Title: A systems approach identifies time-dependent associations of multimorbidities with pancreatic cancer risk

Short title: Multimorbidity and pancreatic cancer

Authors: P. Gomez-Rubio[1], V. Rosato[2,3], M. Marquez[1], C. Bosetti[4], E. Molina-Montes[1], M. Rava[1], J. Piñero[5], C. W. Michalski[6,7], A. Farré[8], X. Molero[9,10,11], M. Löhr[12], L. Ilzarbe[13], J. Perea[14], W. Greenhalf[15], M. O'Rorke[16], A. Tardón[17,11], T. Gress[18], V. M. Barberà[19], T. Crnogorac-Jurcevic[20], L. Muñoz-Bellvís[21], E. Domínguez-Muñoz[22], A. Gutiérrez-Sacristán[5], J. Balsells[9,10,11], E. Costello[15], C. Guillén-Ponce[23], J. Huang[12], M. Iglesias[13], J. Kleeff[6,15], B. Kong[6], J. Mora[8], L. Murray[16], D. O'Driscoll[24], P. Peláez[14], I. Poves[13], R. T. Lawlor[26], A. Carrato[23], , M. Hidalgo[27], A. Scarpa[26], L. Sharp[24,25], L. I. Furlong[5], F. X. Real[28,29], C. La Vecchia[2], N. Malats[1] on behalf of the PanGenEU Study Investigators[30]

Authors' affiliations

(1) Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Center (CNIO), Madrid, and CIBERONC, Spain.

(2) Branch of Medical Statistics, Biometry and Epidemiology "G.A. Maccacaro", Department of Clinical Sciences and Community Health, University of Milan., Milan, Italy.

(3) Unity of Medical Statistics, Biometry and Bioinformatics, National Cancer Institute, IRCCS Foundation, Milan, Italy.

(4) Department of Epidemiology, Mario Negri Institute for Pharmacological Research-IRCCS, Milan, Italy.

(5) Research Programme on Biomedical Informatics (GRIB), Hospital del Mar Research Institute (IMIM), Pompeu Fabra Univeristy (UPF), Barcelona, Spain.

(6) Department of Surgery, Technical University of Munich, Munich, Germany.

(7) Department of Surgery, University of Heidelberg, Heidelberg, Germany.

(8) Department of Gastroenterology, Santa Creu i Sant Pau Hospital, Barcelona, Spain.

(9) Exocrine Pancreas Research Unit and Vall d'Hebron Research Institute (VHIR), Vall d'Hebron University Hospital, Barcelona, Spain.

(10) Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain.

(11) Network of Biomedical Research Centres (CIBER), Hepatic and Digestive Diseases and Epidemiology and Public Health, Madrid, Spain

(12) Gastrocentrum, Karolinska Institutet and University Hospital, Stockholm, Sweden.

(13) Department of Gastroenterology, del Mar—Parc de Salut Mar University Hospital, Barcelona, Spain.

(14) Department of Surgery, University Hospital 12 de Octubre, Madrid, Spain.

(15) Department of Molecular and Clinical Cancer Medicine, The Royal Liverpool University Hospital, Liverpool, UK.

(16) Centre for Public Health, Queen's University Belfast, Belfast, UK.

(17) Department of Medicine, University Institute of Oncology of Asturias, Oviedo, Spain.

(18) Department of Gastroenterology, University Hospital of Giessen and Marburg, Marburg, Germany.

(19) Molecular Genetics Laboratory, General University Hospital of Elche , Elche, Spain.

(20) Centre for Molecular Oncology, Barts Cancer Institute, Queen Mary University of London, John Vane Science Centre, London, UK.

(21) General and Digestive Surgery Department, Salamanca University Hospital, Salamanca, Spain.

(22) Department of Gastroenterology, University Clinical Hospital of Santiago de Compostela, Santiago de Compostela, Spain.

(23) Department of Oncology, Ramón y Cajal Hospital, Madrid, and CIBERONC, Spain.

(24) National Cancer Registry Ireland, Cork, Ireland,

(25) Institute of Health & Society, Newcastle University, United Kindom.

(26) ARC-Net centre for Applied Research on Cancer and Department of Pathology and Diagnostics, University and Hospital trust of Verona, Verona, Italy.

(27) Clara Campal Integrated Oncological Centre, Sanchinarro Hospital, Madrid, Spain.

(28) Epithelial Carcinogenesis Group, Spanish National Cancer Research Centre (CNIO), Madrid, and CIBERONC, Spain.

(29) Department of experimental and health sciences, Pompeu Fabra University, Barcelona, Spain.

(30) PanGenEU Study Investigators (Supplemental Annex S1).

Abbreviations

Pancreatic ductal adenocarcinoma, PDAC; Type 2 diabetes mellitus, T2DM; Metabolic syndrome, MetS; *Helocobacter pylori*, *H. pylori*; multimorbidity pattern, MP.

Corresponding author

Dr. Núria Malats,

Genetic and Molecular Epidemiology Group, Spanish National Cancer Research Center (CNIO), C/Melchor Fernandez Almagro, 3, 28029, Madrid, Spain, Phone: +34-912-246-900 (ext.3330), Fax: +34-912-246-911, E-mail: nmalats@cnio.es

Abstract: 277 words

Text: 3680 words

References: 26

Tables: 2

Figures: 1

Supplementary material: Supplemental Annex, Supplementary Methods, 6 Supplementary Tables, and 3 Supplementary Figure

Abstract

Background

Pancreatic ductal adenocarcinoma (PDAC) is usually diagnosed in late adulthood; therefore, many patients suffer or have suffered from other diseases. Identifying disease-patterns associated with PDAC risk may enable a better characterization of high-risk patients.

Methods

Multimorbidity patterns (MPs) were assessed from 17 self-reported conditions using hierarchical clustering, principal component, and factor analyses in 1705 PDAC cases and 1084 controls from a European population. Their association with PDAC was evaluated using adjusted logistic regression models. Time since diagnosis of morbidities to PDAC diagnosis/recruitment was stratified into recent (<3 years) and long-term (≥3 years). The MPs and PDAC genetic networks were explored with DisGeNET bioinformatics-tool which focuses on gene-diseases associations available in curated databases.

Results

Three MPs were observed: gastric (heartburn, acid regurgitation, *H. pylori* infection, and ulcer), metabolic syndrome (obesity, type-2 diabetes, hypercholesterolemia, and hypertension), and atopic (nasal allergies, skin allergies, and asthma). Strong associations with PDAC were observed for ≥2 recently diagnosed gastric conditions (odds ratio [OR], 6.13; 95%CI 3.01-12.5) and for ≥3 recently diagnosed metabolic syndrome conditions (OR, 1.61; 95%CI 1.11-2.35). Atopic conditions were negatively associated with PDAC (high adherence score OR for tertile III, 0.45; 95%CI 0.36 – 0.55). Combining type-2 diabetes with gastric MP resulted in higher PDAC risk for recent (OR, 7.89; 95%CI 3.9-16.1) and long-term diagnosed conditions (OR, 1.86; 95%CI 1.29-2.67). A common genetic basis between MPs and PDAC was observed in the bioinformatics analysis.

Conclusions

Specific multimorbidities aggregate and associate with PDAC in a time-dependent manner. A better characterization of a high-risk population for PDAC may help in the early diagnosis of this cancer. The common genetic basis between MP and PDAC points to a mechanistic link between these conditions.

Keywords

Pancreatic cancer; Multimorbidity; Risk.

Key Message (400characters max (spaces included)= 397)

We explored the co-occurrence of morbidities and their common effect on pancreatic cancer. We showed that some groups of morbidities cluster together. Given that multimorbidity is frequent in the adult population, evaluating the collective effect of patterns of conditions will allow a better characterization of risk for this malignancy potentially improving the management of affected patients.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is a dreadful disease usually diagnosed in late adulthood; hence, many patients suffer from other diseases [1,2]. PDAC patients present certain morbidities more often than non-PDAC patients, including chronic pancreatitis, type two diabetes mellitus (T2DM) or obesity [3–5]. Few studies focus in constellations of morbidities, and most explore associations between individual conditions and PDAC risk. In PDAC, multimorbidity has been mainly studied in the context of the metabolic syndrome (MetS), a cluster of conditions including abdominal obesity, hypertension, fasting hyperglycemia and serum hypertriglyceridemia, and low HDL cholesterol [6]. Having $\geq$3 MetS conditions has been associated with a 2-fold increased PDAC risk [7] yet, few studies have characterized in detail this relationship [8–12]. Additional multimorbidity patterns (MPs) affecting PDAC might exist, providing key information on disease mechanisms and contributing to the development of strategies to improve risk assessment.

We explored the patterns of co-occurrence of 17 self-reported medical conditions and their association with PDAC in a European case-control population. Identifying MPs and evaluating their effect on PDAC risk provides clues on common (genetic and/or environmental) backgrounds and allow us to obtain a finer characterization of risk. The assessment of time since diagnosis of these conditions allowed us to evaluate whether the morbidities represent a risk factor or an early manifestation of the malignancy.

Methods

Study population

The PanGenEU case-control study recruited PDAC patients $\geq$18 years old from Spain, United Kingdom, Germany, Ireland, Italy, and Sweden between 2009 and 2014 (Supplementary Annex S1, Table S2). Italy contributed with cases only. Controls were hospital-based, except in Ireland and Sweden where controls were population-based. Controls had no history of PDAC and principal diagnosis at admission was unrelated to known risk factors of PDAC (Supplementary Table S1). Response rate was 86.3% in cases and 77.8% in controls. IRB approval and written informed consent were obtained from all centers and participants.

Health monitors performed in-person interviews; subjects answered “yes/no” to “Has your doctor ever told you that you had any of the following illnesses, health problems or procedures?” for 25 medical conditions (Supplementary TableS3). Obesity (Body mass index,

BMI ≥30 kg/m^2) was calculated using reported usual adult height and weight two years before recruitment. Only conditions with prevalence ≥2% were considered. Subjects without information in the medical questions were removed (*n*=268, 8.8%), leaving 1705 cases and 1084 controls. Age, sex and smoking were not statistically different between included and excluded subjects (*p*-value >0.05).

Statistical analysis

Data was imputed using missForest R package (Supplementary Methods). Adjustment variables were selected using the 10% change of estimate method. Statistical significance was considered as *p*-value <0.05. Analyses were performed using R version 3.1.2.

All combinations of two (comorbidity) and three morbidities (trimorbidity) within an individual were considered (Supplementary Methods). MPs were defined separately for cases and controls using hierarchical cluster analysis, principal component (PCA), and factor (FA) analyses (Supplementary Methods). Variables were created by combining presence/absence of conditions (0=none, 1=any one, 2=any two, 3=any three or more) for gastric (*H. pylori*, acid regurgitation, heartburn and ulcer), MetS (T2D, hypertension, obesity and hypercholesterolemia), and atopic (asthma, nasal and skin allergies) patterns. To determine the degree of a subject belonging to the identified patterns, i.e. adherence to a pattern, factor scores were calculated: low and high adherence for the first and third tertile, respectively. Logistic regressions were performed to evaluate the association of each MP and adherence tertiles with PDAC. Time since diagnosis of morbidities to PDAC diagnosis/recruitment was explored through stratification as recent (<3 years) and long-term (≥3 years). Variables whose associations resulted in an OR change >1.5 after stratification by time since diagnosis were combined into a single variable for further analysis. Interaction between MPs and age, sex, and smoking were calculated comparing models with and without interaction terms through likelihood ratio test. Additive interactions of morbidities within each MP were also tested (Supplementary Methods). Excluding patients from Italy did not modify the results. Heterogeneity by country was not significant (*p*-value >0.2). Internal validation of these models was also performed (Supplementary Methods)

Multimorbidity system analysis was performed with DisGeNET (Supplementary Methods).

Results

Demographic and lifestyle information is provided in Supplementary Table S2.

Individual conditions.

Number of morbidities reported was similar among cases and controls (~7% zero morbidities; ~38% ≥4 morbidities). When comparing subjects reporting zero morbidities with those reporting ≥4, the former were younger (mean age 59 vs. 67 years), less heavy smokers (15.5% vs. 20.1%), and more commonly males (61.1% vs. 53.3%).

Comorbidity assessment.

Most morbidities (81%) associated with each other in the same direction in cases and controls, with only moderate differences (Supplementary Figure 1). Nine comorbidities and four trimorbidities were significantly associated with PDAC after multiple test correction (Supplementary Table S4). The strongest positive association was observed for "T2DM and acid regurgitation" (OR, 4.25; 95%CI, 2.55-7.08). The average SPP for these associations was 93.3, i.e. if the analysis were repeated 100 times the association would be statistically significant 93 times after multiple test correction. Significant trimorbidities (average SPP, 93.4) always included T2DM and a combination of other conditions with ORs around 4, suggesting that the association of T2DM with PDAC risk might be potentiated by some morbidities.

Multimorbidity assessment.

Three clear patterns of conditions emerged: 1) Gastric: heartburn, acid regurgitation, *H. pylori*, and ulcer; 2) MetS: obesity, T2DM, hypercholesterolemia and hypertension; and 3) Atopic: nasal allergies, skin allergies and asthma (Figure 1). The remaining conditions did not follow a consistent pattern of co-occurrence and were not considered in further analyses.

A positive association was observed between PDAC risk and the gastric and MetS patterns with a positive trend (p-value <0.01); a negative association was observed in the atopic pattern (Table 1). Similar findings were observed when comparing the highest with the lowest adherence tertiles. An interaction was observed between MetS and sex; ≥3 MetS conditions

was associated with PDAC in males but not in females (*p*-value = 0.032, Supplementary Table S5).

Time since diagnosis.

Associations do not necessarily indicate causality; therefore, we assessed the associations in relation to the period since diagnosis of conditions and PDAC. Recent diagnosis of the gastric pattern showed stronger association with PDAC risk than long-term diagnosis (OR and 95%CI: 6.13, 3.01-12.5; 1.17, 0.79-1.73; respectively). Stratification of the MetS pattern was done based of T2DM diagnosis because the other conditions did not show different associations with PDAC by time since diagnosis: OR, 1.61; 95%CI, 1.11-2.35 for recent and OR, 1.35; 95%CI, 0.99-1.85 for long-term diagnosis. This analysis could not be performed for the atopic pattern because time since diagnosis was only available for asthma.

Gastric MP and T2DM showed the greatest change in their association with PDAC after stratification by time since diagnosis (Supplementary Table S3). Considered altogether, having $\geq$1 of these was strongly associated with PDAC among subjects with recent diagnosis (OR, 7.89; 95%CI, 3.9-16.1 for $\geq$2 conditions) while a moderate estimate was observed among subjects with long-term diagnosis (OR, 1.86; 95%CI, 1.29-2.67 for $\geq$3 conditions, Table 2). The average AUC for this model was 0.78 (SD= 0.011) for recently diagnosed conditions; and 0.72 (SD= 0.01) and 0.74 (SD= 0.009) for long-term and lifetime models, respectively (data not shown). Treating obesity as a confounder showed similar results (Supplementary Table S6).

Bioinformatics analysis showed a stronger genetic link among morbidities included in the metabolic and atopic patterns in comparison to those of the gastric pattern, possibly due to less reported disease-gene associations in the latter. All patterns shared genetic links with PDAC (Supplementary Figure S2), and many of these conditions shared inflammatory related genes such as *TNF*, *CXCL8*, *HIF1A*, and *PTGS2* (Supplementary Figure S3).

Discussion

We analyzed the aggregation and association of 17 self-reported medical conditions with PDAC risk in a large case-control European study. We identified three MPs: the gastric and MetS patterns, positively associated with PDAC, and the atopic pattern, negatively associated.

Moreover, we explored MPs in the temporal context of pancreatic cancer which provides clues about common mechanisms and causal effect.

Among the gastric pattern, sensitivity analysis showed that heartburn and acid regurgitation had the strongest effect on the association. In the MetS pattern, T2DM was the main driving condition since removing T2DM from the pattern resulted in loss of significance. In the atopic pattern, none of the conditions seemed to drive the association individually.

Associations between gastric and MetS patterns with PDAC were significant only among subjects with recently diagnosed conditions. Our results strengthen previous findings with smaller case sample size or pooled heterogeneous populations reporting positive associations between history of gastric or duodenal ulcer with PDAC only among recently diagnosed subjects [2, 13, 14]. Two cohort studies reported a significant association between PDAC and long-term diagnosis (up to 20 years) of gastric but not duodenal ulcer [15, 16]; whether gastric ulcer was related to *H. pylori* infection was not elucidated. A recent meta-analysis including ten case-control studies reported a non-significant association between overall *H. pylori* infection and PDAC [17]. These studies did not consider time since diagnosis and are limited in their ability to assess causality. Less information exists for acid regurgitation or heartburn. One study reported a significant positive association between PDAC and episodes of $\geq$4 weeks of heartburn up to 5 years prior to cancer diagnosis/interview [18]. For the MetS pattern, this observation was attributed to T2DM confirming and extending previous reports of a stronger positive association among new-onset diabetics [4, 19]. Importantly, the significant association between three or more MetS conditions and PDAC was restricted to males, pointing to a potential role of BMI since, contrary to females, there were more obese males among cases (22.6%) than controls (18.9%). Previous studies reported contradictory results on the interaction between MetS and sex [7, 8, 11, 12], more studies are needed to confirm our observations. Although recent asthma diagnosis was reported by <1%, long-term asthmatics were negatively associated with PDAC. Consistently, a significant inverse association between PDAC and nasal allergies unrelated to disease duration has been reported [20].

Five conditions (i.e. T2DM and gastric conditions) stood out by the magnitude of OR change after stratification by time since diagnosis. When consolidated into a single variable, a stronger positive association was observed between having $\geq$1 of these conditions recently

diagnosed and PDAC. T2DM was an important driver of this association, but having ≥3 of these conditions showed a stronger estimate than by T2DM alone highlighting the potential significance of combining morbidities when studying PDAC. Bootstrapped AUCs showed a fair performance of the models including these five conditions; restriction to recent diagnosis resulted in the best performance. Conceivably, additive interactions between morbidities within MPs could explain these results. In this regard we observed a significant additive interaction between asthma and both nasal and skin allergies, and a significant sub-additive interaction between diabetes and the combination of any three or more gastric conditions; however, we were limited by sample size in these analyses. Mechanistically, these associations might be explained in the context of systemic inflammation: early events during PDAC carcinogenesis could favor the development of type 3c diabetes, altogether prompting dysbiosis which could, in turn, be manifested as gastric conditions. Further analysis showed an intricate genetic system among the MPs and pancreatic cancer pointing to a potential mechanistic link between these (Supplementary Figure S2 and S3).

It has been hypothesized that PDAC development might occur during the ten years previous to its detection with disease manifestations in the preceding 1-2 years. Yet, there are arguments in favor of a rapid "catastrophic" evolution shortly before diagnosis [21]. We aimed at discriminating medical conditions acting as true risk/protective factors (long-term diagnosis) vs. early manifestations of PDAC (recent diagnosis). We show that long-term T2DM and asthma are associated with PDAC, pointing to them as true risk and protective factors, respectively. Interestingly, the co-occurrence of long-term T2DM with long-term gastric conditions was associated with a stronger PDAC risk than that conferred by T2DM alone indicating that while T2DM is a main risk factor of PDAC, its overlap with other gastric conditions might further aggravate carcinogenesis. We also provide strong evidence that several conditions associate with PDAC risk only when diagnosed shortly before PDAC, suggesting that these could be PDAC manifestations. This concept has been discussed in the context of T2DM [4, 13, 19, 22, 23], and it is further supported by studies showing improvement of insulin resistance and glucose intolerance after PDAC resection [24, 25]. However, we cannot rule out that pathophysiologically distinct subtypes of T2DM and/or other conditions might accelerate PDAC progression resulting in a shorter time between diagnoses.

Early PDAC diagnosis provides the only opportunity for long-term survival. Establishing a set of conditions alerting the clinical staff of potential PDAC cases could improve diagnosis, management and/or treatment of patients and consequently, the outcome of the disease. However, the symptoms/conditions here described are relatively non-specific hindering their immediate adscription to PDAC. Moreover, whether PDAC stage at diagnosis modifies the reported associations deserves further investigation. While innovative and detailed analyses have been performed to both identify MPs and to assess their risk with PDAC, bias by unmeasured confounders is always a concern (i.e., metformin or statins medications have been associated with decreased risk of PDAC). Future studies should focus on exploring the role of particular treatments in these associations.

Our study has some limitations. The information was self-reported thus, misclassification and bias cannot be completely ruled out; nevertheless, the consistency of basic findings in this report with existing literature argues against this possibility. Our study is one of the largest performed but - despite its size and the restriction of our analyses to conditions with a frequency $\geq$2% in cases and controls – statistical power remains an issue when considering multimorbidities and stratification by time since diagnosis.

To our knowledge, this is the first study that simultaneously considers the association of many morbidities with PDAC. We report three main MPs significantly associated with PDAC and show that the associations change depending on the time since diagnosis of morbidities. Owing the high prevalence of multimorbidities, evaluating MPs may help to improve the characterization of PDAC risk. Confirmation of these results in independent studies will be critical in their generalization and future clinical application. Identification of early manifestations of PDAC may help define subpopulations of patients in whom screening might be cost-effective [26]. Moreover, the discovery of risk and protective associations should lead to more pathophysiological studies. Thus, the knowledge generated could contribute to implement novel preventive and therapeutic strategies.

Acknowledgments

We acknowledge the COST Action BM1204 EUPancreas "An integrated European platform for pancreatic cancer research: from basic science to clinical and public health interventions for a rare disease", coordinators, field and administrative workers, technicians and study participants of the *European Study into Digestive Illnesses and Genetics* (PanGenEU) study.

Funding

The work was partially supported by Fondo de Investigaciones Sanitarias (FIS), Instituto de Salud Carlos III-FEDER, Spain (#PI11/01542, #PI0902102, #PI12/01635, #PI12/00815, #PI13/00082, CP10/00524); Red Temática de Investigación Cooperativa en Cáncer, Spain (#RD12/0036/0034, #RD12/0036/0050, #RD12/0036/0073); World Cancer Research (WCR #15-0391); Acción Especial de Genómica, Spain (#GEN2001-4748-c05-03); *EU-6FP Integrated Project* (#018771-MOLDIAG-PACA), EU-FP7-HEALTH (#259737-CANCERALIA, #256974-EPC-TM-Net, #602783-Cam-Pac), Associazione Italiana Ricerca sul Cancro (AIRC #12182); Cancer Focus Northern Ireland and Department for Employment and Learning; ALF (#SLL20130022), Sweden; Italian Foundation for Cancer Research (FIRC). The funders had no role in the study design or the analysis and interpretation of the data.

Disclosure

The authors have declared no conflicts of interest.

References

1. Olson SH. Selected medical conditions and risk of pancreatic cancer. Mol. Carcinog. 2012; 51(1):75–97.

2. Ko AH, Wang F, Holly E a. Pancreatic cancer and medical history in a population-based case-control study in the San Francisco Bay Area, California. Cancer Causes Control 2007; 18(8):809–819.

3. Raimondi S, Lowenfels AB, Morselli-Labate AM et al. Pancreatic cancer in chronic pancreatitis; aetiology, incidence, and early detection. Best Pract. Res. Clin. Gastroenterol. 2010; 24(3):349–358.

4. Bosetti C, Rosato V, Li D et al. Diabetes, antidiabetic medications, and pancreatic cancer risk: an analysis from the International Pancreatic Cancer Case-Control Consortium. Ann. Oncol. 2014; 25(10):2065–2072.

5. Genkinger JM, Spiegelman D, Anderson KE et al. A pooled analysis of 14 cohort studies of anthropometric factors and pancreatic cancer risk. Int. J. Cancer 2011; 129(7):1708–1717.

6. Eckel RH, Alberti K, Grundy SM, Zimmet PZ. The metabolic syndrome. Lancet 2010; 375(9710):181–183.

7. Rosato V, Tavani A, Bosetti C et al. Metabolic syndrome and pancreatic cancer risk: A case-control study in Italy and meta-analysis. Metabolism. 2011; 60(10):1372–1378.

8. Russo A, Autelitano M, Bisanti L. Metabolic syndrome and cancer risk. Eur. J. Cancer 2008; 44(2):293–297.

9. Inoue M, Noda M, Kurahashi N et al. Impact of metabolic factors on subsequent cancer risk: results from a large-scale population-based cohort study in Japan. Eur. J. Cancer Prev. 2009; 18(3):240–247.

10. Matthews CE, Sui X, Lamonte MJ et al. Metabolic syndrome and risk of death from cancers of the digestive system. Metabolism 2010; 59(8):1231–1239.

11. Johansen D, Stocks T, Jonsson H et al. Metabolic factors and the risk of pancreatic cancer: a prospective analysis of almost 580,000 men and women in the Metabolic Syndrome and Cancer Project. Cancer Epidemiol. Biomarkers Prev. 2010; 19(9):2307–2317.

12. Stocks T, Bjorge T, Ulmer H et al. Metabolic risk score and cancer risk: pooled analysis of seven cohorts. Int. J. Epidemiol. 2015:1–11.

13. Lipworth L, Zucchetto A, Bosetti C et al. Diabetes mellitus, other medical conditions and pancreatic cancer: a case-control study. Diabetes Metab Res Rev 2011; 27(3):255–261.

14. Bosetti C, Lucenteforte E, Bracci PM et al. Ulcer, gastric surgery and pancreatic cancer risk: an analysis from the International Pancreatic Cancer Case-Control Consortium (PanC4). Ann. Oncol. 2013; 24(11):2903–2910.

15. Luo J, Nordenvall C, Nyrén O et al. The risk of pancreatic cancer in patients with gastric or duodenal ulcer disease. Int. J. Cancer 2007; 120(2):368–372.

16. Bao Y, Spiefelman D, Li R et al. History of peptic ulcer diseases and pancreatic cancer risk in men. Gastroenterology 2010; 138(2):541–9.

17. Schulte A, Pandeya N, Fawcett J et al. Association between Helicobacter pylori and pancreatic cancer risk: a meta-analysis. Cancer Causes Control 2015; 26(7):1027–1035.

18. Holly EA, Chaliha I, Bracci PM, Gautam M. Signs and symptoms of pancreatic cancer: a population-based case-control study in the San Francisco Bay area. Clin. Gastroenterol. Hepatol. 2004; 2(6):510–517.

19. Batabyal P, Vander Hoorn S, Christophi C, Nikfarjam M. Association of diabetes mellitus and pancreatic adenocarcinoma: a meta-analysis of 88 studies. Ann. Surg. Oncol. 2014; 21(7):2453–62.

20. Cotterchio M, Lowcock E, Hudson TJ et al. Association between allergies and risk of pancreatic cancer. Cancer Epidemiol. Biomarkers Prev. 2014; 23(3):469–80.

21. Real FX. A “catastrophic hypothesis” for pancreas cancer progression. Gastroenterology 2003; 124(7):1958–1964.

22. Ben Q, Xu M, Ning X et al. Diabetes mellitus and risk of pancreatic cancer: A meta-analysis of cohort studies. Eur. J. Cancer 2011; 47(13):1928–1937.

23. Elena JW, Steplowski E, Yu K et al. Diabetes and risk of pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium. Cancer Causes Control 2013; 24(1):13–25.

24. Pannala R, Leirness JB, Bamlet WR et al. Prevalence and Clinical Profile of Pancreatic Cancer-Associated Diabetes Mellitus. Gastroenterology 2008; 134(4):981–987.

25. Nakamori S, Ishikawa O, Ohigashi H et al. Increased blood proinsulin and decreased C-peptide levels in patients with pancreatic cancer. Hepatogastroenterology 1999; 46(25):16–24.

26. Gerdtsson AS, Malats N, Säll A et al. A Multicenter Trial Defining a Serum Protein Signature Associated with Pancreatic Ductal Adenocarcinoma. Int. J. Proteomics 2015; 2015:587250.

Supplementary Annex. PanGenEU centres and investigators ([1] Principal Investigator in each centre)

Spanish National Cancer Research Centre (CNIO), Madrid, Spain: Núria Malats[1], Francisco X Real[1], Mirari Marquez, Roger Milne, Paulina Gómez, Ana Sagrera.

Verona University, Italy: Aldo Scarpa[1], Rita Lawlor, Giada Bonizzato, Sonia Grimaldi.

National Cancer Registry Ireland, Cork, Ireland: Linda Sharp[1], Damian O'Driscoll.

Hospital Madrid-Norte-Sanchinarro, Madrid, Spain: Manuel Hidalgo[1].

Hospital Ramon y Cajal, Madrid, Spain: Alfredo Carrato[1], Carmen Guillén-Ponce.

Hospital del Mar, Barcelona, Spain: Lucas Ilzarbe[1], Cristina Álvarez-Urturi, Xavier Bessa, Felipe Bory, Lucía Márquez Mosquera, Ignasi Poves Prim, Fernando Burdío, Luis Grande, Mar Iglesias, Javier Gimeno.

Hospital Vall d´Hebron, Barcelona, Spain: Xavier Molero[1], Luisa Guarner, Joaquin Balcells.

Technical University of Munich, Germany: Christoph Michalski[1], Jörg Kleeff, Bo Kong

Karolinska Institute, Stockholm, Sweden: Matthias Löhr[1], Jiaqui Huang.

Hospital 12 de Octubre, Madrid, Spain: José Perea[1], Pablo Peláez.

Hospital de la Santa Creu i Sant Pau, Barcelona, Spain: Antoni Farré[1], Josefina Mora, Marta Martín, Vicenç Artigas, Carlos Guarner Argente, Francesc J Sancho, Mar Concepción, Teresa Ramón y Cajal.

The Royal Liverpool University Hospital, UK: William Greenhalf[1], Eithne Costello.

Queen's University Belfast, UK: Liam Murray[1], Michael O'Rorke, Marie Cantwell.

Laboratorio de Genética Molecular, Hospital General Universitario de Elche, Spain: Víctor M Barberá[1], Javier Gallego.

Instituto Universitario de Oncología del Principado de Asturias, Oviedo, Spain: Adonina Tardón[1], Luis Barneo.

Hospital Clínico Universitario de Santiago de Compostela, Spain: Enrique Domínguez Muñoz[1], Antonio Lozano, Maria Luaces.

Hospital Clínico Universitario de Salamanca, Spain: Luís Muñoz-Bellvís[1], J.M. Sayagués Manzano, M.L. Gutíerrrez Troncoso, A. Orfao de Matos.

University of Marburg, Department of Gastroenterology, Phillips University of Marburg, Germany: Thomas Gress[1], Malte Buchholz, Albrecht Neesse.

Queen Mary University of London, UK: Tatjana Crnogorac-Jurcevic[1], Hemant M Kocher, Satyajit Bhattacharya, Ajit T Abraham, Darren Ennis, Thomas Dowe, Tomasz Radon

Scientific advisors of the PanGenEU Study: Debra T Silverman (NCI, USA) and Douglas Easton (U. of Cambridge, UK)

Figure Legends

Figure 1. Clustering of morbidities in pancreatic ductal adenocarcinoma cases and controls. A) Dendrogram depicting Hierarchical cluster analysis using Yule's Q distance matrix, B) Loadings from principal component analysis, C) Loadings from factor analysis. Loadings $\geq$0.3 are considered to belong to the same component and factor, respectively. Yule's Q distance was calculated as (1 - ((ad-bc)/(ad+bc))), where a indicates presence of both conditions, b and d absence of one condition and presence of the other, and d absence of both conditions. PanGenEU, 2009-2014.

Table 1. Association between multimorbidity patterns and pancreatic ductal adenocarcinoma.

Gastric pattern[a]							
Lifetime gastric conditions							
Number of conditions	**Cases**		**Controls**		**OR (95% CI)**		***p* for trend[e]**
	n=1705	%	*n*=1084	%	Unadjusted	Adjusted[d]	
0	987	57.9	709	65.4	1[Reference]	1[Reference]	
1	319	18.7	182	16.8	1.26 (1.02-1.55)	1.34 (1.07 – 1.68)	
2	292	17.1	135	12.5	1.55 (1.24-1.95)	1.65 (1.29 – 2.10)	
	107	6.3	58	5.4	1.32 (0.95-1.85)	1.46 (1.01 – 2.10)	<0.01
High adherence[f]	615	36.1	315	29.1		1.22 (0.99 – 1.51)	
Recent gastric conditions							
	n=1177	%	*n*=725	%			
0	987	83.9	709	97.8	1[Reference]	1[Reference]	
1	111	9.4	7	1.0	11.4 (5.27-24.6)	12.3 (5.63 – 26.9)	
	79	6.7	9	1.2	6.30 (3.14-12.7)	6.13 (3.01 – 12.5)	<0.01
High adherence[f]	432	36.7	203	28.0		2.17 (1.64 – 2.87)	
Long-term gastric conditions							
	n=1479	%	*n*=1059	%			
0	987	66.7	709	66.9	1[Reference]	1[Reference]	
1	211	14.3	176	16.6	0.86 (0.69-1.07)	0.89 (0.70 – 1.14)	
2	202	13.7	121	11.4	1.19 (0.94-1.53)	1.29 (0.99 – 1.68)	
	79	5.3	53	5.0	1.07 (0.75-1.54)	1.17 (0.79 – 1.73)	0.14
High adherence[f]	491	33	355	33.5		0.79 (0.63 – 0.98)	

				MetS pattern[b]			
				Lifetime T2DM			
	n=1705	%	*n*=1084	%			
	591	34.7	372	34.3	1[Reference]	1[Reference]	
	526	30.9	349	32.2	0.94 (0.78-1.14)	1.10 (0.89 – 1.36)	
	372	21.8	256	23.6	0.91 (0.74-1.12)	1.11 (0.88 – 1.39)	
	216	12.7	107	9.9	1.27 (0.97-1.66)	1.62 (1.22 – 2.17)	<0.01
High adherence[f]	572	33.5	359	33.1		1.24 (1 – 1.53)	
				Recent T2DM			
	n=1458	%	*n*=963	%			
	591	40.5	372	38.6	1[Reference]	1[Reference]	
	466	32.0	322	33.4	0.91 (0.75-1.10)	1.05 (0.85 – 1.30)	
	289	19.8	213	22.1	0.85 (0.69-1.06)	1.05 (0.83 – 1.35)	
	112	7.7	56	5.8	1.26 (0.89-1.78)	1.61 (1.11 – 2.35)	0.06
High adherence[f]	498	34.2	314	32.6		1.37 (1.09 – 1.73)	
				Long-term T2DM			
	n=1523	%	*n*=1060	%			
	591	38.8	372	35.1	1[Reference]	1[Reference]	
	492	32.3	346	32.6	0.89 (0.74-1.08)	1.03 (0.83 – 1.27)	
	287	18.8	248	23.4	0.73 (0.59-0.90)	0.89 (0.70 – 1.13)	
	153	10.0	94	8.9	1.02 (0.77-1.37)	1.35 (0.99 – 1.85)	0.25
High adherence[f]	489	32.1	372	35.1		1 (0.80 – 1.25)	

	Atopic pattern[c]						
	Lifetime atopic conditions						
	n=1705	%	*n*=1084	%			
	1100	64.5	610	56.3	1[Reference]	1[Reference]	
	485	28.4	353	32.6	0.76 (0.64-0.90)	0.77 (0.64 – 0.92)	
	100	5.9	105	9.7	0.53 (0.39-0.71)	0.54 (0.39 – 0.74)	
	20	1.2	16	1.5	0.69 (0.36-1.35)	0.66 (0.32 – 1.37)	<0.01
High adherence[f]	510	29.9	438	40.4		0.45 (0.36 – 0.55)	

Gastric pattern: heartburn, acid regurgitation, *H. pylori* and ulcer.
MetS pattern: obesity, T2DM, hypercholesterolemia and hypertension.
Atopic pattern: skin allergies, nasal allergies and asthma. Time since diagnosis was not available for nasal and skin allergies.
Adjusted for age (continuous), sex, country, smoking (tertiles of pack/years), and number of morbidities excluding those tested in each model.
value for trend was calculated using variables with 4 categories (0,1,2, ≥3) as continuous in the model.
Tertiles of factor scores on the overall group of cases and controls. High adherence (tertile III) compared to low adherence (tertile I).
Abbreviation: *H. pylori, Helicobacter pylori*; MetS, metabolic syndrome; T2DM, type 2 diabetes mellitus.

Table 2. Association between the combination of T2DM and gastric conditions with pancreatic ductal adenocarcinoma.

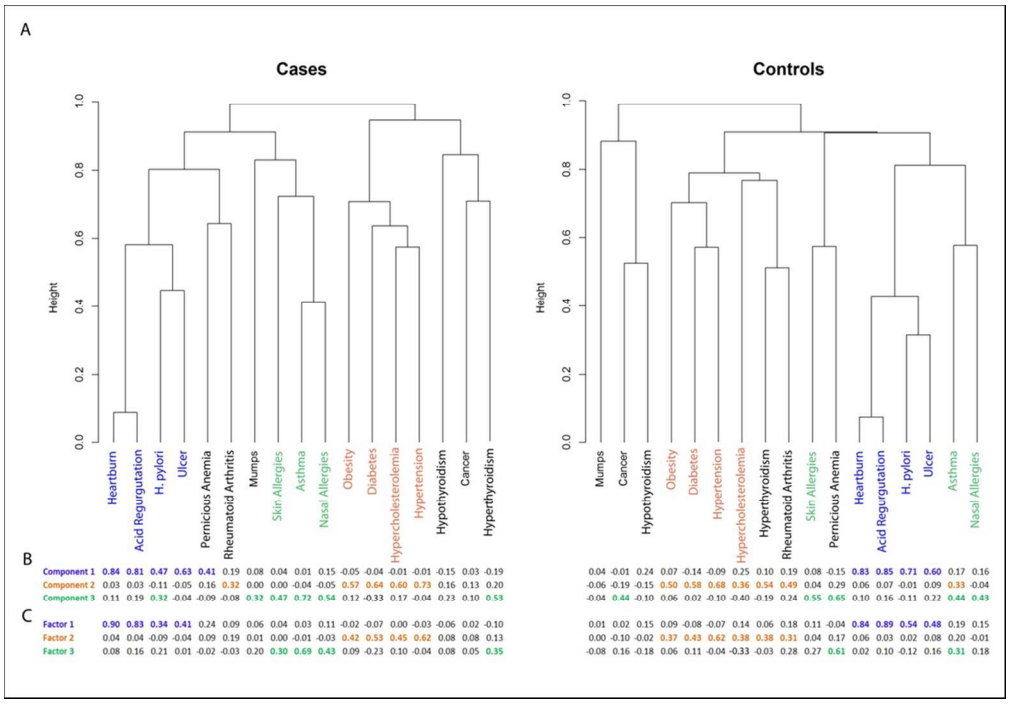
T2DM + Gastric conditions[a]							
Lifetime conditions							
Number of conditions	**Cases**		**Controls**		**OR (95% CI)**		***p* for trend[c]**
	n=1705	%	*n*=1084	%	**Unadjusted**	**Adjusted[b]**	
0	730	42.8	610	56.3	1[Reference]	1[Reference]	
1	482	28.3	256	23.6	1.57 (1.31-1.89)	1.76 (1.43 – 2.16)	
2	322	18.9	150	13.8	1.79 (1.44-2.24)	1.98 (1.55 – 2.51)	
	171	10.0	68	6.3	2.10 (1.56-2.84)	2.58 (1.87 – 3.57)	<0.01
Recent conditions							
	n=1009	%	*n*=640	%			
0	730	72.3	610	95.3	1[Reference]	1[Reference]	
1	193	19.1	21	3.3	7.68 (4.83-12.2)	7.80 (4.84 – 12.6)	
	86	8.5	9	1.4	7.98 (3.98-15.9)	7.89 (3.9 – 16.1)	<0.01
Long-term conditions							
	n=1318	%	*n*=1037	%			
0	730	55.4	610	58.8	1[Reference]	1[Reference]	
1	292	22.2	235	22.7	1.04 (0.85-1.27)	1.18 (0.94 – 1.48)	
2	190	14.4	134	12.9	1.18 (0.92-1.52)	1.29 (0.99 – 1.69)	
	106	8.0	58	5.6	1.53 (1.09-2.14)	1.86 (1.29 – 2.67)	<0.01

Heartburn, acid regurgitation, *H. pylori*, and ulcer.

Adjusted for age (continuous), sex, country, smoking (tertiles of pack/years), and number of morbidities excluding those tested in each model.

value for trend was calculated using variables with 4 categories (0,1,2, $\geq$3) as continuous in the model.

Abbreviation: *H. pylori, Helicobacter pylori*; T2DM, type 2 diabetes mellitus.



Clustering of morbidities in pancreatic ductal adenocarcinoma cases and controls. A) Dendrogram depicting Hierarchical cluster analysis using Yule's Q distance matrix, B) Loadings from principal component analysis, C) Loadings from factor analysis. Loadings >0.3 are considered to belong to the same component and factor, respectively. Yule's Q distance was calculated as (1 - ((ad-bc)/(ad+bc))), where a indicates presence of both conditions, b and d absence of one condition and presence of the other, and d absence of both conditions. PanGenEU, 2009-2014.

202x140mm (150 x 150 DPI)

Supplementary Methods

Imputation

Imputation of missing values (2.4% in cases and 2.6% in controls) was performed with missForest R package. Variables for imputation (% missings) included case-control status, country, age (2.3%), sex (0.2%), smoking (pack/years, 9.4%), and morbidities and time since their diagnosis (0.8%-21.7%). Imputation was performed without maximum number of iteration and 100 trees. Random attribution of missing values to the complete dataset mimicking the missing proportions in the original dataset resulted in an averaged correct assignment of 93.7% after imputation. Out of bag (OOB) error for imputation was $\leq$0.31, mean: 0.04, SD=0.09 (0= good performance, 1= bad performance). Frequencies between imputed and complete-data were not significantly different (*p* value $\geq$0.38).

Comorbidity and trimorbidity models

All combinations of two and three medical conditions within an individual were considered: 136 comorbidities and 680 trimorbidities (62 excluded due to cells with zero counts). Odds ratios (ORs) for each comorbidity and trimorbidity were calculated and visualized by a heatmap using plotrix package [Lemon 2006]. *p* values for co and trimorbidity analyses were adjusted for multiple testing with Bonferroni's method where a Bonferroni corrected p-value of 0.05 equals a p-value before correction of 3x10-3 in comorbidities and a p-value of 8x10-5 in trimorbidities [Bland 1995].. Robustness was assessed with bootstrapped samples (*n*=1,000) drawn with replacement from the original dataset preserving the sample size. The selection probability proportion (SPP) was calculated as the proportion of models that remained significant for each variable after multiple test correction ((number of significant models*100)/1000) [Pineda 2014].

Multimorbidity patterns

Multimorbidity patterns (MPs) were defined separately for cases and controls using hierarchical cluster analysis, principal component (PCA), and factor (FA) analyses. Hierarchical agglomerative cluster analysis was conducted with stats and proxy R packages using the average linkage method and Yule's Q distance (1-((ad-bc)/(ad+bc))) matrix for all possible combinations of the 17 morbidities. PCA and exploratory FA with a principal factor solution were carried out using a tetrachoric correlation

matrix. Scree plots and eigenvalues greater than 1 from both cases and controls were used to determine the number of components. An orthogonal (varimax) rotation of the loading matrix was applied to obtain a simpler structure and improve interpretability while providing uncorrelated factors. The magnitude of each loading indicates the relevance of the correspoinding medical condition to the component/factor. MPs were determined according to those conditions with loadings $\geq$0.30 in their respective components and factors [Laher 2010]. The same number of components/factors (*n*=3) was retained for cases and controls and the three methods. Moreover, the biological rationale of the MPs was considered for their selection.

As internal validation, estimates were calculated using new samples generated by bootstrapping with replacement (*n*=10,000) while preserving the original sample size. Receiver operating characteristic (ROC) curves and the corresponding area under the curve (AUC) were calculated for the new samples generated with pROC R package. Mean and standard deviation (SD) were calculated for the resulting AUCs.

Multimorbidity system analysis

DisGeNET version 3.0, a bioinformatics-tool that correlates human diseases and their associated genes from information available in UniProt, CTD, ClinVar databases [Piñero 2015], was used to obtain information about genes associated with the MPs and PDAC. The Jaccard index (JI) was calculated to assess disease similarity based in common genes while accounting for variation in gene findings, i.e. diseases with larger versus smaller number of genes. The JI is defined as: $|Genes_{dis1} \cap Genes_{dis2}| / |Genes_{dis1} \cup Genes_{dis2}|$, where $Genes_{dis1}$ and $Genes_{dis2}$ are the genes associated with disease 1 and 2, respectively, $\cap$ is the intersection operator, and $\cup$ is the union operator between the two sets of genes. The disease system figure was created using the DisGeNET plugin for Cytoscape 2.8.3.

Additive interactions for multimorbidity patterns

The additive interaction between morbidities within multimorbidity patterns was evaluated using the Relative Excess Risk due to Interaction (RERI) with the EpiR package in R. Variables were created that included all or none of three morbidities of a pattern (or two en the case of the atopic pattern) and compared with the presence (yes) or absence (no) of the fourth (or third) morbidity of a pattern; all

possible combinations were tested. For the pattern of diabetes and gastric comorbidities the additive interaction was tested between each morbidity and the combination of any three or more morbidities (we were unable to test the full model due to low sample size).

References

Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995; 310(6973):170.

Laher S. Using exploratory factor analysis in personality research: Best-practice recommendations. J. Ind. Psychol. 2010; 36(1):1–7.

Lemon J. Plotrix: a package in the red light district of R. *R-News*. 2006;6(4):8-12.

Pineda S, Milne RL, Calle ML, et al. Genetic variation in the TP53 pathway and bladder cancer risk. A comprehensive analysis. *PLoS One*. 2014;9(5):1-8.

Piñero J, Queralt-Rosinach N, Bravo Á et al. DisGeNET: A discovery platform for the dynamical exploration of human diseases and their genes. Database 2015; 2015:1–17.

Supplementary Figure Legends

Supplementary Figure S1. Heatmap of the associations between comorbidities in pancreatic ductal adenocarcinoma cases and controls. White cells represent an OR=1, shades of blue represent ORs below 1, and shades of red represent ORs above 1. ORs are presented for the association between the condition in the left and the condition in the bottom of each square. All models were adjusted for age, sex, country, smoking (pack/years) and number of morbidities. The asterisks represent statistical significance after Bonferroni´s correction. PanGenEU, 2009-2014.

Supplementary Figure S2. Genetic system for multimorbidity patterns and pancreatic cancer. Association between pancreatic cancer and medical conditions of the detected multimorbidity patterns based on the number of shared genes. Edge width represents the Jaccard index for each disease pair; Jaccard indexes were multiplied by 100 for better visualization. Node size represents the number of genes obtained through DisGeNET for each medical condition.

Supplementary Figure S3. Venn diagrams showing the number of shared genes between each multimorbidity pattern and pancreatic cancer. A) Overlapping genes for atopic pattern conditions and pancreatic cancer; B) Overlapping genes for gastric pattern conditions and pancreatic cancer; C) Overlapping genes for metabolic syndrome pattern conditions and pancreatic cancer. The squares show the genes shared between different medical conditions and pancreatic cancer.

Supplementary Material

Supplementary Table S1. Diagnosed conditions at admission for hospital controls. PanGenEU, 2009-2014.

Disease Group	Conditions
Diseases of the circulatory system	Hemorrhoids
Diseases of the digestive system	Appendicitis, appendiceal abscess, inguinal hernia and other abdominal hernias, fissure and fistula of anal and rectal regions, anal prolapse, rectal prolapse, stenosis of anus and rectum
Diseases of the skin and subcutaneous tissue	Cutaneous abscess, furuncle, cellulitis, pilonidal cyst
Diseases of the genitourinary system	Varicocele, spermatocele, congenital spermatocele, hydrocele and spermatocele, torsion of testis. Exclude: congenital hydrocele
Injury, poisoning and other consequences of external causes	Fractures including prosthesis due to fractures, dislocations, sprains, internal injuries of thorax, abdomen and pelvis. Open wounds of the superior and inferior extremities. Other diseases of external causes, burns, plastic surgery (due to burns, accidents)
Diseases of the musculoskeletal system and connective tissue	Acquired deformities of toe (exclude congenital deformations), change of orthopedic prosthesis due infection, sacral cyst
Diseases of the eye and adnexa and diseases of the ear and mastoid process	Cornea transplant, retinal detachments and breaks, cholesteatoma (cleaning of), tympanoplasty due perforation of tympanic membrane.
Operations on the nose, mouth, and pharynx	Nasal septoplasty, turbinectomy, uvulopalatoplasty
Diseases of the respiratory system	Spontaneous pneumothorax

Supplementary Table S2. Distribution of cases and controls according to selected variables. PanGenEU, 2009-2014.

	Cases		Controls	
	n=1705	%	n=1084	%
Country				
Spain	842	49.4	595	54.9
England	121	7.1	22	2.0
Germany	140	8.2	110	10.1
Ireland	173	10.1	290	26.8
Italy	292	17.1	0	
Sweden	137	8.0	67	6.2
Sex				
Female	741	43.5	518	47.8
Male	964	56.5	566	52.2
Age				
< 54	363	21.3	221	20.4
55 - 64	413	24.2	230	21.2
65 - 74	592	34.7	333	30.7
> 75	337	19.8	300	27.7
Smoking (pack/years)				
Never	673	39.5	540	49.8
< 18.5	315	18.5	233	21.5
18.6 - 40.9	365	21.4	145	13.4
> 41	352	20.6	166	15.3
Number of morbidities				
0	117	6.9	76	7.0
1	267	15.7	171	15.8
2	348	20.4	215	19.8
3	328	19.2	207	19.1
4	292	17.1	194	17.9
5	177	10.4	117	10.8
>6	176	10.3	104	9.6

Supplemetary Table S3. Frequency of medical conditions in cases and controls, and corresponding odds rations and 95% condifence intervals. PanGenEU, 2009-2014.

	Cases		Controls		OR[a] (95%CI)
	n=1705	%	n=1084	%	
Rheumatoid arthritis					
No	1597	93.7	1000	92.3	1 [Reference]
Yes	108	6.3	84	7.7	0.75 (0.54-1.03)
Pernicious anemia					
No	1630	95.6	1014	93.5	1 [Reference]
Yes	75	4.4	70	6.5	0.72 (0.50-1.02)
Hyperthyroidism					
No	1648	96.7	1041	96.0	1 [Reference]
Yes	57	3.3	43	4.0	0.83 (0.54-1.28)
Hypothyroidism					
No	1612	94.5	1023	94.4	1 [Reference]
Yes	93	5.5	61	5.6	1.07 (0.74-1.54)
Mumps					
No	1023	60.0	616	56.8	1 [Reference]
Yes	682	40.0	468	43.2	0. 83 (0.69-0.99)
Cancer					
No	1502	88.1	953	87.9	1 [Reference]
Yes	203	11.9	131	12.1	1.03 (0.80-1.33)
Diabetes					
No	1276	74.8	939	86.6	1 [Reference]
Yes	429	25.2	145	13.4	2.25 (1.79 - 2.81)
<3y	182	10.7	24	2.2	5.27 (3.38 - 8.23)
≥3y	247	14.5	121	11.2	1.61 (1.25 - 2.07)
Hypertension					
No	1070	62.8	659	60.8	1 [Reference]
Yes	635	37.2	425	39.2	0.94 (0.79 - 1.12)
<3y	90	5.3	50	4.6	0.96 (0.65 - 1.42)
≥3y	545	32.0	375	34.6	0.93 (0.78 - 1.12)
Hypercholesterolemia					
No	1167	68.4	691	63.7	1 [Reference]
Yes	538	31.6	393	36.3	1.05 (0.88 - 1.25)
<3y	106	6.2	69	6.4	1.25 (0.89 - 1.75)
≥3y	432	25.3	324	29.9	1 (0.83 - 1.21)
Obesity[b]					
No	1350	79.2	855	78.9	1 [Reference]
Yes	355	20.8	229	21.1	1.05 (0.86 - 1.29)
Helicobacter pylori					
No	1556	91.3	1016	93.7	1 [Reference]
Yes	149	8.7	67	6.3	1.31 (0.94-1.82)
<3y	42	2.5	6	0.6	3.64 (1.49 - 8.94)
≥3y	107	6.3	61	5.6	1.08 (0.75 - 1.56)
Acid Regurgitation					
No	1293	75.8	888	81.9	1 [Reference]
Yes	412	24.2	196	18.1	1.59 (1.29 - 1.96)
<3y	126	7.4	18	1.7	5.02 (2.99 - 8.42)
≥3y	286	16.8	178	16.4	1.22 (0.97 - 1.53)

Supplementary Table S3. Frequency of medical conditions in cases and controls and corresponding odds ratios, and 95% confidence intervals. PanGenEU, 2009-2014. (continued)

	Cases		Controls		OR[a] (95%CI)
	n=1705	%	n=1084	%	
Heartburn					
No	1217	71.4	829	76.5	1 [Reference]
Yes	488	28.6	255	23.5	1.42(1.17 – 1.73)
<3y	130	7.6	11	1.0	8.23 (4.36 – 15.3)
≥3y	358	21.0	244	22.5	1.09 (0.88 – 1.34)
Ulcer					
No	1516	88.9	966	89.1	1 [Reference]
Yes	186	11.1	118	10.9	1.03 (0.78 – 1.34)
<3y	24	1.4	4	0.4	2.89 (0.97 – 8.61)
≥3y	162	9.5	114	10.5	0.91 (0.69 – 1.20)
Asthma					
No	1580	92.7	969	89.4	1 [Reference]
Yes	125	7.3	115	10.6	0.61 (0.45 – 0.82)
<3y	8	0.5	4	0.4	1.75 (0.49 – 6.2)
≥3y	117	6.9	111	10.2	0.58 (0.42 – 0.78)
Nasal allergies[c]					
No	1465	85.9	865	79.8	1 [Reference]
Yes	240	14.1	219	20.2	0.66 (0.53 – 0.83)
Skin allergies[c]					
No	1325	77.7	807	74.4	1[Reference]
Yes	380	22.3	277	25.6	0.84 (0.68 – 1.03)
Chronic pancreatitis					
No	1689	99.1	1083	99.9	1 [Reference]
Yes	16	0.9	1	0.1	6.01 (0.74 – 4.85)
Lupus					
No	1696	99.5	1081	99.7	1 [Reference]
Yes	9	0.5	3	0.3	1.88 (0.47-7.51)
Scleroderma					
No	1697	99.5	1080	99.6	1 [Reference]
Yes	8	0.5	4	0.4	0.79 (0.2 – 3.12)
Polymyalgia					
No	1692	99.2	1065	98.2	1 [Reference]
Yes	13	0.8	19	1.8	0.38 (0.17-0.85)
Crohn's disease					
No	1701	99.8	1079	99.5	1 [Reference]
Yes	4	0.2	5	0.5	0.27 (0.46-1.62)
Ulcerative colitis					
No	1690	99.1	1068	98.5	1 [Reference]
Yes	15	0.9	16	1.5	0.48 (0.20-1.15)
Celiac disease					
No	1698	99.6	1078	99.4	1 [Reference]
Yes	7	0.4	6	0.6	0.59 (0.17-2.11)
Addison's disease					
No	1698	99.6	1078	99.4	1 [Reference]
Yes	7	0.4	6	0.6	0.87 (0.28-2.68)

Adjusted for age, sex, country, smoking (pack/years), and number of morbidities.
Calculated using weight at two years before recruitment.
Information on time of diagnosis of the disease was not collected.

Supplemetary Table S4. Co and trumorbidities significantly associated with pancreatic ductal Adenocarcinoma after multiple test correction. PanGenEU, 2009-2014

	Cases/Controls[a]	OR[b] (95%CI)	*Corrected p*-value[c]	SPP
Comorbidities				
Diabetes-Acid Regurgitation	91/23	4.25 (2.55-7.08)	<0.01	99.9
Diabetes-Heartburn	113/31	3.94 (2.49- 6.2)	<0.01	100
Diabetes-Obesity	131/39	2.39 (2.49-6.21)	<0.01	95.9
Diabetes-Hypercholesterolemia	184/65	2.36 (1.69-3.32)	<0.01	99.4
Diabetes-Hypertension	226/85	1.9 (1.38-2.59)	0.01	94.3
Heartburn-Acid Regurgitation	317/152	1.62 (1.24-2.10)	0.04	84.1
Nasal allergies-Hypertension	79/86	0.48 (0.32-0.70)	0.02	88.5
Nasal allergies- Skin allergies	75/78	0.47 (0.32-0.69)	0.02	91.0
Skin allergies-Pernicious anemia	19/31	0.32 (0.17-0.59)	0.04	87.0
Trimorbidities				
Diabetes-Hypercholesterolemia-Acid regurgitation	41/11	4.81 (2.28-10.12)	0.02	90.7
Diabetes-Hypercholesterolemia- Heartburn	48/14	4.54 (2.31-8.96)	<0.01	94.0
Diabetes-Heartburn-Acid regurgitation	62/17	4.40 (2.39-8.08)	<0.01	96.7
Diabetes-Hypertension-Acid regurgitation	52/14	4.29 (2.18-8.43)	0.01	92.2

Subjects that report having all medical conditions

Adjusted for age, sex, country, smoking (pack/years), and number of morbidities. Reference category: none of the two or three medical conditions

Bonferroni corrected p-value. A Bonferroni corrected p-value of 0.05 equals a p-value before correction of 3x10-3 in comorbidities and of 8x10-5 in trimorbidities.

SPP: selection probability proportion

Supplementary Table S5. Association between metabolic syndrome pattern and pancreatic ductal adenocarcinoma stratified by sex. PanGenEU, 2009-2014.

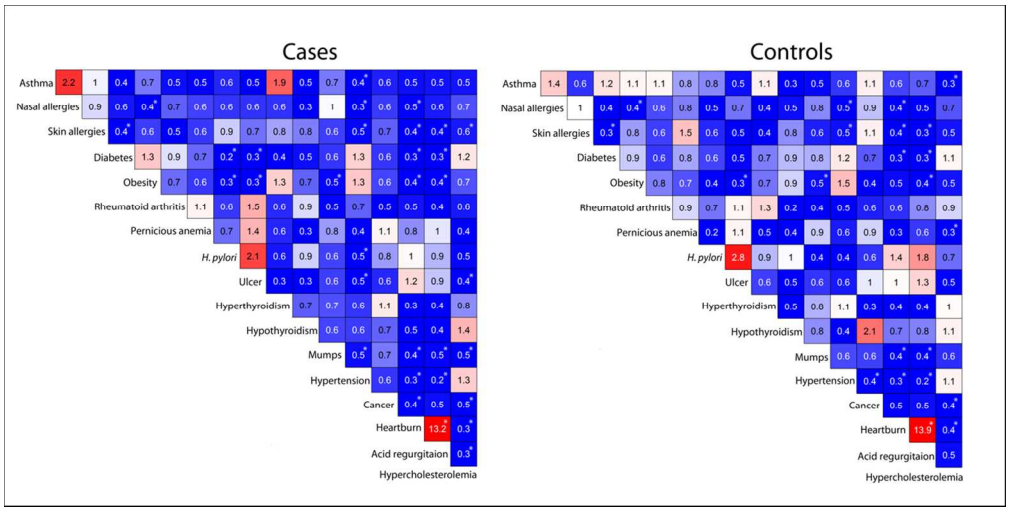
	Females					Males				
	Cases		Controls			Cases		Controls		
Number of conditions	n=741	%	n=518	%	OR[a] (95%CI)	n=964	%	n=566	%	OR[a] (95%CI)
MetS pattern										
0	267	36.0	162	31.3	1 [Reference]	324	33.6	210	37.1	1 [Reference]
1	237	32.0	162	31.3	1.09 (0.79-1.51)	289	30.0	187	33.0	1.08 (0.82-1.43)
2	155	20.9	139	26.8	0.92 (0.65-1.29)	217	22.5	117	20.7	1.31 (0.96-1.78)
	82	11.1	55	10.6	1.27 (0.83-1.98)	134	13.9	52	9.2	1.93 (1.31-2.85)

MetS: Metabolic Syndrome (i.e. type 2 diabetes, hypertension, hypercholesterolemia, and obesity)
Adjusted for age, country, smoking (pack/years), and number of morbidities
value of interaction between MetS and sex: 0.032

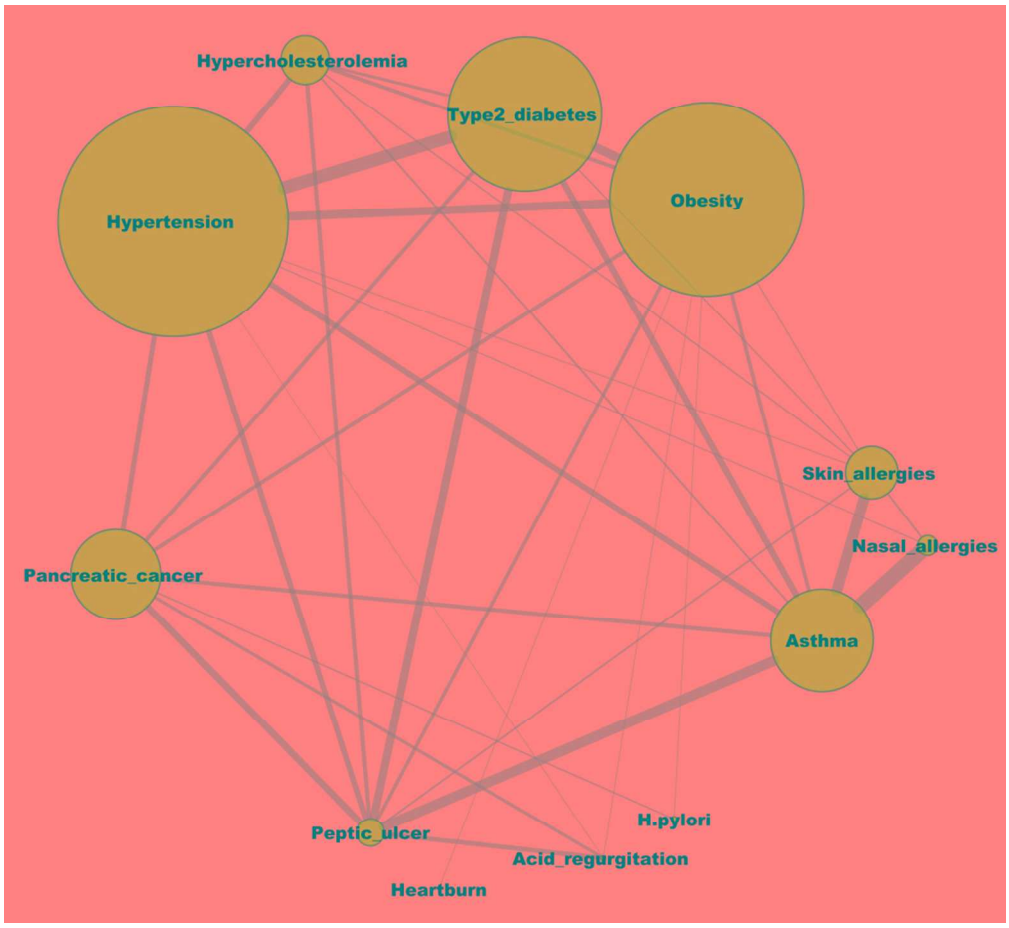
Supplementary Table S6. Association between lifetime multimorbidity patterns and pancreatic ductal adenocarcinoma further adjusting for obesity. PanGenEU, 2009-2014.

Number of conditions	Cases n=1705	%	Controls n=1084	%	OR*	95%CI
Gastric pattern						
	987	57.9	709	65.4	1	Ref
	319	18.7	182	16.8	1.35	1.08-1.69
	292	17.1	135	12.5	1.65	1.29-2.10
	107	6.3	58	5.4	1.47	1.02-2.11
MetS pattern (excluding obesity)						
	684	40.1	428	39.5	1	Ref
	549	32.2	391	36.1	0.98	0.81-1.19
	363	21.3	223	20.6	1.27	1.01-1.59
	109	6.4	42	3.9	1.97	1.97-2.94
Atopic pattern						
	1100	64.5	610	56.3	1	Ref
	485	28.4	353	32.6	0.77	0.64-0.92
	100	5.9	105	9.7	0.54	0.39-0.75
	20	1.2	16	1.5	0.66	0.32-1.36
T2DM + Gastric conditions						
	730	42.8	610	56.3	1	Ref
	482	28.3	256	23.6	1.76	1.43-2.16
	322	18.9	150	13.8	1.98	1.56-2.52
	171	10	68	6.3	2.59	1.87-3.59

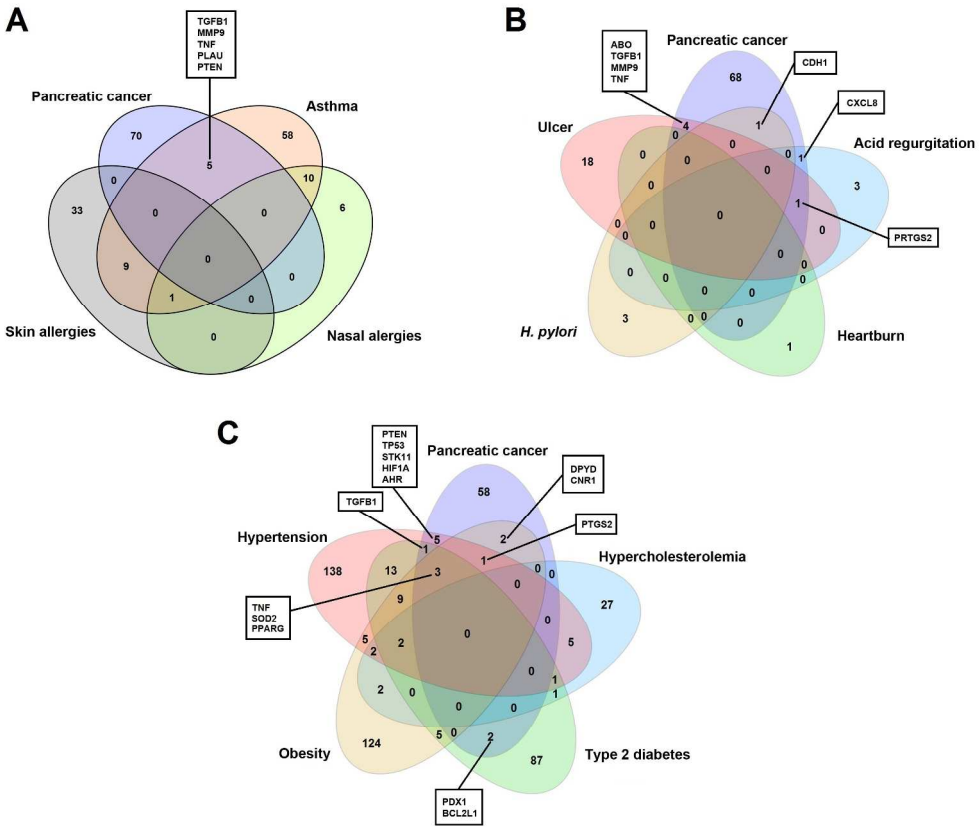
* Adjusted for age (continuous), sex, country, smoking (tertiles of pack/years), obesity (BMI <30 kg/m^2 / BMI $\geq$30 kg/m^2) and number of morbidities excluding those tested and adjusted in each model.



247x124mm (150 x 150 DPI)



398x364mm (72 x 72 DPI)



750x631mm (100 x 100 DPI)