicting. Whereas a study by Esposito *et al*[11] did not find any association between MS and prostate cancer, a recent publication by Gacci *et al*[12] found a slight correlation. Data and publications for less prevalent types of cancer (incidence <6/100 000/year)[13] and for thyroid cancer, multiple myeloma, non-Hodgkin lymphoma and leukaemias are currently scarce. However, a recent meta-analysis found a higher risk of haematological cancer[14] in patients with MS.

The pathophysiology that underlies the association of MS with cancer is as yet poorly understood. Components of MS such as obesity,[15] glucose levels[16] and high blood pressure[16 17] are associated with higher risk of various types of cancer. In the pathophysiology of MS, adipose tissue is known to deregulate cytokine production, causing a chronic inflammatory state that can ultimately mediate tumorous development. Additionally, adipose tissue has been linked to insulin resistance,[18] although insulin resistance is not found in all obese patients.[19] It is thus possible for the association between MS and cancer to be mediated by the coexistence of obesity and insulin resistance as the main factors.[8] There are some patients who besides having MS also have hyperinsulinaemia, because the pancreas compensates the insulin resistance with insulin extrasecretion. Other studies suggest that high-density lipoprotein cholesterol (HDL-C) cholesterol is inversely correlated with cancer.[6 20]

Whether MS carries a higher risk than the risk of each separate MS component is unclear. Recent research looks to elucidate if the effect of MS as a whole is stronger than the effect of each component individually. Indeed, our study would allow us to analyse this aspect. For instance, in the case of endometrial cancer, obesity on its own carries a higher risk than MS as a whole.[21] MS is slightly associated with prostate cancer, but a stronger association exists between prostate cancer and hypertension.[12] With respect to kidney cancer, the time of progression to cancer has shown to be shorter in patients with MS, whereas it remains unchanged when considering any single component of MS.[22] In postmenopausal breast cancer, the association with MS is also higher than with individual components.[21] In conclusion, it is yet unclear if the risk of MS is higher than the risk of MS single components in some types of cancer.

This project will contribute new evidence on the association between MS and cancer using a large database with electronic primary care records of people that attended primary care centres in Catalonia during the 2006–2017 period. In addition, we will be able to determine the time of exposure to MS, that is, from the MS diagnosis until the diagnosis of cancer. With these data, we will be able to evaluate the role of MS in cancer risk and the risk of cancer with one or two MS components. Additionally, the use of an anonymised and validated databases facilitates the agile retrieval of retrospective data. Our database, known as the Information System for the Development of Research in Primary Care (SIDIAP) (www.sidiap.org),[23] includes approximately 6 million people, that is, 80% of the Catalan population, and thus facilitates the study of less prevalent cancer types that have been scarcely studied to date. Research studies with databases provide large amounts of information (clinical, laboratory and pharmaceutical data) at a reasonable cost. In addition, long follow-up times are made possible since the electronic medical records started in 2005. Arguably, the main justification of this study on MS and its relationship with cancer is that MS can be modified by adopting healthy lifestyles that reduce obesity, cholesterol and blood pressure. Indeed, healthy lifestyles can contribute to reduce the prevalence of MS in the adult population.[24]

The current study will provide scientific evidence for the future implementation on new preventive strategies in primary healthcare and to obtain significant results in the fight against cancer. Patients with MS could be encouraged to be screened for cancer and, more importantly, to adopt healthier lifestyles as prevention strategy.[8]

The main objective of this project will be to investigate the association of MS with 14 types of cancer diagnosed in the electronic primary care records from 1 January 2006 to 31 December 2017 in SIDIAP database, as well as to compare the association between MS and cancer versus 0, 1 or 2 single MS components. As secondary objective, from the cohort of patients with MS obtained from the case–control study, we will analyse the time elapsed between diagnosis of MS and onset of cancer.

METHODS AND ANALYSIS
Design
Case–control study for the main objective and in a subgroup of patients, longitudinal study to evaluate time of MS exposure to onset of cancer.

Data collection
This study collects all information from SIDIAP database. In Spain, primary care doctors and nurses have a pivotal role in the public health system since they have primary responsibility for the health of the population. The Spanish public health system has a coverage of over 98% of the population, most of which visit their GP at least once a year. The computerisation of primary care medical records in Catalonia started in 2005. The SIDIAP is a reliable database with information from primary care records for use in biomedical research. From 2005, the SIDIAP collects information of approximately 6 million patients[23 25] and is thus representative of the whole Catalan region. The SIDIAP includes information collected by primary care professionals during visits such as physical measurements, clinical diagnoses, laboratory tests, treatments, referral to specialist, sociodemographic and lifestyles information.

Sample size
In the SIDIAP database, we have over 350 000 patients with cancer, of which 171 275 correspond to the subtypes considered in the study. Furthermore, we estimate that

approximately 400 000 patients could be exposed to MS. Assuming an OR=1.3 (power of 80%, alpha risk 5%, no losses, 18% controls exposed to MS and four controls per case), we would need a minimum of 939 cases of cancer and 3756 controls for the main objective. This study involves all cases in the SIDIAP, and thus, we exceed the minimum cancer cases required. There are more than 2000 patients for each type of cancer registered in SIDIAP and included in the study, with the exception of Hodgkin's disease. In this case, we will have approximately 725 patients, and assuming an OR=1.3, the power will be 69%.

Study population and inclusion/exclusion criteria

All participants in the SIDIAP database of ≥40 years of age collected from 1 January 2006 to 31 December 2017.

The case – control study aims to analyse the cases of cancer that have been exposed to MS or to 0, 1 or 2 single components of MS and evaluate if the risk of MS on cancer is greater than the risk provided by 1 or 2 components. Patients with incident cancer between 1 January 2008 and 31 December 2017 will be selected. It is required that the cancer case has to be exposed for a minimum of two years to the MS; in other words, the onset of exposure to MS must be at least two years before the diagnosis of cancer. Secondary cancers and metastases will be excluded. The date of the diagnosis of cancer will be considered the index date. Four controls will be selected for each case, considering as index date of the control the date of the selection of the case. The onset of the exposure to MS in the controls, as in the cases, must begin at least two years before to index date. Each paired case– control will be of the same sex and age (± 1 year).

The cohort of patients diagnosed with MS selected in the case–control study will be followed, the time elapsed until the onset of cancer will be evaluated and the cancer-free time will be thus calculated and compared by sex.

Study variables

MS and components

For the construction of the variable MS, the criteria defined by the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity[1 26 27] published by Alberti *et al* in 2009 was used:

1. High blood pressure (defined using the International Statistical Classification of Diseases and Related Health Problems, version 10 (ICD-10) codes)): I10–I15 or dispensed pharmacological treatment of any antihypertensive drug or blood pressure ≥130/85 mm Hg.
2. Waist circumference ≥102 cm for men and ≥88 cm for women.
3. HDL-Cholesterol (High Density Lipoprotein) or drug treatment for reduced HDL-C or blood levels <40 mg/dL in men and <50 mg/dL in women.
4. Triglycerides) ≥150 mg/dL.
5. Fasting glucose levels ≥100 mg/dL.

For the diagnosis of MS, at least 3 of 5 cardiometabolic parameters are required.

MS and its components will be used to construct a composite variable: 0, 1, 2 and ≥3 (MS) components.

The first pathological value of one of these variables registered in the database will be taken as the reference date, and from that moment, it will be considered that a patient is exposed to this factor in a constant and fixed manner until the end of the study. In a sensitivity analysis, we will consider two measures of parameters separated at least by 2 weeks (maximum 1 year) to ensure that the patient has that pathological component of MS.

We will assume two considerations in relation to the WHO definition of MS[27]: since we anticipate that waist circumference will be mostly unavailable for the majority of individuals, we will also consider a body mass index >30 kg/m^2 as an MS component. Since this represents a strong limitation, we will perform a sensitivity analysis including only those individuals who have a registered waist circumference measurement in the database.

A first diagnosis can indicate the first time a condition is recorded in the data base but not necessarily the moment from which the patient is exposed to a specific condition. Specifically, this situation is a fact in those cases where the data of registration is in 2006, when the data start to be collected in the database. For this reason, we will carry out a sensitivity analysis including only those patients who have the first registration of a component from 2007.

Cancer case

We will use the ICD-10 to define cancer and we will only consider incident cases. We will select the most prevalent types of cancer with published evidence of association with MS such as colon (ICD-10 code: C18), rectum (ICD-10 code: C20), prostate (ICD-10 code: C61), liver (ICD-10 code: C22), bladder (ICD-10 code: C67), endometrium (ICD-10 code: C54), pancreas (ICD-10 code: C25) and breast (ICD-10 code: C50). Although few published data are available, we will also include lung cancer (ICD-10 code: C34) and kidney cancer (ICD-10 code: C64) because of their high prevalence. Finally, we will include types of cancer for which little evidence available, which might also be associated with MS such as thyroid cancer (ICD-10 code: C73), Hodgkin disease (ICD-10 code: C81), non-Hodgkin lymphoma (ICD-10 code: C82–85) and leukaemias (ICD-10 code: C91–95).

Other covariables

We will retrieve the following variables from the SIDIAP: sex; age; the MEDEA index ("Mortalidad en áreas pequeñas Españolas y Desigualdades Socioeconómicas y Ambientales": a deprivation index for urban census sectors that identifies the areas with the most unfavourable socioeconomic conditions, categorised in quintiles)[28]; smoking (non-smoker, smoker and ex-smoker); alcohol consumption calculated in standard units (the categories are determined by an alcohol calculator incorporated

Table 1 Study variables from the SIDIAP database in Catalonia, 2006–2017

	Variables	Period
Main variable (independent)	Metabolic syndrome.	First diagnosis 2006–2017.
	Components: HBP, waist circumference, BMI, HDL, TG and fasting glucose.	First diagnosis 2006–2017 of 0, 1, 2 MS components and no diagnosis of 3 MS components.
Dependent variable	Incident cancers: ICD-10 codes and location.	Incident cancer 2008–2017 (at least 2 years of exposure).
Covariables	Sex, age, the MEDEA index, smoking status, alcohol consumption, physical activity, drugs, urban/rural setting, number of visits to the GP, number of pregnancies, age at menarche and menopause.	Collected 2 years before cancer diagnosis

MEDEA index: deprivation index for urban census sectors that identifies the areas with the most unfavourable socioeconomic conditions.
BMI, body mass index; GP, general practitioner; HDL, high-density lipoprotein; HPB, high blood pressure; ICD-10, International Statistical Classification of Diseases and Related Health Problems, version 10; MS, metabolic syndrome; SIDIAP, Information System for the Development of Research in Primary Care; TG, triglycerides.

into the electronic clinical record) (no alcohol, low, moderate and high consumption); physical activity (none, low, moderate and high); use of drugs (ATC classification system (Anatomical Therapeutic Chemical Classification) and information collected from drugs dispensed in pharmacies); area (rural and urban); number of visits to the primary healthcare centre; number of pregnancies; age at menarche; and menopause (yes/no) and its date. These data will be collected at least two years before the diagnosis of cancer.

A summary of the variables collected from SIDIAP is presented in table 1.

Statistical analysis

Descriptive statistics will be used for each variable to identify asymmetric distributions. The continuous variables will be analysed as mean (SD) or median (25th and 75th centiles) based on the normality/non-normality of the distribution, and categorical variables will be described as percentages. To evaluate baseline differences, the appropriate statistics will be applied based on the type of variable and their distribution (χ^2, F-distribution, Student's t-distribution, analysis of variance, Mann-Whitney U and Kruskal-Wallis).

For the case–control study, we will carry out a logistic regression model adjusted for the potential confounders described in the covariables section to analyse the risk of MS on cancer and the risk of cancer comparing the effect of MS versus the individual components of MS. A descriptive analysis will determine the time elapsed between exposure to MS to cancer event. Kaplan-Meier curve will be computed by each sex and logrank, and Breslow tests will be applied to compare curves.

The analysis will be stratified by type of cancer.
Sensitivity analysis:
1. Two measures will be necessary to consider a value of specific factor as pathological.
2. We will conduct an analysis in those cases that waist circumference is registered.

3. We will conduct an analysis with those patients who in 2006 have not registered any parameter and the first registration is in 2007.

All analyses will be carried out with the statistical packages SPSS V.24 and Stata V.15.

Patient and public involvement

The study is based on real world data in which the anonymity and confidentiality of the data are guaranteed. Even so, although patient are not involved in the development of study, patient involvement will be guaranteed in the dissemination of data through different channels of communication, adapting the results of the study to citizen participation and adapting it to the natural language of citizens/patients. The dissemination of the possible effects of the MS on cancer risk can be useful to encourage citizens to reduce MS through and improvement of lifestyles.

This study is not a randomised clinical trial.

DISCUSSION

This project will contribute new evidence on the association between MS and 14 site-specific cancers in a large database containing the electronic primary care records of approximately 6 million people in Catalonia during the 2006–2017 period.

Recent studies have suggested that MS might be a risk factor for different cancer sites. However, to date, the role of MS in cancer development is still uncertain. The possible link between MS and cancer has been scarcely studied in our setting.

MS can be modulated by adopting healthy lifestyles that reduce obesity, cholesterol and blood pressure. Since cancer is one of the main causes of mortality worldwide, the possible association with MS could have a significant clinical and financial impact on health services. If the results of the study show that there is a statistically significant association, these results will be added to existing

recommendations on cancer, the risk of MS as a whole rather than only the risk originating from individual MS components. The training of primary care physicians and nurses will be enhanced by the results of this study, specifically with regard to cancer risk factors and the reduction of MS by means of healthy lifestyles.

We foresee the following potential limitations in this study: the possible underestimation of the prevalence of some diseases and a left truncation, that is, the onset of some diseases and risk factors will remain unknown since the database starts in 2005, when in reality the diagnosis of MS or some of its components might precede this date. In addition, we need to take into account the possible under-recording and quality of variables in the primary care setting and the lack of data on diet, a lifestyle item that can affect MS components and cancer risk. Data on cancer diagnoses in the SIDIAP database have been shown to be reliable, as shown in preliminary results from a validation study performed in collaboration with two populations' cancer registries in Catalonia (results still not published). Another possible limitation is the difficulty in obtaining adequate controls. The design of the study with paired controls selected from the SIDIAP minimises this potential bias.

Author affiliations
[1]Fundació Institut Universitari per a la recerca a l'Atenció Primària de Salut Jordi Gol i Gurina (IDIAPJGol), Barcelona, Spain
[2]Foundation Network of Study Groups of Diabetes in Primary Care (redGDPS), Sabadell, Spain
[3]Primary Care Centre Sant Martí de Provençals. Primary Care Management Barcelona Ciutat, Catalan Institute of Health, Barcelona, Spain
[4]CAP Salou, Institut Català de la Salut, Tarragona, Spain
[5]Universitat Autònoma de Barcelona, Bellaterra, Cerdanyola del Vallès, Spain

REFERENCES

1. Alberti KG, Eckel RH, Grundy SM, *et al*. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation Task Force on Epidemiology and Prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International Atherosclerosis Society; and International Association for the Study of Obesity. *Circulation* 2009;120:1640–5.
2. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *JAMA* 2002;287:356–9.
3. Fernández-Bergés D, Cabrera de León A, Sanz H, *et al*. Metabolic syndrome in Spain: prevalence and coronary risk associated with harmonized definition and WHO proposal. DARIOS study. *Rev Esp Cardiol* 2012;65:241–8.
4. Guallar-Castillón P, Pérez RF, López García E, *et al*. Magnitude and management of metabolic syndrome in Spain in 2008-2010: the ENRICA study. *Rev Esp Cardiol* 2014;67:367–73.
5. Blanc-Lapierre A, Spence A, Karakiewicz PI, *et al*. Metabolic syndrome and prostate cancer risk in a population-based case-control study in Montreal, Canada. *BMC Public Health* 2015;15:913.
6. Harding J, Sooriyakumaran M, Anstey KJ, *et al*. The metabolic syndrome and cancer: Is the metabolic syndrome useful for predicting cancer risk above and beyond its individual components? *Diabetes Metab* 2015;41:463–9.
7. Chen Y, Zhang L, Liu W, *et al*. Case-control study of metabolic syndrome and ovarian cancer in Chinese population. *Nutr Metab* 2017;14:21.
8. Esposito K, Chiodini P, Colao A, *et al*. Metabolic syndrome and risk of cancer: a systematic review and meta-analysis. *Diabetes Care* 2012;35:2402–11.
9. Fiori E, Lamazza A, De Masi E, *et al*. Association of liver steatosis with colorectal cancer and adenoma in patients with metabolic syndrome. *Anticancer Res* 2015;35:2211–4.
10. Montella M, Di Maso M, Crispo A, *et al*. Metabolic syndrome and the risk of urothelial carcinoma of the bladder: a case-control study. *BMC Cancer* 2015;15:720.
11. Esposito K, Chiodini P, Capuano A, *et al*. Effect of metabolic syndrome and its components on prostate cancer risk: meta-analysis. *J Endocrinol Invest* 2013;36:132–9.
12. Gacci M, Russo GI, De Nunzio C, *et al*. Meta-analysis of metabolic syndrome and prostate cancer. *Prostate Cancer Prostatic Dis* 2017;20:146–55.
13. Gatta G, van der Zwan JM, Casali PG, *et al*. Rare cancers are not so rare: the rare cancer burden in Europe. *Eur J Cancer* 2011;47:2493–511.
14. Li C, Liu P, Liu L, *et al*. Metabolic syndrome in hematologic malignancies survivors: a meta-analysis. *Med Oncol* 2015;32:422.
15. Bhaskaran K, Douglas I, Forbes H, *et al*. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5·24 million UK adults. *Lancet* 2014;384:755–65.
16. Stocks T, Lukanova A, Johansson M, *et al*. Components of the metabolic syndrome and colorectal cancer risk; a prospective study. Int J Obes (Lond. 2008;32:304–14.
17. Stocks T, Van Hemelrijck M, Manjer J, *et al*. Blood pressure and risk of cancer incidence and mortality in the Metabolic Syndrome and Cancer Project. *Hypertension* 2012;59:802–10.
18. Mendonça FM, de Sousa FR, Barbosa AL, *et al*. Metabolic syndrome and risk of cancer: which link? *Metabolism* 2015;64:182–9.
19. Abbasi F, Brown BW, Lamendola C, *et al*. Relationship between obesity, insulin resistance, and coronary heart disease risk. *J Am Coll Cardiol* 2002;40:937–43.
20. Jafri H, Alsheikh-Ali AA, Karas RH. Baseline and on-treatment high-density lipoprotein cholesterol and the risk of cancer in randomized controlled trials of lipid-altering therapy. *J Am Coll Cardiol* 2010;55:2846–54.
21. Esposito K, Chiodini P, Capuano A, *et al*. Metabolic syndrome and endometrial cancer: a meta-analysis. *Endocrine* 2014;45:28–36.
22. Kriegmair MC, Mandel P, Porubsky S, *et al*. Metabolic Syndrome Negatively Impacts the Outcome of Localized Renal Cell Carcinoma. *Horm Cancer* 2017;8:127–34.

23. Bolíbar B, Fina Avilés F, Morros R, *et al*. [SIDIAP database: electronic clinical records in primary care as a source of information for epidemiologic research]. *Med Clin* 2012;138:617–21.
24. Giugliano D, Ceriello A, Esposito K. Are there specific treatments for the metabolic syndrome? *Am J Clin Nutr* 2008;87:8–11.
25. García-Gil MM, Hermosilla E, Prieto-Alhambra D, *et al*. Construction and validation of a scoring system for the selection of high-quality data in a Spanish population primary care database (SIDIAP). *Inform Prim Care* 2011;19:135–45.
26. Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults. Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III). *JAMA* 2001;285:2486–97.
27. Grundy SM, Brewer HB, Cleeman JI, *et al*. Definition of metabolic syndrome: report of the National Heart, Lung, and Blood Institute/ American Heart Association conference on scientific issues related to definition. *Circulation* 2004;109:433–8.
28. Felícitas Domínguez-Berjón M, Borrell C, Cano-Serral G, *et al*. Construcción de un índice de privación a partir de datos censales en grandes ciudades españolas (Proyecto MEDEA). *Gac Sanit* 2008;22:179–87.