GastroIntestinal Cancer in Children and Adolescents (GICCA study)

(Version November 2022)

PROTOCOL TITLE: 'Gastrointestinal cancer in children and adolescents (GICCA): a SEER population-based study'

Protocol ID	**GICCA**
Short title	**GI cancer in children and adolescents**
EudraCT number	**N/A**
Version	**1**
Date	**November 6, 2022**
Coordinating investigators /project leaders	**Caroline E. Veldhuis-Hogen Esch, M.D. Ph.D.** **Division of Pediatric Gastroenterology** **Juliana Children's Hospital** **c.veldhuis-hogenesch@hagaziekenhuis.nl** **Daniëlle M. Hendriks, M.D.** **Division of Pediatric Gastroenterology** **Juliana Children's Hospital** **d.hendriks@hagaziekenhuis.nl**
Principal investigator	**Paul G. Kemps, M.D.** **(1) Department of Pathology** **Leiden University Medical Center** **(2) Division of Pediatric Gastroenterology** **Juliana Children's Hospital** **p.g.kemps@lumc.nl**
Sponsor	**Paul G. Kemps (Sponsor-Investigator)**
Subsidising party	**N/A**
Independent expert(s)	**N/A**

PROTOCOL SIGNATURE SHEET

Name	Signature	Date
Principal investigator: **Paul G. Kemps, M.D.**	Kemps	November 7, 2022
Study statistician: **Jelle J. Goeman, Ph.D.** *Professor of Biostatistics* *Dept. of Biomedical Data Sciences* *Leiden University Medical Center*		December 22, 2022

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND RELEVANT DEFINITIONS

ABR	**General Assessment and Registration form (ABR form), the application form that is required for submission to the accredited Ethics Committee; in Dutch: Algemeen Beoordelings- en Registratieformulier (ABR-formulier)**
AE	**Adverse Event**
AR	**Adverse Reaction**
CA	**Competent Authority**
CCMO	**Central Committee on Research Involving Human Subjects; in Dutch: Centrale Commissie Mensgebonden Onderzoek**
CV	**Curriculum Vitae**
DSMB	**Data Safety Monitoring Board**
EU	**European Union**
EudraCT	**European drug regulatory affairs Clinical Trials**
GCP	**Good Clinical Practice**
GDPR	**General Data Protection Regulation; in Dutch: Algemene Verordening Gegevensbescherming (AVG)**
IB	**Investigator's Brochure**
IC	**Informed Consent**
IMP	**Investigational Medicinal Product**
IMPD	**Investigational Medicinal Product Dossier**
METC	**Medical research ethics committee (MREC); in Dutch: medisch-ethische toetsingscommissie (METC)**
(S)AE	**(Serious) Adverse Event**
SPC	**Summary of Product Characteristics; in Dutch: officiële productinformatie IB1-tekst**
Sponsor	**The sponsor is the party that commissions the organisation or performance of the research, for example a pharmaceutical company, academic hospital, scientific organisation or investigator. A party that provides funding for a study but does not commission it is not regarded as the sponsor, but referred to as a subsidising party.**
SUSAR	**Suspected Unexpected Serious Adverse Reaction**

UAVG	**Dutch Act on Implementation of the General Data Protection Regulation; in Dutch: Uitvoeringswet AVG**
WMO	**Medical Research Involving Human Subjects Act; in Dutch: Wet Medisch-wetenschappelijk Onderzoek met Mensen**

SUMMARY

Rationale: Gastrointestinal malignancies are rare in children and adolescents. Consequently, data on clinical features, treatments and outcomes of these patients are limited.
Objective: We aim to describe the patient and tumor characteristics, treatments, and outcomes of children and adolescents diagnosed with primary gastrointestinal malignancies, and to analyze these data for trends over time. In addition, we want to explore independent prognostic factors for overall survival. Finally, we want to describe subsequent primary malignancies diagnosed among patients from the study cohort, as well as to calculate their risk of developing a second primary cancer relative to the general population. Data on children (aged 0-17 years at diagnosis) and young people aged 18-24 years at diagnosis are analyzed separately. Special focus is laid on patients with primary carcinomas of the gastrointestinal tract; particularly colorectal carcinoma.
Study design: This is an observational population-based cohort study based on the publicly available Surveillance, Epidemiology, and End Results (SEER) database.
Study population: Children (aged 0-17 years) and young people aged 18-24 years with a diagnosis of a primary gastrointestinal malignancy registered as their first primary malignancy in the SEER 17 database between 2000 and 2019.
Intervention: Not applicable.
Main study parameters/endpoints: Descriptive statistics of patient and tumor characteristics at diagnosis, first course treatments, overall survival, and subsequent primary malignancies.
Secondary study parameters/endpoints: 1) Time trends in incidence, stage at diagnosis, first course treatments, mortality, and overall survival. 2) Associations between the variables age at diagnosis, sex, race and origin, year of diagnosis, stage at diagnosis, and histologic category (when relevant) with overall survival in multivariable analysis. 3) Standardized incidence of second primary malignancies in patients who survived (i) two or more months, or (ii) five or more years after initial gastrointestinal cancer diagnosis, relative to the general population.
Nature and extent of the burden and risks associated with participation, benefit and group relatedness: No burden and/or risks; no benefit.

1. INTRODUCTION AND RATIONALE

Gastrointestinal (GI) malignancies – also called "digestive cancers" – are rare in children and adolescents[1]. These malignancies develop in the GI tract and/or accessory digestive organs, including the liver, biliary tract and pancreas[2,3]. Notably, the mouth, pharynx and salivary glands are also part of the GI system, but malignancies in these locations are generally classified within the group of head and neck cancers[3,4]. Therefore, like in prior studies of digestive cancers[2], these malignancies are not investigated in this study; the term "GI tract" is used to refer to the alimentary canal running from the esophagus to the anus. Because of the rarity of GI malignancies in children and adolescents, data on clinical features, treatments and outcomes of these patients are limited.

Childhood is defined as the developmental period between 0 and 17 years of age[5]. Importantly, much debate still exists about the definition of adolescence[5,6]. The United Nations (UN) and World Health Organization define adolescence as the period between 10 and 19 years of age. However, this definition dates back to the mid-20th century, with significant changes in human biological and social development since that time[5]. Therefore, Sawyer *et al.* recently proposed an expanded definition of adolescence, which encompasses the period between 10 and 24 years of age[5]. This new definition overlaps with those of the terms youth (defined by the UN as people aged 15-24 years) and young people (defined as people aged 10-24 years)[5,6]. Adolescents have often been overlooked in health care, research and policy[7]; one growing response is that from pediatricians to more embrace adolescent health[8]. A recent survey among pediatricians in high-income North America revealed a mean preferred upper age limit of pediatric practice of 21.7 years, which was significantly higher than the mean current upper age limit of the pediatricians' practices of 19.2 years[8]. To comply with this increasing focus of pediatricians on adolescent health, our study will include patients aged 0-24 years at primary gastrointestinal cancer diagnosis. Yet, the upper age limit of pediatric care is still below 18 years in many – particularly lower-income – countries[8]. Therefore, data on children (aged 0-17 years) and young people aged 18-24 years will be (preferably) analyzed separately – to ensure that the study results are appropriate to readers from different health care settings.

One common site of GI malignancies in children and adolescents is the liver. There are different types of primary liver cancer. The most common types in children are hepatoblastoma (HB), hepatocellular carcinoma (HCC) and undifferentiated embryonal sarcoma of the liver (UESL)[9]. Several (population-based) cohort studies investigating children and/or adolescents with hepatic malignancies have already been published[10–13]. However, a comprehensive overview of population-based data on the different hepatic malignancies and other gastrointestinal cancers in children and adolescents is still lacking. Regarding malignancies with the GI tract as their primary site, lymphoma is one of the most common subtypes in children – often representing intestinal Burkitt lymphoma (BL)[1]. Although the true primary site of hematologic malignancies is the hematopoietic system, we deliberately do not exclude these malignancies, because we want to investigate all cancers with the GI system as their primary location. Several (population-based) cohort studies investigating children and/or adolescents with BL have been published[1,14]; however, these did not specifically focus on patients with gastrointestinal BL. Finally, limited data is currently available on children and adolescents with primary carcinomas of the GI tract, which predominantly include colorectal carcinoma (CRC) and gastric carcinoma (GC)[15–18]. To address this limited data, our study will focus on patients with these rare cancers.

Lastly, recent studies have identified genetic predisposition in a significant proportion of children and adolescents with cancer, including young people with colorectal cancer[19,20]. This partly explains the increased risk of subsequent malignant neoplasms among childhood cancer survivors[21]. In addition, specific treatments can also increase the risk of second primary malignancies (SPM)[22]. Many prior studies have investigated SPM in five-year childhood and/or adolescent cancer survivors[22–27]. These studies exclude second cancers diagnosed within five years of initial cancer diagnosis; thereby, they reduce surveillance bias. However, recent population-based studies investigating SPM in children with cancer that survived two or more months revealed that a significant proportion of second cancers actually develop within those first five years after initial cancer diagnosis[28–30]. In this study, the standardized incidence of SPM is therefore investigated in both groups: (i) patients that survived two or more months, and (ii) patients that survived five or more years after initial gastrointestinal cancer diagnosis.

2. OBJECTIVES

Primary Objective: To describe the patient and tumor characteristics, first course treatments, survival and subsequent primary malignancies of children and adolescents diagnosed with primary gastrointestinal malignancies.

Secondary Objective(s): 1) To investigate time trends in incidence, stage at diagnosis, first course treatments, mortality, and overall survival. 2) To explore independent prognostic factors for overall survival. 3) To calculate the standardized incidence of second primary malignancies in childhood and adolescent gastrointestinal cancer survivors, relative to the general population.

3. STUDY DESIGN

This is an observational population-based cohort study based on publicly available data of the Surveillance, Epidemiology, and End Results (SEER) database. The SEER Program collects cancer incidence and survival data from multiple population-based cancer registries in the United States of America (U.S.). SEER began collecting data in 1973 with a limited number of registries. Since 1973, the SEER Program has been expanded to cover numerous additional areas. SEER 17 is the largest registry grouping with complete data, and includes cancer cases registered from 2000 onwards in 17 cancer registries covering approximately 26.5% of the U.S. population (https://seer.cancer.gov/registries/data.html). Importantly, SEER recommends that when performing an analysis, researchers subset based on year or registry so that they have consistent geographic coverage for the period of the analysis. Accordingly, SEER does not provide datasets with varying registries and years of data (e.g. 1975-2000 data from SEER 8 registries and 2000-2019 data from SEER 17 registries). SEER 17 incidence data is used for this study, because the SEER 17 database contains the most cancer cases to investigate, when compared to SEER 12 (1992-2019) or SEER 8 (1975-2019) (https://seer.cancer.gov/data-software/documentation/seerstat/nov2021/).

The SEER Program registries routinely collect data on patient demographics, primary tumor site, tumor morphology, stage at diagnosis, first course of treatment, and follow-up for vital status. The mortality data reported by SEER are provided by the National Center for Health Statistics (NCHS). Notably, NCHS granted SEER permission to make available mortality data of the total U.S. population – based on death certificates – for the period 1969-2019. The population data used in calculating cancer rates is obtained from the Census Bureau.

4. STUDY POPULATION

4.1 Population (base)

All cases registered in the SEER 17 database between (and including) 2000 and 2019 (last data year in the April 2022 SEER data release).

4.2 Inclusion criteria

In order to be eligible for this study, a case must meet all of the following criteria:

- Age (at primary gastrointestinal cancer diagnosis) <25 years
- Diagnosis of a primary gastrointestinal malignancy as an individual's first primary malignancy, based on:
    - Primary site "digestive organs" (ICD-O-3 topography codes C015-C026)
    - Behavior code "malignant" (ICD-O-3 behavior code 3)

Selection of only those cases with a primary gastrointestinal malignancy as their first primary malignancy is done by setting the Multiple Primary Selection option on the Selection tab of SEER*Stat software to First Malignant Primary Only (Non-reported Assumed Malignant).

4.3 Exclusion criteria

None

4.4 Sample size calculation

Not applicable, given that the primary objective of this study is of descriptive nature.

5. TREATMENT OF SUBJECTS

Not applicable

5.1 Investigational product/treatment

Not applicable

5.2 Use of co-intervention (if applicable)

Not applicable

5.3 Escape medication (if applicable)

Not applicable

6. INVESTIGATIONAL PRODUCT

Not applicable

6.1 Name and description of investigational product(s)

Not applicable

6.2 Summary of findings from non-clinical studies

Not applicable

6.3 Summary of findings from clinical studies

Not applicable

6.4 Summary of known and potential risks and benefits

Not applicable

6.5 Description and justification of route of administration and dosage

Not applicable

6.6 Dosages, dosage modifications and method of administration

Not applicable

6.7 Preparation and labelling of Investigational Medicinal Product

Not applicable

6.8 Drug accountability

Not applicable

7. NON-INVESTIGATIONAL PRODUCT

Not applicable

7.1 Name and description of non-investigational product(s)

Not applicable

7.2 Summary of findings from non-clinical studies

Not applicable

7.3 Summary of findings from clinical studies

Not applicable

7.4 Summary of known and potential risks and benefits

Not applicable

7.5 Description and justification of route of administration and dosage

Not applicable

7.6 Dosages, dosage modifications and method of administration

Not applicable

7.7 Preparation and labelling of Non Investigational Medicinal Product

Not applicable

7.8 Drug accountability

Not applicable

8. METHODS

A list of data items obtained from the SEER 17 database is provided in Table 1.

Table 1. Retrieved SEER Research Plus Data Items

Data Item name	Category name
Patient ID	Other
Age recode with single ages and 85+	Race and Age (case data only) or Age at Diagnosis
Sex	Race, Sex, Year Dx, Registry, County
Race recode (White, Black, Other)	Race, Sex, Year Dx, Registry, County
Origin recode NHIA (Hispanic, Non-Hisp)	Race and Age (case data only) or Race, Sex, Year Dx, Registry, County
Race and origin recode (NHW, NHB, NHAIAN, NHAPI, Hispanic)	Race and Age (case data only) or Race, Sex, Year Dx, Registry, County
Year of diagnosis	Race, Sex, Year Dx, Registry, County
Sequence number	Multiple Primary Fields
Record number recode	Multiple Primary Fields
Total number of in situ/malignant tumors for patient	Multiple Primary Fields
Total number of benign/borderline tumors for patient	Multiple Primary Fields
First malignant primary indicator	Multiple Primary Fields
Primary by international rules	Multiple Primary Fields
Diagnostic Confirmation	Site and Morphology
Site recode ICD-O-3/WHO 2008	Site and Morphology
Site recode ICD-O-3/WHO 2008 (for SIRs)	Site and Morphology
Primary Site - labeled	Site and Morphology
Laterality	Site and Morphology
ICD-O-3 Hist/behav	Site and Morphology
Grade (thru 2017)	Site and Morphology
Grade Clinical (2018+)	Site and Morphology
Grade Pathological (2018+)	Site and Morphology
Site recode - rare tumors	Site and Morphology
ICCC site recode extended 3rd edition/IARC 2017	Site and Morphology
AYA site recode 2020 Revision	Site and Morphology
Lymphoid neoplasm recode 2021 Revision	Site and Morphology
SEER Brain and CNS Recode	Site and Morphology
TNM 7/CS v0204+ Schema (thru 2017)	Site and Morphology
CS Schema - AJCC 6th Edition	Site and Morphology
Schema ID (2018+)	Site and Morphology
AJCC ID (2018+)	Site and Morphology
EOD Schema ID Recode (2010+)	Site and Morphology
Combined Summary Stage (2004+)	Stage - Summary/Historic
Summary stage 2000 (1998-2017)	Stage - Summary/Historic
Derived EOD 2018 Stage Group (2018+)	Stage - 8th edition
Derived SEER Cmb Stg Grp (2016-2017)	Stage - 7th edition
Derived AJCC Stage Group, 7th ed (2010-2015)	Stage - 7th edition
Derived AJCC Stage Group, 6th ed (2004-2015)	Stage - 6th edition
SEER modified AJCC stage 3rd (1988-2003)	Stage - Other
Lymphoma - Ann Arbor Stage (1983-2015)	Stage - Other
Tumor Size Summary (2016+)	Extent of Disease
CS tumor size (2004-2015)	Extent of Disease
EOD 10 - size (1988-2003)	Extent of Disease - Historic
Months from diagnosis to treatment	Therapy
RX Summ--Surg Prim Site (1998+)	Therapy
Reason no cancer-directed surgery	Therapy
Radiation recode	Therapy
Chemotherapy recode (yes, no/unk)	Therapy
B Symptoms Recode (2010+)	Site-Specific Data Items
CEA Pretreatment Interpretation Recode (2010+)	Site-Specific Data Items
Survival months	Cause of Death (COD) and Follow-up
Survival months flag	Cause of Death (COD) and Follow-up
Vital status recode (study cutoff used)	Cause of Death (COD) and Follow-up
COD to site recode	Cause of Death (COD) and Follow-up
SEER cause-specific death classification	Cause of Death (COD) and Follow-up
SEER other cause of death classification	Cause of Death (COD) and Follow-up

Patients are categorized into two age groups: children (aged 0-17 years at diagnosis) and young people aged 18-24 years at primary gastrointestinal cancer diagnosis. In principle, data is analyzed and presented by 1-year age category or by age group (<18 or 18-24 years at diagnosis), except for age distributions of specific patient populations (e.g. patients with colorectal carcinoma). When data from children and young people aged 18-24 years are analyzed together, then this is clearly stated in the accompanying text. Race is categorized into White, Black, and Other. Origin is categorized into Hispanic and non-Hispanic. Together, race and origin are reported in four mutually exclusive categories: non-Hispanic White, non-Hispanic Black, non-Hispanic Other, and Hispanic (all races).

Year of diagnosis is categorized into single years or four time periods of five years (2000-2004; 2005-2009; 2010-2014; 2015-2019). Gastrointestinal malignancies are classified into different subtypes based on the Rare Cancer Classification[31]. For this classification, the SEER Rare Cancer Classification variable is used (https://seer.cancer.gov/seerstat/variables/seer/raresiterecode/). In addition, some subtypes are specified with other selections – based on International Classification of Diseases for Oncology, third edition, second revision (ICD-O-3.2) codes – because they are not (separately) classified in the Rare Cancer Classification. A complete list of cancer subgroup selections is provided in Table 2.

Table 2. Gastrointestinal cancer subtypes and their selection methods

Group	Source | Group(s) | Recode value(s)	ICD-O-3.2 Histology code(s)	ICD-O-3.2 Site code(s)
Hematologic malignancies	Rare Cancer Classification | Tumor groups 63-68 | Values 275-300	See classification	See classification
• Lymphomas	SEER Site Recode ICD-O-3/WHO 2008 Definition | Site group Lymphoma | Values 33011, 33012, 33041, 33042	See SEER website	See SEER website
□ **Burkitt lymphoma (BL)**	ICCC Recode Third Edition ICD-O-3/IARC 2017 | Site group IIc | Value 014 (extended recode)	9687	C015-C026
Sarcomas	Rare Cancer Classification | Tumor groups 50-53 | Values 202-229	See classification	See classification
• **Undifferentiated embryonal sarcoma of the liver (UESL)**	ICCC Recode Third Edition ICD-O-3/IARC 2017 | Site group VIIa.3 | Value 43 (extended recode)	8991	C022
Neuroendocrine neoplasms (NEN)	Rare Cancer Classification | Tumor group 54 | Values 230-233*	See classification	See classification
• Well differentiated, neuroendocrine tumors (NETs)	Rare Cancer Classification | Tumor groups 54.1 and 54.2 | Values 230-231	See classification	See classification
• Poorly differentiated, neuroendocrine carcinomas (NECs)	Rare Cancer Classification | Tumor group 54.3 | Value 232*	See classification	See classification
• Mixed neuroendocrine-non-neuroendocrine neoplasms (MiNEN)	Rare Cancer Classification | Tumor group 54.4 | Value 233	See classification	See classification
Carcinomas and blastomas	Rare Cancer Classification | Tumor groups 7-15, 44, 46 | Values 19-63, 190, 192	See classification	See classification
• **Hepatoblastoma (HB)**	Rare Cancer Classification | Tumor group 44 | Value 190	See classification	See classification
• Carcinomas of the liver and intrahepatic bile tract	Rare Cancer Classification | Tumor group 14 | Values 52-59*	See classification	See classification
□ **Hepatocellular carcinoma (HCC)**	Rare Cancer Classification | Tumor groups 14.1 and 14.2 | Values 52-53	See classification	See classification
• Carcinomas of the gallbladder and extrahepatic bile tract	Rare Cancer Classification | Tumor group 15 | Values 60-63*	See classification	See classification
• Pancreatoblastoma	Rare Cancer Classification | Tumor group 46 | Value 192	See classification	See classification
• Carcinomas of the pancreas	Rare Cancer Classification | Tumor group 13 | Values 43-51*	See classification	See classification
• Carcinomas of the gastrointestinal tract	Rare Cancer Classification | Tumor groups 7-12 | Values 19-42*	See classification	See classification
□ Carcinomas of the esophagus	Rare Cancer Classification | Tumor group 7 | Values 19-23*	See classification	See classification
□ **Carcinomas of the stomach (GC)**	Rare Cancer Classification | Tumor group 8 | Values 24-28*	See classification	See classification
□ Carcinomas of the small intestine	Rare Cancer Classification | Tumor group 9 | Values 29-31*	See classification	See classification
□ **Carcinomas of the colon and rectum (CRC)**	Rare Cancer Classification | Tumor groups 10 and 11 | Values 32-38*	See classification	See classification
□ Carcinomas of the anal canal	Rare Cancer Classification | Tumor group 12 | Values 39-42*	See classification	See classification
• Carcinomas of digestive organs, NOS	Custom selection*	Histology codes from Rare Cancer Classification tumor groups 7-15	C026
Rare other and unknown subtypes	Remaining cases	Remaining cases	Remaining cases

Legend Table 2. Abbreviation: NOS, not otherwise specified. Symbols: * Contrary to the Rare Cancer Classification, cancers with ICD-O-3 morphology codes 8000-8009 ("neoplasms, NOS") and ICD-O-3 topography codes C015-C026 ("digestive organs") are not classified (e.g. morphology code 8002 as NEN; morphology codes 8000-8001 and 8003-8005 as epithelial tumors). Instead, these cancers are grouped within "rare other and unknown subtypes" in our study. This decision is made because – in our study population – the likelihood that digestive cancers of unspecified histologic subtype are of neuroendocrine or epithelial origin is not necessarily highest. In our young patient population, malignancies in these anatomic locations are relatively often other cancers, such as lymphomas, sarcomas or blastomas.

Seven common gastrointestinal cancers are specifically distinguished: Burkitt lymphoma (BL), undifferentiated embryonal sarcoma of the liver (UESL), neuroendocrine neoplasms (NEN), hepatoblastoma (HB), hepatocellular carcinoma (HCC), gastric carcinoma (GC), and colorectal carcinoma (CRC). These cancers are depicted in **blue** in Table 2. Subgroup analyses of patients with these malignancies are performed, with a special focus on patients with CRC.

For some of these common cancers, different histologic subtypes are distinguished (Table 3). Grading is based on ICD-O-3 and categorized as 1) Well differentiated, Grade I; 2) Moderately differentiated, Grade II; 3) Poorly differentiated, Grade III; 4) Undifferentiated or anaplastic, Grade IV. Stage at diagnosis is categorized into localized, regional, or distant according to the SEER Summary Staging system (https://training.seer.cancer.gov/staging/systems/summary/) by using the Combined Summary Stage variable (available for cases diagnosed between 2004-2019) or the Summary Stage 2000 variable (available for cases diagnosed between 2000-2003). In addition, the Ann Arbor classification is available for staging BL (for cases diagnosed between 2000-2015), and American Joint Committee on Cancer (AJCC) staging schemes are available for NEN, HCC, GC and CRC (Table 3). In any publications of analyses involving staging data, it is clearly stated that different staging schemes were used to stage cancers diagnosed in different years (e.g. AJCC 3rd edition for CRC diagnosed in 2000, and AJCC 8th edition for CRC diagnosed in 2019) – thereby acknowledging the limitations associated with these analyses.

Regarding the first course of treatment, surgery of the primary site, radiation and chemotherapy are dichotomously captured as "yes" or "no/unknown". Similarly, total hepatectomy with liver transplant is specified as "yes" or "no/unknown" for patients with primary hepatic malignancies – including UESL, HB, and HCC (Table 3). In any publications of analyses involving treatment data, the limitations of these data items are acknowledged, as outlined on the SEER website (https://seer.cancer.gov/data-software/documentation/seerstat/nov2021/treatment-limitations-nov2021.html).

Table 3. Categorization of data items analyzed for common cancer subtypes

Subtype	Primary site classification | subgroups	Additional staging scheme | subgroups	Grading scheme | subgroups	Histologic classification | subgroups	Surgery primary site (RX Summ--Surg Prim Site (1998+) variable)	Total hepatectomy with transplant (RX Summ--Surg Prim Site (1998+) variable)	Radiation (Radiation recode variable)	Chemotherapy (Chemotherapy recode (yes, no/unk) variable)
BL	ICD-O-3 | C015-C026	Ann Arbor | stage I-IV	NR	NR	Yes (values 10-90); No or unknown (values 0 and 99)	NR	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown
UESL	NR	N/A	NR	NR	Yes (values 10-90); No or unknown (values 0 and 99)	Yes (values 61 and 75); No or unknown (other values)	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown
NEN	ICD-O-3 | C015-C017, C018.1 (appendix), C018 excl. C018.1, C019-C026	AJCC | stage I-IV	ICD-O-3 | Grade I-III	Rare Cancer Classification | **Cat. 1**: 230-231 (well differentiated) vs. **Cat. 2**: 232-233 (poorly differentiated & MiNEN)	Yes (values 10-90); No or unknown (values 0 and 99)	NR	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown
HB	NR	N/A	ICD-O-3 | Grade I-IV	NR	Yes (values 10-90); No or unknown (values 0 and 99)	Yes (values 61 and 75); No or unknown (other values)	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown
HCC	NR	AJCC | stage I-IV	ICD-O-3 | Grade I-IV	Rare Cancer Classification | **Cat. 1**: 52 (HCC, NOS) vs. **Cat. 2**: 53 (HCC, fibrolamellar)	Yes (values 10-90); No or unknown (values 0 and 99)	Yes (values 61 and 75); No or unknown (other values)	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown
GC	ICD-O-3 | C016.0-C016.9	AJCC | stage I-IV	ICD-O-3 | Grade I-IV	ICD-O-3 | **Cat. 1**: 8140 (AC, NOS), **Cat. 2**: 8210 + 8220 + 8221 + 8261 + 8263 (AC in polyp/adenoma), **Cat. 3**: 8480-8482 (MAC), **Cat. 4**: 8490 (SRCC), **Cat. 5**: remaining cases (others/unknown)[16]	Yes (values 10-90); No or unknown (values 0 and 99)	NR	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown
CRC	ICD-O-3 | C018.0-C018.9, C019.9, C020.9	AJCC | stage I-IV	ICD-O-3 | Grade I-IV	ICD-O-3 | **Cat. 1**: 8140 (AC, NOS), **Cat. 2**: 8210 + 8220 + 8221 + 8261 + 8263 (AC in polyp/adenoma), **Cat. 3**: 8480-8482 (MAC), **Cat. 4**: 8490 (SRCC), **Cat. 5**: remaining cases (others/unknown)[16]	Yes (values 10-90); No or unknown (values 0 and 99)	NR	Yes (all subtypes); No or unknown (incl. "refused" and "recommended, unknown if administered")	Yes; No or unknown

New abbreviations: NR, not relevant; N/A, not available; Cat, category; AC, adenocarcinoma; MAC, mucinous adenocarcinoma; SRCC, signet ring cell carcinoma.

8.1 Study parameters/endpoints

8.1.1 Main study parameter/endpoint

Descriptive statistics of patient and tumor characteristics, first course treatments, overall survival, and subsequent primary malignancies.

8.1.2 Secondary study parameters/endpoints

1. Time trends in incidence, stage at diagnosis, first course treatments, mortality, and overall survival.

2. Associations between the variables age at diagnosis, sex, race and origin, year of diagnosis, stage at diagnosis, and histologic category (when relevant) with overall survival in multivariable analysis.

3. Standardized incidence of second primary malignancies in patients who survived (i) two or more months, or (ii) five or more years after gastrointestinal cancer diagnosis, relative to the general U.S. population.

8.1.3 Other study parameters (if applicable)

Not applicable.

8.2 Randomization, blinding and treatment allocation

Not applicable

8.3 Study procedures

Study subjects do not undergo any procedures for this study.

8.4 Withdrawal of individual subjects

Not applicable

8.4.1 Specific criteria for withdrawal (if applicable)

Not applicable

8.5 Replacement of individual subjects after withdrawal

Not applicable

8.6 Follow-up of subjects withdrawn from treatment

Not applicable

8.7 Premature termination of the study

Not applicable

9. SAFETY REPORTING

9.1 Temporary halt for reasons of subject safety

Not applicable

9.2 AEs, SAEs and SUSARs

9.2.1 Adverse events (AEs)

Not applicable

9.2.2 Serious adverse events (SAEs)

Not applicable

9.2.3 Suspected unexpected serious adverse reactions (SUSARs)

Not applicable

9.3 Annual safety report

Not applicable

9.4 Follow-up of adverse events

Not applicable

10. STATISTICAL ANALYSIS

Analyses are restricted to the first primary gastrointestinal malignancy of included patients. Subsequent primary malignancies diagnosed among patients from the study cohort, including subsequent primary gastrointestinal malignancies, are described separately. Included patients must have a primary gastrointestinal malignancy as their first primary cancer (see Chapter 4.2: inclusion criteria). As a consequence of these decisions, not all (patients with) gastrointestinal malignancies during childhood or adolescence registered in the SEER 17 database are analyzed, which affects the results of our study. For example, calculated incidence rates of the different gastrointestinal malignancies may be slightly underestimated. However, these decisions were deliberately made to facilitate easy interpretation of the study cohort and corresponding data. For example, patients with a diagnosis of a non-gastrointestinal cancer prior to the diagnosis of a primary gastrointestinal cancer during childhood or adolescence represent a small but special group of patients, whose clinical characteristics, treatments and outcomes may be different from children and adolescents with a primary gastrointestinal malignancy as their first primary cancer – the population of interest. To gain insight in the degree of underestimation of incidence rates, additional analyses of all primary gastrointestinal malignancies registered among children and adolescents in the SEER 17 database may be performed.

For analyses that include histologic subgroups or grade data (Table 3), only cases that are microscopically confirmed are analyzed. For analyses including data regarding tumor characteristics (except primary site), treatment, survival, and/or subsequent primary malignancies, cases diagnosed with primary gastrointestinal cancers at autopsy and cases with this diagnosis based on death certificate only are excluded.

10.1 Primary study parameter(s)

Data is presented quantitatively (numerically). Continuous variables are summarized with medians and ranges, and categorical variables are summarized with frequencies and proportions. Age at diagnosis can also be summarized with frequencies and proportions, using 1-year age categories. Overall survival (OS) is estimated by the Kaplan-Meier method. OS is defined as the time from the date of diagnosis until the date of death from any cause (event) or until the date of last follow-up. Survival curves may be compared with the log-rank test. Median follow-up is estimated using the reverse Kaplan-Meier method[32].

The following descriptive statistics are minimally obtained:
Frequencies and proportions of the different gastrointestinal malignancies – specified in Table 2 – by age group (<18; 18-24 years at diagnosis) and by 1-year age categories (0-24 years). Age distributions of patients diagnosed with common gastrointestinal malignancies (BL, UESL, NEN, HCC, HB, GC, and CRC).

Stratified by age group (<18; 18-24 years at diagnosis):
Frequencies and proportions of the different gastrointestinal malignancies by primary anatomic site (ICD-O-3 topography codes C015-C026). Summarized sex and race and origin data of patients diagnosed with common gastrointestinal malignancies. Frequencies and proportions of surgery, radiation, and chemotherapy as part of the first course of treatment of patients with common gastrointestinal malignancies. Frequencies and proportions of different stages at diagnosis, different grades, and different histologic subtypes of patients diagnosed with CRC. Frequencies and proportions of CRC cases by primary anatomic site (ICD-O-3 topography codes C018.0-C18.9, C19.9, and C20.9). Kaplan-Meier curves depicting OS of patients diagnosed with common gastrointestinal malignancies. Estimated 1- and 5-year OS (with 95% confidence intervals) of patients diagnosed with common gastrointestinal malignancies. Kaplan-Meier curves depicting OS of CRC patients by i) stage at diagnosis, ii) by grade, and iii) by histologic subtype. Frequencies and proportions of patients with multiple primary cancers registered in 2000-2019 by common cancer subgroup (Table 3). Median total number of in situ/malignant tumors of patients diagnosed with common gastrointestinal malignancies.

10.2 Secondary study parameter(s)

Data will be presented quantitatively (numerically); stratified by age group (<18; 18-24 years at diagnosis).

Age-adjusted incidence and mortality rates are calculated as the average annual number of cases or deaths per million person-years at risk, respectively. For mortality rates, calendar year and age (<18; 18-24 years) represent year and age at death, respectively. Incidence rates of the common gastrointestinal malignancies are based on SEER 17 incidence data (2000-2019); mortality rates of malignant neoplasms of the digestive organs (ICD-O-3 topography codes C015-C026) are based on mortality data of the total U.S. population (1969-2019). In addition, incidence-based mortality (IBM) rates – for cases registered in SEER 17 – may be calculated to partition mortality by digestive cancer subtype (BL, UESL, NEN, HCC, HB, GC, and CRC)[33]. The annual population size is obtained from the Census Bureau. Rates are age-standardized using the age structure of the 2000 U.S. standard population. Rates are calculated using SEER*Stat software and plotted using 3-year moving averages[34]. When 3-year moving averages vary greatly, rates may be plotted using 5- or 7-year moving averages. Trends over time in incidence and mortality are analyzed by Poisson regression, with the calendar year as a continuous regressor variable. In addition, trends over time may be evaluated by linear regression modelling – a method frequently used in prior studies[34,35] – to allow easier comparisons between studies, as well as with Jointpoint regression modelling to identify change points in the trends. Results are reported as average annual percent changes (AAPC) along with 95% confidence intervals and *P* values.

Trends over time in stage at diagnosis or first course treatments of patients with common GI malignancies are tested with logistic regression, with year of diagnosis as a continuous variable[34]. The proportions of each stage at diagnosis or treatment category are plotted as weighted 3-year moving averages in percentages. When weighted 3-year moving averages vary greatly, results may be plotted using weighted 5- or 7-year moving averages.

Trends over time in OS of patients with common GI malignancies are evaluated by Cox regression analysis, with year of diagnosis as a continuous variable. The estimated 1- and 5-year OS of patients with common GI malignancies are plotted as percentages by diagnostic time period. For the diagnostic period 2015-2019, 5-year OS is not available due to insufficient follow-up. In addition, changes over time in OS of patients diagnosed with common GI malignancies may be visualized by plotting separate Kaplan-Meier curves for each diagnostic time period[16].

To explore independent prognostic factors for overall survival of patients with common GI malignancies, multivariable Cox regression analysis is performed with age at diagnosis, sex, race and origin, year of diagnosis, stage at diagnosis, and histologic category (when relevant; see Table 3) as included variables.

Table 4. Multivariable Cox regression analysis

	BL			UESL			NEN			HB			HCC			GC			CRC		
	No. at risk	HR (95% CI)	*P*	No. at risk	HR (95% CI)	*P*	No. at risk	HR (95% CI)	*P*	No. at risk	HR (95% CI)	*P*	No. at risk	HR (95% CI)	*P*	No. at risk	HR (95% CI)	*P*	No. at risk	HR (95% CI)	*P*
Age at diagnosis																					
Sex																					
Male		Reference			Reference			Reference			Reference			Reference			Reference			Reference	
Female																					
Race & origin																					
NH-White		Reference			Reference			Reference			Reference			Reference			Reference			Reference	
NH-Black																					
NH-Other/unknown																					
Hispanic																					
Year of diagnosis																					
Stage																					
Localized		Reference			Reference			Reference			Reference			Reference			Reference			Reference	
Regional																					
Distant																					
Unknown/unstaged																					
Histology		NR			NR						NR										
Category 1								Reference						Reference			Reference			Reference	
Category 2																					
Category 3																					
Category 4																					
Category 5																					

NH, non-hispanic; NR, not relevant

Standardized incidence ratios (SIR) and 95% confidence intervals of second primary malignancies (SPM) in patients that survived i) two or more months, or ii) five or more years after initial gastrointestinal cancer diagnosis are estimated using the multiple primary (MP-SIR) session of SEER*Stat software. Analyses are performed by common gastrointestinal cancer subgroup (BL, UESL, NEN, HCC, HB, GC, and CRC). SIR calculations compare the observed second cancer incidence with the expected first primary cancer incidence in the age-, sex-, race- and calendar year-matched general population[29]. The general population rates are available by cancer site, 5-year age categories, sex, race and 5-year diagnostic time periods.

A *P* value < 0.05 is considered statistically significant.

10.3 Other study parameters

Not applicable

10.4 Interim analysis (if applicable)

Not applicable

11. ETHICAL CONSIDERATIONS

11.1 Regulation statement

In analogy to studies based on data from the Netherlands Cancer Registry[34], this observational study of anonymous data does not require approval from an ethics committee in the Netherlands, according to the Central Committee on Research involving Human Subjects (CCMO).

11.2 Recruitment and consent

Not applicable

11.3 Objection by minors or incapacitated subjects (if applicable)

Not applicable

11.4 Benefits and risks assessment, group relatedness

No risk or benefits

11.5 Compensation for injury

Not applicable

11.6 Incentives (if applicable)

Not applicable

12. ADMINISTRATIVE ASPECTS, MONITORING AND PUBLICATION

12.1 Handling and storage of data and documents

Retrieved data are anonymous. Access to SEER Research Plus data is provided under the SEER Research Data Use Agreement (https://seer.cancer.gov/data-software/documentation/seerstat/nov2021/seer-dua-nov2021.html). Data are available in SEER*Stat software. Exported data are stored on a secured network drive with regular back-ups.

12.2 Monitoring and Quality Assurance

No monitoring will take place.

12.3 Amendments

Amendments will be recorded and filed by the investigators.

12.4 Annual progress report

No progress report will be made.

12.5 Temporary halt and (prematurely) end of study report

Not applicable

12.6 Public disclosure and publication policy

The investigators are committed to publishing the results of the study in an international medical scientific journal; preferably open access. All investigators (including the study statistician) are co-author of the manuscript, following ICMJE guidelines.

13. STRUCTURED RISK ANALYSIS

Not applicable

13.1 Potential issues of concern

Not applicable

13.2 Synthesis

Not applicable

14. REFERENCES

1. Ladd AP. Gastrointestinal tumors in children and adolescents. *Semin. Pediatr. Surg.* 2006;15(1):37–47.
2. Reulen RC, Wong KF, Bright CJ, et al. Risk of digestive cancers in a cohort of 69 460 five-year survivors of childhood cancer in Europe: the PanCareSurFup study. *Gut*. 2020;70(8):1520–1528.
3. Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer statistics, 2022. *CA. Cancer J. Clin.* 2022;72(1):7–33.
4. Geiger JL, Ismaila N, Beadle B, et al. Management of Salivary Gland Malignancy: ASCO Guideline. *J. Clin. Oncol.* 2021;39(17):1909–1941.
5. Sawyer SM, Azzopardi PS, Wickremarathne D, Patton GC. The age of adolescence. *Lancet. Child Adolesc. Heal.* 2018;2(3):223–228.
6. McDonagh JE, European Training Effective Care and Health Faculty A-E, Boisen KA, et al. The age of adolescence…and young adulthood. *Lancet. Child Adolesc. Heal.* 2018;2(4):e6.
7. Patton GC, Sawyer SM, Santelli JS, et al. Our future: a Lancet commission on adolescent health and wellbeing. *Lancet*. 2016;387(10036):2423–78.
8. Sawyer SM, McNeil R, Francis KL, et al. The age of paediatrics. *Lancet. Child Adolesc. Heal.* 2019;3(11):822–830.
9. Litten JB, Tomlinson GE. Liver Tumors in Children. *Oncologist*. 2008;13(7):812–820.
10. Darbari A, Sabin KM, Shapiro CN, Schwarz KB. Epidemiology of primary hepatic malignancies in U.S. children. *Hepatology*. 2003;38(3):560–6.
11. Feng J, Polychronidis G, Heger U, et al. Incidence trends and survival prediction of hepatoblastoma in children: a population-based study. *Cancer Commun*. 2019;39(1):62.
12. Lau CSM, Mahendraraj K, Chamberlain RS. Hepatocellular Carcinoma in the Pediatric Population: A Population Based Clinical Outcomes Study Involving 257 Patients from the Surveillance, Epidemiology, and End Result (SEER) Database (1973-2011). *HPB Surg.* 2015;2015:670728.
13. Pan L, Yin L, Liu X-C, Ying R-C, Kong W-C. Adult versus paediatric undifferentiated embryonal sarcoma of the liver: a SEER database analysis. *ANZ J. Surg.* 2021;91(12):2690–2694.

14. Mbulaiteye SM, Biggar RJ, Bhatia K, Linet MS, Devesa SS. Sporadic childhood Burkitt lymphoma incidence in the United States during 1992-2005. *Pediatr. Blood Cancer*. 2009;53(3):366–70.
15. Hill DA, Furman WL, Billups CA, et al. Colorectal carcinoma in childhood and adolescence: a clinicopathologic review. *J. Clin. Oncol.* 2007;25(36):5808–14.
16. Sultan I, Rodriguez-Galindo C, El-Taani H, et al. Distinct features of colorectal cancer in children and adolescents. *Cancer*. 2010;116(3):758–765.
17. Poles GC, Clark DE, Mayo SW, et al. Colorectal carcinoma in pediatric patients: A comparison with adult tumors, treatment and outcomes from the National Cancer Database. *J. Pediatr. Surg.* 2016;51(7):1061–6.
18. Tessler RA, Dellinger M, Richards MK, et al. Pediatric gastric adenocarcinoma: A National Cancer Data Base review. *J. Pediatr. Surg.* 2019;54(5):1029–1034.
19. Jongmans MCJ, Zhang J, Kuiper RP, et al. Genetic Cancer Susceptibility in Adolescents and Adults 25 Years or Younger With Colorectal Cancer. *Gastroenterology*. 2022;162(3):969-974.e6.
20. de Voer RM, Diets IJ, van der Post RS, et al. Clinical, Pathology, Genetic, and Molecular Features of Colorectal Tumors in Adolescents and Adults 25 Years or Younger. *Clin. Gastroenterol. Hepatol.* 2021;19(8):1642-1651.e8.
21. Turcotte LM, Neglia JP, Reulen RC, et al. Risk, Risk Factors, and Surveillance of Subsequent Malignant Neoplasms in Survivors of Childhood Cancer: A Review. *J. Clin. Oncol.* 2018;36(21):2145–2152.
22. Teepen JC, van Leeuwen FE, Tissing WJ, et al. Long-Term Risk of Subsequent Malignant Neoplasms After Treatment of Childhood Cancer in the DCOG LATER Study Cohort: Role of Chemotherapy. *J. Clin. Oncol.* 2017;35(20):2288–2298.
23. Friedman DL, Whitton J, Leisenring W, et al. Subsequent neoplasms in 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. *J. Natl. Cancer Inst.* 2010;102(14):1083–95.
24. Reulen RC, Frobisher C, Winter DL, et al. Long-term Risks of Subsequent Primary Neoplasms Among Survivors of Childhood Cancer. *JAMA*. 2011;305(22):2311.
25. Turcotte LM, Whitton JA, Friedman DL, et al. Risk of Subsequent Neoplasms During the Fifth and Sixth Decades of Life in the Childhood Cancer Survivor Study Cohort. *J. Clin. Oncol.* 2015;33(31):3568–75.

26. Lee JS, DuBois SG, Coccia PF, et al. Increased risk of second malignant neoplasms in adolescents and young adults with cancer. *Cancer*. 2016;122(1):116–23.
27. Turcotte LM, Liu Q, Yasui Y, et al. Temporal Trends in Treatment and Subsequent Neoplasm Risk Among 5-Year Survivors of Childhood Cancer, 1970-2015. *JAMA*. 2017;317(8):814–824.
28. Olsen JH, Möller T, Anderson H, et al. Lifelong cancer incidence in 47,697 patients treated for childhood cancer in the Nordic countries. *J. Natl. Cancer Inst.* 2009;101(11):806–13.
29. Abrahão R, Ribeiro RC, Brunson A, Keegan THM. The burden of second primary cancers among childhood cancer survivors. *Ann. Cancer Epidemiol.* 2020;4(0):7–7.
30. Youlden DR, Baade PD, Green AC, et al. Second primary cancers in people who had cancer as children: an Australian Childhood Cancer Registry population-based study. *Med. J. Aust.* 2020;212(3):121–125.
31. Gatta G, Capocaccia R, Botta L, et al. Burden and centralised treatment in Europe of rare tumours: results of RARECAREnet—a population-based study. *Lancet Oncol.* 2017;18(8):1022–1039.
32. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control. Clin. Trials*. 1996;17(4):343–346.
33. Howlader N, Forjaz G, Mooradian MJ, et al. The Effect of Advances in Lung-Cancer Treatment on Population Mortality. *N. Engl. J. Med.* 2020;383(7):640–649.
34. Reedijk AMJ, Zijtregtop EAM, Coebergh JWW, et al. Improved survival for adolescents and young adults with Hodgkin lymphoma and continued high survival for children in the Netherlands: a population-based study during 1990-2015. *Br. J. Haematol.* 2020;189(6):1093–1106.
35. Hoogendijk R, van der Lugt J, van Vuurden D, et al. Survival rates of children and young adolescents with CNS tumors improved in the Netherlands since 1990: A population-based study. *Neuro-Oncology Adv.* 2022;4(1):vdab183.