

STUDY PROTOCOL

Open Access



The TEOGIC study project: a comprehensive characterization of early onset gastrointestinal cancer in the Northern area of Spain

R. Vera^{1,2}, N. Castro¹, I. Labiano¹, A. Lecumberri^{1,2}, A. E. Huerta¹, H. Arasanz^{1,2}, I. Caseda¹, F. Ruiz-Pace³, C. Viaplana³, V. Arrazubi^{1,2}, I. Hernandez-Garcia^{1,2}, E. Mata^{1,2}, D. Gomez^{1,2}, S. Laguna², J. Suarez⁴, I. Fernandez-De-los-Reyes^{5,6}, M. Rullan^{7,8}, F. Estremera⁷, V. Alonso⁹, R. Pazo-Cid⁹, A. Gil-Negrete¹⁰, A. Lafuente¹¹, A. Martin-Carnicero¹¹, R. Dientsmann^{3,12} and M. Alsina^{1,2*}

Abstract

Background Gastrointestinal cancers represent one of the most prevalent diseases worldwide. Strikingly, the incidence of Early Onset Gastrointestinal Cancer (EOGIC) has been rising during the last decades and changes in lifestyle and environmental exposure seem to play a role. EOGIC has been defined as a different entity compared to on-average gastrointestinal cancer, with distinct clinical and molecular characteristics. Inherent to the particularities of younger age, there is an unmet need for a tailored approach for the management of these patients. The TEOGIC proposes a comprehensive study to characterize EOGIC patients in the northern of Spain.

Methods Patients with histologically confirmed new diagnosis of colorectal, gastroesophageal and pancreatic adenocarcinoma will be considered for two cohorts: EOGIC (≤ 50 years old) and non-EOGIC (60–75 years old), with a ratio of 1:2. Two hundred and forty patients will be recruited in 4 Public Hospitals from northern Spain. After receiving unified informed consent, demographic and clinical data of the patients will be collected in a REDCap database. Lifestyle related data will be obtained in questionnaires assessing diet, physical activity and the general quality of life of the patients before diagnosis. Biological samples prior to any onco-specific treatment will be obtained for the analyses of circulating inflammatory proteins, gut microbiota, and the proteome of the tumor microenvironment. Histologic characteristics and routine biomarkers will be also collected. Thereafter, data will be integrated and analyzed to assess tumor specific, pan-tumor and sex-associated differential characteristics of EOGIC.

Discussion The underlying risk factors and differential characteristics of EOGIC remain poorly studied, particularly in our geographical area. Although limited by the exploratory nature and the small sample size estimated to be recruited, TEOGIC represents the first attempt to comprehensively characterize these young patients, and thus attend to their special needs. Findings derived from this study could contribute to raise awareness and preventive behaviors in the population. In parallel, molecular studies could lead to the identification of potential novel non-invasive biomarkers and therapeutic targets that would help in the development of the tailored clinical management of these patients, focusing on screening programs for early diagnosis and precision medicine.

*Correspondence:

M. Alsina

maria.alsina.maqueda@navarra.es

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Keywords Gastrointestinal cancer, Colorectal cancer, Gastroesophageal cancer, Pancreatic cancer, Early onset gastrointestinal cancer, Environmental risk factors, Microbiome, Tumor microenvironment, Precision medicine

Background

Gastrointestinal (GI) cancer includes malignant tumors of the GI tract such as colon, rectum, esophagus, stomach, pancreas, liver and biliary tract. In 2020, colorectal was the cancer with the third highest incidence and the second cause of cancer-related deaths worldwide. The global incidence and prevalence of other GI cancers was also relevant that year, in this regard, both gastroesophageal and pancreatic cancer represented 11.7% of all new cancer diagnoses and 17.9% of all cancer-related deaths [1]. In Spain, estimated incidences by 2024 ranked colorectal, pancreatic and gastroesophageal as the first, seventh and ninth most incident cancers, respectively [2]. Even though a global decreasing incidence of colorectal and gastroesophageal cancer is observed, some studies show an increase in Early-Onset GastroIntestinal Cancer (EOGIC), defined as that diagnosed in the population below 50 years old. Considering pancreatic cancer, both on average-onset and early-onset incidences are gradually growing [3–6]. Specific data regarding cancer incidence among young adults are still lacking in our country.

Several studies have assessed the occurrence of germline mutations in patients with EOGIC [7, 8], however only a small fraction is linked to a hereditary cancer predisposition syndrome, since most of EOGIC cases are sporadic. Therefore, the increase in EOGIC incidence in the last years seems to be influenced by environmental factors such as changes in diet and lifestyle. Obesity, sedentary habits and some dietary patterns are well-known cancer risk factors and are prevalent among young adults developing GI tumors [5, 9–12]. The increased incidence in developed countries of early onset colorectal cancer (EOCRC) or gastroesophageal cancer (EOGEC) also support the hypothesis of environmental factors playing a role in this trend. In this regard, alterations in the gut microbiota due to antibiotic consumption, eradication of *H. pylori* infections or autoimmune atrophic gastritis have been associated to EOCRC and EOGEC [3, 13–16].

Some studies have assessed differential clinical and molecular characteristics of EOGIC, proposing the recognition of a novel entity with special requirements for diagnosis and clinical management. EOGIC usually presents in advanced clinical stages and with aggressive phenotypes, including poor differentiation, signet ring histology and vascular-perineural invasion [17]. EOCRC typically occurs in the distal colon and rectum, with a higher proportion of microsatellite unstable tumors, fewer cases of mutations in *APC*, *KRAS*, and *BRAF*^{V600E},

and mostly with a molecular subtype related to infiltration of immune cells [18–20]. Interestingly, it has been linked to a distinctive microbiota [21]. EOGEC commonly associates with atrophic gastritis from autoimmune phenomena. It presents higher tumor grades, higher proportion of poorly cohesive histology and signet-ring cell subtype, with an increment of *CDH1* mutations, as well as the molecular subtype EBV (TCGA classification) [20, 22]. Early onset pancreatic cancer has lower proportion of *KRAS*, while *SMAD4* mutations are more frequent. Different signaling pathways like Hedgehog, TGF β , hypoxia or PI3K/AKT have also been reported in these patients [23, 24].

Additionally, EOGIC patients have distinctive characteristics that affect their clinical management. Due to their extended life expectancy, they face long-term effects from therapeutic strategies, which may impact their quality of life and more specifically their sexual and reproductive health. Furthermore, cancer diagnosis at a young age normally represents a more pronounced psychological disorder derived from its impact on professional, social and familiar commitments. This is why it is essential to treat these patients individually, considering all these factors and providing a more personalized and precise approach [6, 25–30].

In summary, available evidence suggests that environmental factors underlie most of the increased incidence of EOGIC, although with geographical differences [3]. Available studies also suggest that EOGIC represents a distinct entity with differential molecular signatures that affect the tumor physiopathology and patients' outcomes. Besides, young patients need a distinct approach considering their longer life expectancy and unique familiar and social commitments. In Spain, comprehensive studies focused on the incidence and characterization of this population are still lacking. In a retrospective study analyzing locally advanced rectal cancer recently published by our group, we observed that patients diagnosed at a young age (≤ 50 years old) showed higher rates of response to neoadjuvant therapy compared to older patients, which suggests that these patients could benefit from a tailored clinical management [31].

TEOGIC study methods

Study design and aim

TEOGIC is an observational, transversal and multicenter study with the objective of characterizing EOGIC (i.e. diagnosed in patients ≤ 50 years old) in comparison to

non EOGIC from a comprehensive approach, covering aspects from tumor biology to patients' lifestyle.

Project investigators comprise a multidisciplinary team involving Public Hospitals, a Biomedical Research Center, and a research group specialized in oncology data science. The project is led by the Health Research Institute of Navarra (IdiSNA), including the University Hospital of Navarra (HUN)-Pamplona-Iruña and the Biomedical Research Center Navarrabiomed. Three more Hospitals of the Northern area of Spain participate in patient recruitment: Miguel Servet Hospital (HMS)-Zaragoza, University Hospital Donostia (HUD)-San Sebastián-Donostia and San Pedro Hospital (HSP) Logroño. ODysSey, from the Vall d'Hebron Institute of Oncology (VHIO)-Barcelona will be in charge of the management, integration and interpretation of multiomics and clinical data.

Patient selection and recruitment

Patients will be selected and recruited in routine medical appointments by specialized medical oncologists at the participating Hospitals during the period of 01/2024–12/2025. Patients will be included consecutively based on inclusion and exclusion criteria, following a ratio of 1 EOGIC: 2 non-EOGIC.

Inclusion criteria

- Patients with newly diagnosed, histologically confirmed colorectal adenocarcinoma (CRC, group A), gastroesophageal adenocarcinoma (GEC, group B) and pancreatic adenocarcinoma (PC, group C).
- Age 18–50 (both included) years old for the EOGIC group, and 60–75 (both included) years old for the non-EOGIC group.

Exclusion criteria

- Eastern Cooperative Oncology Group (ECOG) performance status of ≥ 3 .
- Presence of comorbidities which may compromise participation of the patient in the study (mainly in terms of capability to complete the questionnaires).
- Having received oncological treatment directed at the newly diagnosed cancer (i.e., at the time of inclusion the patient has not yet started any type of treatment, including surgery).

Although there is a lack of consensus in the age threshold to define EOGIC, it is generally set at 50 years old, in agreement with the initiation of most of the screening programs [6]. Regarding the non-EOGIC group, we set the lower age at 60 years old, in order to ensure a clear

differentiation between groups; while the upper limit is established at 75 years old in order to exclude the very elderly population. Again, although the lack of consensus in the age to define elderly patients, several studies in the field consider 75 years old as the lower limit to define this population, and present evidence on their differential characteristics compared to on-average patients [32–36]. Age range is also defined with the objective to minimize age related confounding effect.

We estimate the following number of patients to be included in each group: CRC-group A, 40 EOGIC and 80 non-EOGIC; GEC- group B, 20 EOGIC, 40 non-EOGIC; PC-group C, 20 EOGIC, 40 non-EOGIC (Fig. 1a). Estimated number of patients has been calculated based on incidence of EOGIC in HUN in the years 2019–2022 (Fig. 1b), and assuming a similar incidence in the other participating hospitals, taking into account that all participating Hospitals cover a similar target population. The rate of patient inclusion will be evaluated periodically and if necessary, recruitment period will be extended or additional Hospitals of the geographical area will be invited to participate in the study in order to increase patient recruitment.

As aforementioned, patients will be included in a ratio of 1 EOGIC: 2 non-EOGIC with the aim of increasing sample-size in the control group, thus optimizing statistical analyses and ensuring study viability. Assuming a statistical power of 80% and a significance level of 5%, this sample size allows for the detection of statistically significant differences between the EOGIC and the non-EOGIC group in an exploratory approach. For continuous variables at a Cohen's *d* of 0.385-medium size effect; and for categorical variables, with variables at a 10% prevalence in the non-EOGIC group, a 15% of size-difference between groups; with variables at a 30–70% prevalence in the non-EOGIC group, a 20% difference between groups. Additionally, semi-supervised automatic learning will be employed for pan-tumor comparisons, a sample size of 60 EOGIC and 140 non-EOGIC patients is necessary to detect differences in an interest cluster of 20%.

This study is in agreement with international, national and local regulations for clinical studies (latest version of the Declaration of Helsinki, The Spanish Law for Biomedical Research and local regulations "Normas de Buena Práctica Clínica y Orden Foral 125/2009) as well as personal data protection (General Data Protection Regulation (RGPD-Regulation (EU) 2016/679), Spanish Organic Law 3/2018 and the local regulation "Resolution 1387–2017 of the Managing Director of the SNS-O"). Ethics Committee for Investigation with medicinal products from Navarra, have approved this project and a unified informed consent and patient-information sheet has been developed following guidelines of the Spanish

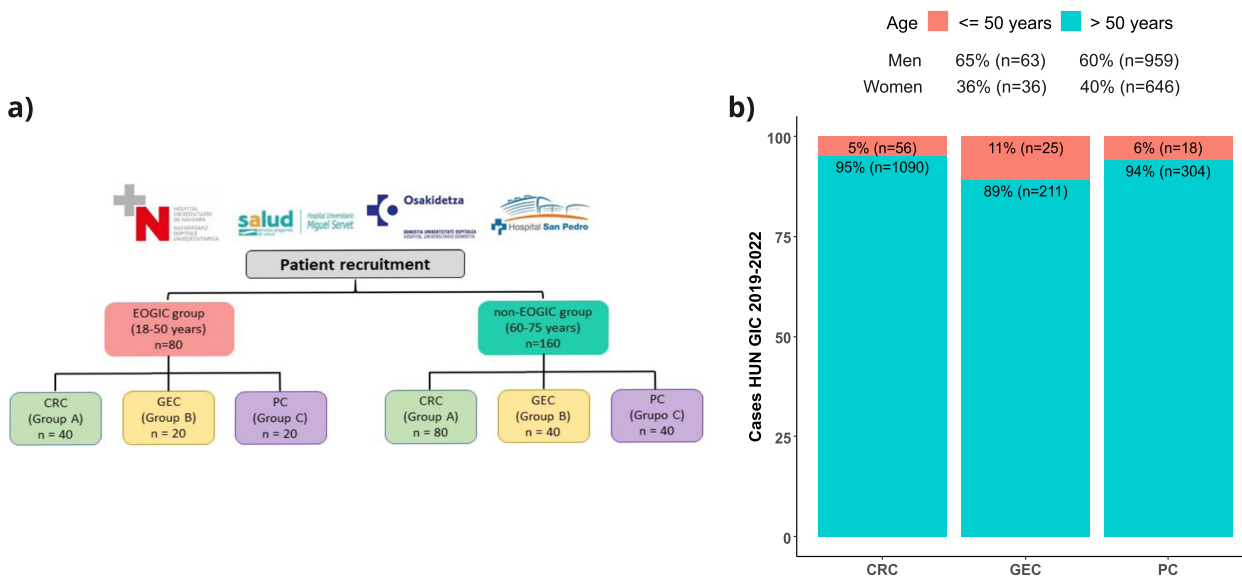


Fig. 1 Estimated patient recruitment. **a** Diagram of the recruiting centers and estimated number of patients to be recruited by group. **b** Number of cases with GI cancer treated in Medical Oncology Unit of the HUN in the years 2019–2022. Red depicts EOGIC, while green depicts non-EOGIC (age > 50 years) CRC: colorectal cancer; EOGIC: early onset gastrointestinal cancer. GEC: gastroesophageal cancer; PC: pancreatic cancer

Medication and Sanitary Product Agency (AEMPS). Previous to the inclusion, informed consent will be obtained for each patient.

Data collection and analyses

Figure 2 shows a general outlook of the study. After signing the informed consent, patients will be registered and study data will be collected and managed using REDCap electronic data capture tools hosted at VHIO internal server [37, 38]. An automated ID will be generated for each patient. Demographic, clinical and histopathological data will be recorded; patient-reported questionnaires will be collected, and biological samples will be obtained for further analyses.

Demographic and clinical data and histopathological data

Recruiting medical personnel will be in charge of obtaining the variables in REDCap via a complete clinical interview and screening of the digital clinical record.

Demographic information related to the patients’ environment will be recorded, including, town of residence, country of origin and year of migration if applicable, housing status, social network, educational level and occupation.

In parallel, a comprehensive clinical profile of the patient will also be recorded, including ECOG Performance Status at diagnosis, height, current and usual body weight, body mass index, waist-hip ratio, allergy record, current medication, non-steroidal anti-inflammatory drug (NSAID) and antibiotics consumption, toxic habits,

co-morbidities (*H. Pylori* infections and treatment, particularly for patients with GEC), previous surgeries in the affected organs, previously diagnosed malignancies and familial oncologic history as well as gynecological history for women. Additionally, a complete blood analysis covering a hemogram and a complete biochemistry as well as the tumoral markers carcinoembryonic antigen (CEA) and Carbohydrate antigen 19–9 (CA 19–9) will also be included.

Histopathological data and tissue proteome

General histopathological and molecular characteristics of the tumor will be comprehensively studied, via specific determinations relevant for each of the etiologies.

General histopathological characteristics, tumor location and clinical tumor stage will be recorded. Colorectal and gastroesophageal cancer staging will be assessed according to The Eighth Edition AJCC Cancer Staging Manual [39]. Pancreatic cancer will be classified as resectable, borderline resectable, locally advanced unresectable and metastatic. Number of metastatic sites and their location will also be recorded in stage IV of the disease. Histology type and differentiation grade will be assessed in formalin-fixed paraffin-embedded (FFPE) tissue. Molecular characteristics to be recorded in FFPE tissue include MMR/MSI status by immunohistochemical analysis of mismatch-repair proteins (*MLH1*, *MSH2*, *MSH6*, and *PMS2*) as well as CD3 and CD8 infiltration in all tumors. Additional molecular characteristics will also be measured according to tumor type: *HER2*

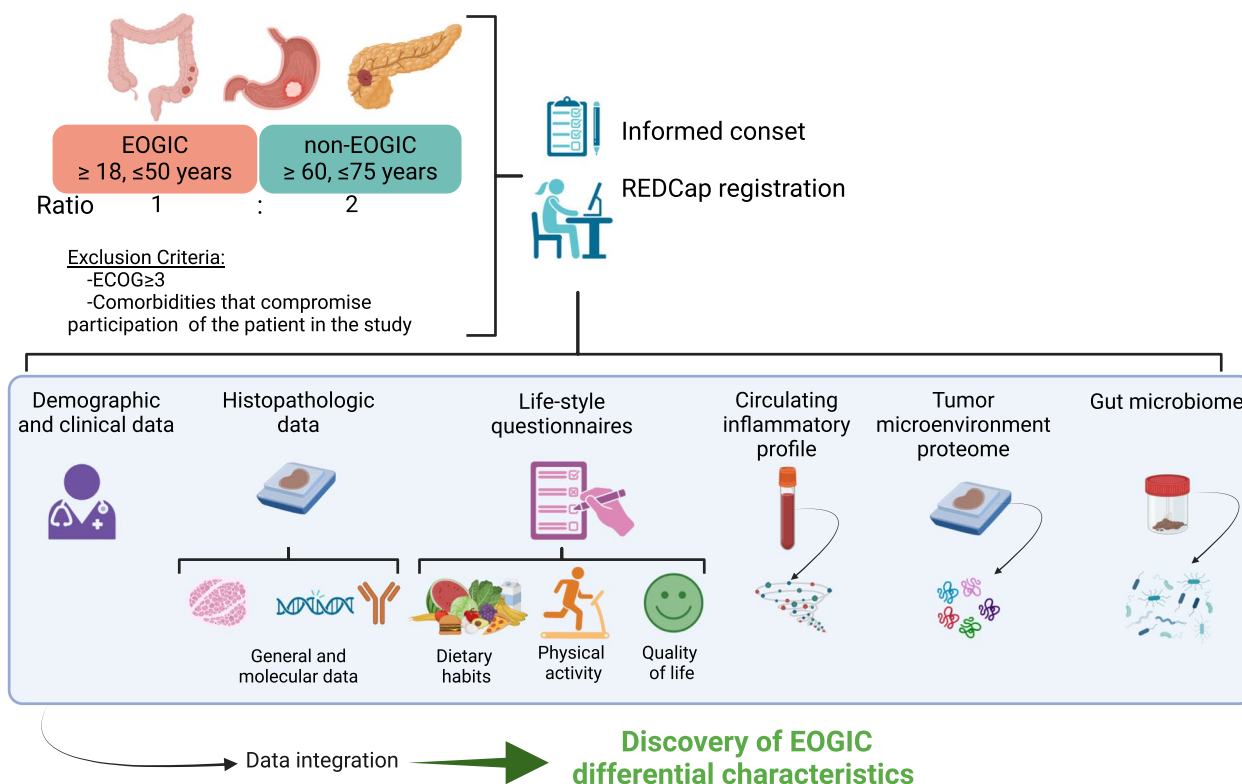


Fig. 2 General diagram of the study. Diagram summarizing patient recruitment criteria and data to be recorded in the TEOGIC study. EOGIC: early onset gastrointestinal cancer; ECOG: Eastern Cooperative Oncology Group

amplification status, *KRAS*, *NRAS* and *BRAF* mutations for CRC; *HER2* overexpression/amplification status, PDL1 (CPS) and CLDN 18.2 and FGFR2b expression for GEC; and *KRAS* mutations for PC. For all tumors, Next-Generation Sequencing (NGS) analyses targeting cancer-related genes will be performed if a suitable sample is available.

Histopathological characteristics will be preferentially assessed in the recruitment center. If available after routine determinations for diagnosis and clinical management, FFPE tissue samples will be sent to the Pathology Department of HUN in order to perform a centralized assessment of missing determinations.

Lifestyle associated variables

Questionnaires related to dietary habits, physical activity and general quality of life will be collected for each patient.

Dietary habits questionnaire: a food-frequency questionnaire (FFQ) consisting on 136 items that evaluate intake of total calories, protein, carbohydrates, fats (saturated, mono and polyunsaturated), alcohol, cholesterol, fiber, vitamin A and vitamin C in the last year will be employed. This questionnaire was developed fully in Spanish and based on dietary habits in the country [40],

since its development it has also been updated in agreement with food composition tables for Spain [41]. Due to its complexity it requires the assistance of trained personnel for its completion [40]. The questionnaire has been validated in several epidemiological studies, including studies in patients with cancer [41–44]. This questionnaire will always be answered with the assistance of trained personnel via telephone calls with the patient.

Physical activity questionnaire: the global physical activity questionnaire (GPAQ) developed by the World Health Organization consisting of 16 items that evaluate physical activity and sedentary behavior will be employed. This questionnaire has been extensively validated, also in patients with cancer [45, 46] and in its Spanish version [47]. Although initially developed for face-to-face interviews with trained personnel, several studies have validated administration of the questionnaire as self-reported [48–51].

Quality of life questionnaire: the short form health survey (SF-36) will be employed. The SF-36 questionnaire was developed in the USA to assess general health status in the Medical Outcomes Study [52]. Since then, it has been translated to numerous languages including Spanish [53, 54] and has been extensively validated in patients with cancer [55–57], within the Spanish context [58–60].

This questionnaire was initially developed to be self-reported [52], however, several studies have also reported its validity via telephone-calls or online [61–63].

In this study, the FFQ questionnaire will always be completed with the assistance of trained personnel via previously scheduled telephone calls with the patients. The GPAQ and SF-36 will be self-reported, and when needed, patients will be offered the assistance of trained personnel via telephone calls.

Of note, GPAQ and SF-36 questionnaires are designed to assess physical activity and quality of life at the time of questionnaire completion. Nevertheless, in this study, patients will be recruited after the diagnosis of cancer, which may heavily affect patients' lifestyle. Moreover, it is the focus of this study to compare EOGIC and non-EGOIC patient's general lifestyle and its impact on EOGIC development. Therefore, questionnaires will be adapted to evaluate patients' lifestyle 1 year before diagnosis.

Circulating inflammatory protein profile

The circulating inflammatory protein profile will be obtained for each patient. For this purpose, blood samples will be obtained at the time of recruitment, and will always be treatment-naïve including surgery (i.e. samples will be obtained before any kind of onco-specific therapeutic intervention, including resection intended surgery). A sample of 10 ml of total blood in vacutainer® EDTA tubes will be drawn by trained nurses. Thereafter, samples will be processed for plasma separation within 2 h since collection, following standardized work protocols in agreement with the Biobank of Navarrabiomed, and stored at -80°C. Samples will be sent to Navarrabiomed-HUN for centralized storage and subsequent analyses.

In order to assess the circulating inflammatory profile of the patients, a panel of 96 key proteins involved in tumoral and inflammatory processes will be assessed (Olink® Target 96 Immuno-Oncology) by the proximity extension assay (PEA) technology.

Feces samples: microbiome analysis

Fecal microbiome analyses will be conducted by the sequencing the *16S* gene. Feces samples will be obtained at the time of recruitment, and will always be treatment-naïve including surgery. Samples will be self-collected by the patients following detailed instructions with the DANASTOOL Sample Collection MICROBIOME kit (Danagen). This kit consists of a sample collection instrument and a tube holding a stabilization solution that enables stabilization of DNA for several months at room temperature and indefinitely at -80°C. Upon reception, feces samples will be homogenized and stored at -80°C,

thereafter they will be sent to Navarrabiomed-HUN for further processing and analyses.

DNA extraction will be conducted using the DANASTOOL Microbiome Fecal DNA kit (Danagen), and the quantity and quality of the extracted DNA will be assessed with NanoDrop ND-1000 (Thermo Fisher Scientific) and the Qubit Fluorometer (Thermo Fisher Scientific), and stored at -80°C. For microbiome analyses, libraries will be prepared with the xGen Amplicon Panels for 16S and IST1 kit (IDT), which allows for the PCR-based enrichment on the *16S* v2 V1-V9 rRNA regions. Thereafter, sequencing will be carried out on the Illumina NovaSeq platform. After quality control (QC) assessment, data will be processed with QIIME to obtain Operational Taxonomic Units (OTUs) against reference databases (Greengenes or Silva).

Tumor microenvironment proteomic analysis

FFPE tissue samples will be obtained from diagnostic biopsies or resection specimens naïve of treatment. Proteomic analyses will be performed after all needed determinations for diagnosis and clinical management have been conducted. Proteomic analyses will be conducted in the Proteomics unit of Navarrabiomed (ISO9001:2015 certified). Samples will undergo "in-gel" digestion, peptide reconstitution, and purification before identification by liquid chromatography-tandem mass spectrometry (LC-MS/MS), using the Ultimate 3000 RSLCnano system coupled with an Orbitrap Exploris 480 mass spectrometer (Thermo Fisher Scientific). Spectra resulting from the analysis will be processed using MaxQuant software and compared with the most updated *H.sapiens* reference proteome from the UniProt database.

Statistical analyses

Descriptive and comparative analyses will be performed separately for each type of tumor: CRC (Group A), GEC (Group B), and PC (Group C) as well as at a pan-tumor level, comparing all EOGIC to non-EOGIC patients.

Population characteristics will be described using measures of central tendency and dispersion for continuous variables and frequencies and percentages for categorical variables. Normality of the sample will be analyzed for continuous variables through histogram exploration, the Shapiro–Wilk test, and homoscedasticity by Levene's test. Group comparisons will be made using Student's t-test or Mann–Whitney U test for continuous variables and Pearson's Chi-square test or Fisher's exact test for categorical variables. Correlation between different measurements and conditions of interest will be assessed using Pearson's or Spearman's correlation tests.

Analyses for circulating inflammatory proteins, the microbiota and the tumor microenvironment proteome,

will be conducted utilizing specialized software for each application. Raw data will be processed, applying quality criteria, and differential signatures between groups will be identified. For signature identification, previously mentioned tests, including correction methods for multiple comparisons such as the false discovery rate (FDR).

For pan-tumor and gender/sex-level comparisons, machine learning algorithms will be used, allowing integration of all patient clinical and lifestyle profile data with variables derived from circulating inflammatory proteins, microbiome and tumor microenvironment proteome profiles, identifying clusters of patients with similar patterns.

Statistical analysis will be performed using IBM Statistical Package for the Social Sciences (SPSS) version 24.0, GraphPad Prism v9, and R 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>).

Discussion

EOGIC incidence is rising worldwide at an alarming rate and due to the relatively low proportion of hereditary cases, it is thought to be associated with changes in lifestyle and environmental exposure [6]. Some studies have shown that EOGIC has differential clinical and molecular features [6, 8, 18, 20, 24], but prospective and comprehensive studies are still lacking, especially in our geographical area [3, 4].

Patients with EOGIC represent an especially vulnerable population, with distinct challenges and needs. The diagnosis of EOGIC can lead to the interruption of education, professional activity and/or familial responsibilities and have a significant effect on their overall quality of life. Cancer associated morbidity is also distinctive in these patients that may have to face premature mortality. Moreover, the longer life expectancy of these patients requires special attention to the long-term side effects of chemo-radiotherapy treatments, especially those affecting cardiovascular, reproductive and sexual capacity; as well as the effect on body image after aggressive surgeries [6, 25–30].

The TEOGIC study aims to characterize the distinctive features of EOGIC through a comprehensive approach that considers parameters encompassing clinical and demographic, as well as lifestyle characteristics of the patient, together with the biological characteristics of the tumor. The identification of unique biomarkers holds the potential for integration across preventive and therapeutic levels, contributing to the development of precision medicine. Our overarching goal is to enhance the prognosis and quality of life for EOGIC patients. This study has the potential to significantly influence the prevention and awareness of gastrointestinal cancer development

within the population. By identifying lifestyle-associated factors linked to EOGIC, we can enhance preventive measures. Additionally, our molecular studies (microbiome and circulating proteins) could contribute to the identification of new sources of non-invasive biomarkers with a potential for early diagnosis and screening programs currently lacking in gastroesophageal and pancreatic cancer [17]. Similarly, the characterization of the proteome of the tumor microenvironment could unravel novel potential therapeutic targets and thus lead to the generation of new hypotheses in need of validation in upcoming studies.

The main limitations of the present study are the inclusion of a relatively small number of patients and a recruitment restricted to a limited geographical area. The TEOGIC study represents the first initiative to characterize EOGIC in our area and would serve as the founding experience to establish the collaborations and synergies of multidisciplinary teams that would enable the development of extended studies in the future. Results derived from this study will also contribute to raising awareness about the importance of new risk factors in the development of EOGIC and thus, serve to promote lifestyle changes in the general population. Additionally, as aforementioned, the TEOGIC project features an exploratory nature to finding new sources of biomarkers and therapeutic targets.

Conclusion

TEOGIC is a transversal study of Early-Onset Gastrointestinal Cancer conducted across four hospitals within the national public healthcare system, a biomedical research center, and a research group specialized in data integration and interpretation in oncology. Its primary objective is to identify distinctive characteristics associated with patients with EOGIC. The findings from this study are expected to have an impact in preventive measures and therapeutic interventions in this group of patients.

Abbreviations

APC	Adenomatous polyposis coli
CA19-9	Carbohydrate antigen 19–9
CEA	Carcinoembryonic antigen
CLDN 18.2	Claudin 18.2
CRC	Colorectal cancer
CDH1	Cadherin 1
EBV	Epstein-Barr virus
ECOG	Eastern Cooperative Oncology Group
EOGIC	Early onset Gastrointestinal Cancer
EOCRC	Early onset colorectal cancer
EOGEC	Early onset gastroesophageal cancer
FDR	False discovery rate
FFQ	Food frequency questionnaire
FFPE	Formalin-fixed paraffin-embedded
FGFR2b	Fibroblast growth factor receptor 2b
GPAQ	Global physical activity questionnaire
GEC	Gastroesophageal cancer

GI	Gastrointestinal cancer
HMS	Miguel Servet Hospital
HSP	San Pedro Hospital
HUD	University Hospital Donostia
HUN	University Hospital of Navarra
KRAS	Kirsten rat sarcoma virus
MMR	Mismatch repair
MSI	Microsatellite instability
NSAID	Non-steroidal anti-inflammatory drug
OTUs	Operational Taxonomic Units
PC	Pancreatic cancer
PDL1	Programmed death-ligand 1
PEA	Proximity extension assay
PIK3	Phosphatidylinositol 3-kinase
QC	Quality control
TEOGIC	Transversal study on early-Onset Gastrointestinal Cancer
TGFβ	Transforming growth factor β
USA	United States of America
VHIO	Vall d'Hebron Institute of Oncology

Acknowledgements

The authors acknowledge in advance to all patients and families for their participation.

Authors' contributions

Conceptualization, design, and writing: RV, IL, NC, AL, AEH, SL, HA and MA. Patient selection and recruitment: RV, MA, VA, IHG, EM, DG, JS, MR, FE, AGN, AL, VA, AMC and RPC. Histopathological data: IFR. Data collection and analyses: IL, NC, AL, AEH, IC, RD, FRP and CV. All authors have read and agreed to the published version of the manuscript.

Funding

The TEOGIC project has been previously scientifically peer reviewed by the Carlos III Health Institute, with a funding decision within the competitive Research Projects AES 2023 call, ref: PI23/01514. IL is financed by National Agency of Research (AEI) in the "Juan de la Cierva-Postdoctoral formación" (8FJC2021-046521-I); AL is supported by the "Clínico Junior en el Territorio AECC 2023" grant from the Spanish Association Against Cancer (AECC) (CLJUN234885LECU). HA is supported by the "Clínico Junior 2019" grant from the Spanish Association Against Cancer (AECC) (CLJUN19010ARAS); MA is partially financed by the Government of Navarra in the La Caixa Program. AMC is supported by the "AECC SEED IDEAS 2023" grant from the Spanish Association Against Cancer (AECC) and "Knowledge Generation Projects 2022" grant from the State Scientific, Technical and Innovation Research Plan 2021—2023 from the Ministry of science, innovation and universities. Spanish government.

Availability of data and materials

No datasets were generated or analysed during the current study.

Declarations

Ethics approval and consent to participate

Ethics Committee for Investigation with medicinal products from Navarra, have approved this project with the following record ID, CI_2023/144. This collection has been included in to the National Institute of Health Carlos III (ISCIII) registry with the following reference, C.0008110.

Consent for publication

Not applicable.

Competing interests

RV has been involved as a consultant for advisory roles with Servier, Roche and Merck Sharp and has received speaker honoraria from Roche, Amgen, Merck Sharp and Dohme, AstraZeneca. IL, NC, AEH, IC, FRP, CV, DG, JS, IFR, MR, FE, AGN, AL, VA, RPC and SL declare no conflict of interest. AL has received speaker honoraria from Pierre-Fabre. HA has been involved as a consultant for advisory roles from AstraZeneca and for trial coordination from Ferrer Farma and received speaker honoraria from Takeda. VA has been involved as a consultant for advisory roles and received speaker honoraria from MSD, Bristol, Lilly, Astra-Zeneca and Pierre-Fabre. IHG has received speaker honoraria from

AstraZeneca. EM has been involved as a consultant for advisory roles with Servier, Roche and Merck Sharp and Dohme. RD declares advisory role for Roche, Foundation Medicine, received a speaker's fee from Roche, Ipsen, Amgen, Servier, Sanofi, Libbs, Merck Sharp & Dohme, Lilly, AstraZeneca, Janssen, Takeda, Bristol Myers Squibb, GlaxoSmithKline, Gilead, research grants from Merck, Novartis, Daiichi-Sankyo, GlaxoSmithKline and AstraZeneca, and is investor in Trialing Health, S.L. AMC declares advisory role for Servier, Novartis, Pharmamar and Bristol Myers Squibb, received speaker honoraria from Novartis, Roche, Merck Sharp and Dohme, Merck, Pierre Fabre, Amgen, Sanofi, Servier and Bristol Myers Squibb. MA has been involved as a consultant for advisory roles with Amgen, BMS, MSD, Lilly and Servier.

Author details

¹Oncobiona Group, Navarrabiomed-Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Spain. ²Department of Medical Oncology, Hospital Universitario de Navarra (HUN), Pamplona, Spain. ³Oncology Data Science Group, Vall d'Hebron Institute of Oncology (VHIO), Barcelona, Spain. ⁴Department of Surgery, Hospital Universitario de Navarra (HUN), Pamplona, Spain. ⁵Department of Pathology, Hospital Universitario de Navarra (HUN), Pamplona, Spain. ⁶Molecular Pathology of Cancer Group, Navarrabiomed, Universidad Pública de Navarra (UPNA), Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Spain. ⁷Department of Gastroenterology and Hepatology, Hospital Universitario de Navarra (HUN), Pamplona, Spain. ⁸Digestive System and Metabolism Diseases Group, Instituto de Investigación Sanitaria de Navarra (IdiSNA), Pamplona, Spain. ⁹Department of Medical Oncology, Hospital Universitario Miguel Servet, IISA, Zaragoza, Spain. ¹⁰Department of Medical Oncology, Hospital Universitario Donostia, San Sebastian, Spain. ¹¹Department of Medical Oncology, Hospital San Pedro, Logroño, Spain. ¹²University of Vic – Central University of Catalonia, Vic, Spain.

Received: 12 February 2024 Accepted: 30 May 2024

Published online: 01 June 2024

References

- Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49.
- (SEOM) SEEdOM. Las Cifras del Cancer en España 2024. <https://seom.org/ejercicio-contra-el-cancer/las-cifras-del-cancer-en-espana-20242024>. Available from: https://www.seom.org/images/LAS_CIFRAS_2024.pdf.
- Siegel RL, Torre LA, Soerjomataram I, Hayes RB, Bray F, Weber TK, et al. Global patterns and trends in colorectal cancer incidence in young adults. *Gut*. 2019;68(12):2179–85.
- Araghi M, Soerjomataram I, Bardot A, Ferlay J, Cabasag CJ, Morrison DS, et al. Changes in colorectal cancer incidence in seven high-income countries: a population-based study. *Lancet Gastroenterol Hepatol*. 2019;4(7):511–8.
- Sung H, Siegel RL, Rosenberg PS, Jemal A. Emerging cancer trends among young adults in the USA: analysis of a population-based cancer registry. *Lancet Public Health*. 2019;4(3):e137–47.
- Ben-Aharon I, van Laarhoven HWM, Fontana E, Obermannova R, Nilsson M, Lordick F. Early-onset cancer in the gastrointestinal tract is on the rise: evidence and implications. *Cancer Discov*. 2023;13(3):538–51.
- Liao H, Cai S, Bai Y, Zhang B, Sheng Y, Tong S, et al. Prevalence and spectrum of germline cancer susceptibility gene variants and somatic second hits in colorectal cancer. *Am J Cancer Res*. 2021;11(11):5571–80.
- Pocurull A, Herrera-Pariente C, Carballal S, Llach J, Sánchez A, Carot L, et al. Clinical, molecular and genetic characteristics of early onset gastric cancer: analysis of a large multicenter study. *Cancers (Basel)*. 2021;13(13):3132.
- Lauby-Secretan B, Scoccianti C, Loomis D, Grosse Y, Bianchini F, Straif K, et al. Body fatness and cancer-viewpoint of the IARC Working Group. *N Engl J Med*. 2016;375(8):794–8.
- Nguyen LH, Liu PH, Zheng X, Keum N, Zong X, Li X, et al. Sedentary Behaviors, TV Viewing Time, and Risk of Young-Onset Colorectal Cancer. *JNCI Cancer Spectr*. 2018;2(4):pky073.

11. Nguyen LH, Cao Y, Hur J, Mehta RS, Sikavi DR, Wang Y, et al. The sulfur microbial diet is associated with increased risk of early-onset colorectal cancer precursors. *Gastroenterology*. 2021;161(5):1423–32.e4.
12. Carroll KL, Frugé AD, Heslin MJ, Lipke EA, Greene MW. Diet as a risk factor for early-onset colorectal adenoma and carcinoma: a systematic review. *Front Nutr*. 2022;9:896330.
13. Petrillo A, Federico P, Marte G, Liguori C, Seeber A, Ottaviano M, et al. Non-hereditary early onset gastric cancer: an unmet medical need. *Curr Opin Pharmacol*. 2023;68:102344.
14. Song H, Held M, Sandin S, Rautelin H, Eliasson M, Söderberg S, et al. Increase in the prevalence of atrophic gastritis among adults age 35 to 44 years old in Northern Sweden between 1990 and 2009. *Clin Gastroenterol Hepatol*. 2015;13(9):1592–600.e1.
15. Arnold M, Ferlay J, van Berge Henegouwen MI, Soerjomataram I. Global burden of oesophageal and gastric cancer by histology and subsite in 2018. *Gut*. 2020;69(9):1564–71.
16. Chen C, Yang Y, Li P, Hu H. Incidence of gastric neoplasms arising from autoimmune metaplastic atrophic gastritis: a systematic review and case reports. *J Clin Med*. 2023;12(3):1062.
17. Fontana E, Meyers J, Sobrero A, Iveson T, Shields AF, Taieb J, et al. Early-onset colorectal adenocarcinoma in the IDEA database: treatment adherence, toxicities, and outcomes with 3 and 6 months of adjuvant fluoropyrimidine and oxaliplatin. *J Clin Oncol*. 2021;39(36):4009–19.
18. Zaborowski AM, Abdile A, Adamina M, Aigner F, d'Allens L, Allmer C, et al. Characteristics of early-onset vs late-onset colorectal cancer: a review. *JAMA Surg*. 2021;156(9):865–74.
19. Lieu CH, Golemis EA, Serebriiskii IG, Newberg J, Hemmerich A, Connelly C, et al. Comprehensive genomic landscapes in early and later onset colorectal cancer. *Clin Cancer Res*. 2019;25(19):5852–8.
20. Willauer AN, Liu Y, Pereira AAL, Lam M, Morris JS, Raghav KPS, et al. Clinical and molecular characterization of early-onset colorectal cancer. *Cancer*. 2019;125(12):2002–10.
21. Kong C, Liang L, Liu G, Du L, Yang Y, Liu J, et al. Integrated metagenomic and metabolomic analysis reveals distinct gut-microbiome-derived phenotypes in early-onset colorectal cancer. *Gut*. 2023;72(6):1129–42.
22. Pérez-Wert P, Custodio A, Jimenez-Fonseca P, Carmona-Bayonas A, Lecumberri A, Cacho Lavin D, et al. Efficacy and safety of chemotherapy in young patients with advanced gastroesophageal adenocarcinoma: data from the Spanish AGAMENON-SEOM registry. *Gastric Cancer*. 2024;27(1):131–45.
23. Varghese AM, Singh I, Singh R, Kunte S, Chou JF, Capanu M, et al. Early-onset pancreas cancer: clinical descriptors, genomics, and outcomes. *J Natl Cancer Inst*. 2021;113(9):1194–202.
24. Castet F, Fabregat-Franco C, Castillo G, Navarro V, Sierra A, Acosta DA, et al. Clinical and genomic characterisation of early-onset pancreatic cancer. *Eur J Cancer*. 2023;194:113338.
25. Incrocci L, Jensen PT. Pelvic radiotherapy and sexual function in men and women. *J Sex Med*. 2013;10(Suppl 1):53–64.
26. Cercek A, Siegel CL, Capanu M, Reidy-Lagunes D, Saltz LB. Incidence of chemotherapy-induced amenorrhea in premenopausal women treated with adjuvant FOLFOX for colorectal cancer. *Clin Colorectal Cancer*. 2013;12(3):163–7.
27. Perl G, Nordheimer S, Lando S, Benedict C, Brenner B, Perry S, et al. Young patients and gastrointestinal (GI) tract malignancies - are we addressing the unmet needs? *BMC Cancer*. 2016;16:630.
28. Zebrack BJ, Mills J, Weitzman TS. Health and supportive care needs of young adult cancer patients and survivors. *J Cancer Surviv*. 2007;1(2):137–45.
29. Smith AW, Parsons HM, Kent EE, Bellizzi K, Zebrack BJ, Keel G, et al. Unmet support service needs and health-related quality of life among adolescents and young adults with cancer: the AYA HOPE Study. *Front Oncol*. 2013;3:75.
30. Sanford SD, Zhao F, Salsman JM, Chang VT, Wagner LI, Fisch MJ. Symptom burden among young adults with breast or colorectal cancer. *Cancer*. 2014;120(15):2255–63.
31. Suarez J, Alsina M, Castro N, Marin G, Llanos C, Ornoz B, et al. Higher rate of pathologic complete response in patients with early-onset locally advanced rectal cancer. *ESMO Gastrointestinal Oncol*. 2024;3:100033.
32. Schendel J, Jost E, Mah M, Mack L, McCall M, Gu N, et al. Gastric cancer management in elderly patients: a population-based study of treatment patterns and outcomes in gastric cancer patients ≥ 75 years from Alberta, Canada. *Am J Surg*. 2021;221(4):839–43.
33. Natori A, Chan BA, Sim HW, Ma L, Yokom DW, Chen E, et al. Outcomes by treatment modality in elderly patients with localized gastric and esophageal cancer. *Curr Oncol*. 2018;25(6):366–70.
34. Hagerty BL, Aversa JG, Dominguez DA, Davis JL, Hernandez JM, McCormick JT, et al. Age determines adjuvant chemotherapy use in resected stage II colon cancer. *Dis Colon Rectum*. 2022;65(10):1206–14.
35. Henry AC, Schouten TJ, Daamen LA, Walma MS, Noordzij P, Cirkel GA, et al. Short- and long-term outcomes of pancreatic cancer resection in elderly patients: a nationwide analysis. *Ann Surg Oncol*. 2022;29(9):6031–42.
36. Oba A, Wu YHA, Lieu CH, Meguid C, Colborn KL, Beaty L, et al. Outcome of neoadjuvant treatment for pancreatic cancer in elderly patients: comparative, observational cohort study. *Br J Surg*. 2021;108(8):976–82.
37. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–81.
38. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, et al. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform*. 2019;95:103208.
39. Amin MB, Greene FL, Edge SB, Compton CC, Gershenwald JE, Brookland RK, et al. The Eighth Edition AJCC Cancer Staging Manual: continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging. *CA Cancer J Clin*. 2017;67(2):93–9.
40. Martin-Moreno JM, Boyle P, Gorgojo L, Maisonneuve P, Fernandez-Rodriguez JC, Salvini S, et al. Development and validation of a food frequency questionnaire in Spain. *Int J Epidemiol*. 1993;22(3):512–9.
41. Romanos-Nanclares A, Gea A, Martínez-González M, Zazpe I, Gardeazabal I, Fernandez-Lazaro CI, et al. Carbohydrate quality index and breast cancer risk in a Mediterranean cohort: the SUN project. *Clin Nutr*. 2021;40(1):137–45.
42. Gardeazabal I, Ruiz-Canela M, Sánchez-Bayona R, Romanos-Nanclares A, Aramendia-Beitia JM, Shivappa N, et al. Dietary inflammatory index and incidence of breast cancer in the SUN project. *Clin Nutr*. 2019;38(5):2259–68.
43. Fernandez-Lazaro CI, Martínez-González M, Aguilera-Buenosvinos I, Gea A, Ruiz-Canela M, Romanos-Nanclares A, et al. Dietary antioxidant vitamins and minerals and breast cancer risk: prospective results from the SUN Cohort. *Antioxidants (Basel)*. 2021;10(3):340.
44. Fernández-Ballart JD, Piñol JL, Zazpe I, Corella D, Carrasco P, Toledo E, et al. Relative validity of a semi-quantitative food-frequency questionnaire in an elderly Mediterranean population of Spain. *Br J Nutr*. 2010;103(12):1808–16.
45. Baker JL, Di Meglio A, Gbenou AS, El Mouhebb M, Iyengar NM, Michiels S, et al. Association between physical activity and neoadjuvant chemotherapy completion and pathologic complete response in primary breast cancer: the CANTO study. *Br J Cancer*. 2022;127(5):886–91.
46. Lee J, Min J, Lee DH, Kang DW, Jeon JY. Intensity- and domain-specific physical activity levels between cancer survivors and non-cancer diagnosis individuals: a propensity score matching analysis. *Support Care Cancer*. 2021;29(2):661–8.
47. Ruiz-Casado A, Alejo LB, Santos-Lozano A, Soria A, Ortega MJ, Pagola I, et al. Validity of the physical activity questionnaires IPAQ-SF and GPAQ for cancer survivors: insights from a Spanish Cohort. *Int J Sports Med*. 2016;37(12):979–85.
48. Wanner M, Hartmann C, Pestoni G, Martin BW, Siegrist M, Martin-Diener E. Validation of the Global Physical Activity Questionnaire for self-administration in a European context. *BMJ Open Sport Exerc Med*. 2017;3(1):e000206.
49. Hoos T, Espinoza N, Marshall S, Arredondo EM. Validity of the Global Physical Activity Questionnaire (GPAQ) in adult Latinas. *J Phys Act Health*. 2012;9(5):698–705.
50. Sitthipornvorakul E, Janwantanakul P, van der Beek AJ. Correlation between pedometer and the Global Physical Activity Questionnaire on physical activity measurement in office workers. *BMC Res Notes*. 2014;7:280.
51. Chu AH, Ng SH, Koh D, Müller-Riemschneider F. Reliability and Validity of the Self- and Interviewer-Administered Versions of the Global Physical Activity Questionnaire (GPAQ). *PLoS ONE*. 2015;10(9):e0136944.

52. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*. 1992;30(6):473–83.
53. Alonso J, Prieto L, Antó JM. The Spanish version of the SF-36 Health Survey (the SF-36 health questionnaire): an instrument for measuring clinical results. *Med Clin (Barc)*. 1995;104(20):771–6.
54. Bullinger M, Alonso J, Apolone G, Leplège A, Sullivan M, Wood-Dauphinee S, et al. Translating health status questionnaires and evaluating their quality: the IQOLA Project approach. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51(11):913–23.
55. Schurr T, Loth F, Lidington E, Piccinin C, Arraras JI, Groenvold M, et al. Patient-reported outcome measures for physical function in cancer patients: content comparison of the EORTC CAT Core, EORTC QLQ-C30, SF-36, FACT-G, and PROMIS measures using the International Classification of Functioning, Disability and Health. *BMC Med Res Methodol*. 2023;23(1):21.
56. Brown JC, Damjanov N, Courneya KS, Troxel AB, Zemel BS, Rickels MR, et al. A randomized dose-response trial of aerobic exercise and health-related quality of life in colon cancer survivors. *Psychooncology*. 2018;27(4):1221–8.
57. Hara T, Kogure E, Iijima S, Fukawa Y, Kubo A, Kakuda W. Minimal clinically important difference in postoperative recovery among patients with gastrointestinal cancer. *Support Care Cancer*. 2022;30(3):2197–205.
58. Vilagut G, Ferrer M, Rajmil L, Rebollo P, Permanyer-Miralda G, Quintana JM, et al. The Spanish version of the Short Form 36 Health Survey: a decade of experience and new developments. *Gac Sanit*. 2005;19(2):135–50.
59. Vilagut G, Valderas JM, Ferrer M, Garin O, López-García E, Alonso J. Interpretation of SF-36 and SF-12 questionnaires in Spain: physical and mental components. *Med Clin (Barc)*. 2008;130(19):726–35.
60. Santillán Coello JM, Sánchez Barrueco Á, González Galán F, Díaz Tapia G, Mahillo Fernández I, Villacampa Aubá JM, et al. Validation of a Spanish chronic obstructive sialadenitis quality of life questionnaire (CSOC). *Acta Otorrinolaringol Esp (Engl Ed)*. 2023;74(2):116–23.
61. Açma A, Carrat F, Hejblum G. Comparing SF-36 scores collected through web-based questionnaire self-completions and telephone interviews: an ancillary study of the SENTIPAT multicenter randomized controlled trial. *J Med Internet Res*. 2022;24(3):e29009.
62. Wettergren L, Mattsson E, von Essen L. Mode of administration only has a small effect on data quality and self-reported health status and emotional distress among Swedish adolescents and young adults. *J Clin Nurs*. 2011;20(11–12):1568–77.
63. Basnov M, Kongsved SM, Bech P, Hjollund NH. Reliability of short form-36 in an Internet- and a pen-and-paper version. *Inform Health Soc Care*. 2009;34(1):53–8.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.