Methodology for the Health Report on Cancer 2024

Supplementary document to the report

‘Zweittumore in der Schweiz
(Seconds cancers primaires en Suisse)’

A joint project by

Nationale Krebsregistrierungsstelle
Organe national d’enregistrement du cancer
Servizio nazionale di registrazione dei tumori
National Agency for Cancer Registration

Kinderkrebsregister
Registre du cancer de l’enfant
Registro dei tumori pediatrici
Childhood Cancer Registry

On behalf of the Confederation
in collaboration with experts.

Table of Contents

1 Introduction

This methods' report provides information on the data sources, measures and statistical methodology used in the Health Report on Cancer 2024 (Gesundheitsberichterstattung über Krebs 2024), which focuses on the risk for Second Primary Cancer (SPC) after a first primary cancer (FPC). Chapter 2 describes the different data sources utilized in this Health Report and the process of cancer registration in Switzerland. Chapter 3 presents the three analysis cohorts, and the classification systems used to code cancers depending on the age at FPC diagnosis. Key definitions related to SPC are provided in chapter 4. Chapter 5 explains the data preparation process, including the construction of a unified dataset, and discusses data quality issues. Chapter 6 describes the measures used to estimate the risk of SPC and the statistical methods applied, as well as the interpretation of these measures.

2 Description of data sources

The results reported in the Health Report on Cancer 2024 are based on available data from cancer registries, which in Switzerland are organized at the cantonal level for adults and at the national level for children and adolescents. Calculating SPC risks requires the ability to identify multiple cancers in the same person, even if a SPC occurs much later or after relocation to a different canton, which necessitates the use of a reliable person identifier. Additionally, vital status - whether individuals with a prior cancer diagnosis are still alive and at risk of developing an SPC at any given time after their FPC - is essential. The report also compares SPC risk with the risk for FPC in the general population. Calculating the baseline FPC risk requires population data.

2.1 Cancer registration in Switzerland

The Health Report on Cancer 2024 is based on all available cancer registry data for cancers diagnosed between 1990 and 2019. This period predates the mandatory cancer registration introduced on January 1, 2020, under the Cancer Registration Act (CRA). Before 2020, coverage of cancer registration across Switzerland was incomplete and varied by canton.

The CRA mandates the nationwide registration of all cancers diagnosed in adults, children, and adolescents, ensuring comprehensive and uniform cancer data collection. Its primary goals include measuring the cancer burden in Switzerland and providing a foundation for advances in cancer prevention, early detection, and treatment.

2.1.1 Registration of cancers in adults

Under the CRA, the registration of cancers diagnosed in individuals aged 20 years or older falls under the jurisdiction of the cantons. All cantons are required to maintain a cantonal cancer registry (CCR), either independently or in collaboration with other cantons. New cancer diagnoses are reported to the CCR of the canton where the patient resides at the time of diagnosis. Each CCR operates with its institutional structure, collaborating closely with hospitals, pathology laboratories, and other organizations that provide information about cancer patients. In Switzerland there are thirteen CCRs, responsible for registering cancer cases among adults.

Before the CRA came into force, not all cantons maintained a CCR and existing CCRs varied in the breadth of data collected. Some cantons began registering cancer cases as early as the 1970s - 1980s, while others began much later (for more details, see the following map on the website of the Federal Statistical Office: Cantonal cancer registries: Start of data registration). In few cantons, no CCR existed prior to 2020. The data used in the current Health Report is estimated to cover 73% of the Swiss adult population and cancer cases during the 1990 - 2019 period. Despite the absence of mandatory registration during these years, the completeness of case ascertainment was shown to be high.[1]

For national cancer monitoring and reporting, CCRs submit pseudonymized data on individual cancer diagnoses to the National Agency for Cancer Registration (NACR). The NACR manages the National Cancer Dataset and performs quality checks on the submitted data. Data quality issues are reported back to the CCRs for correction and resubmission. The NACR publishes an annual report about the quality of the National Cancer Dataset on its website.

2.1.2 Registration of cancers in childhood and adolescence

Prior to the CRA, childhood cancer diagnoses were recorded by the former Swiss Childhood Cancer Registry (SCCR), which was established in 1976.[2] Initially focusing on children enrolled in clinical trials, the SCCR expanded registration to achieve high nationwide coverage of cancers diagnosed under the age of 16 years by the late 1980s. By 1995, its coverage was estimated to exceed 95%.[3] The SCCR also collected data on cancers diagnosed at the age of 16-19 years. However, coverage for this age group was less complete and relied heavily on data exchanges with existing CCRs.

Since 2020, under the CRA, the Childhood Cancer Registry of Switzerland (ChCR) is responsible for recording all cancers diagnosed among children and adolescents under the age

of 20 years. Data from the former SCCR have been integrated into the ChCR dataset. The ChCR transfers data of new cases to CCRs, ensuring that their cantonal datasets cover all age groups.

The ChCR publishes basic data quality indicators on its website. During the 1990 - 2019 period covered by the Health Report on Cancer 2024, these show that data quality was somewhat lower for cancer diagnosis in the adolescent group (age 15-19 years) than among children (age 0-14 years) (see Data Quality Indicators published on the ChCR's website).

2.1.3 Unique person and case identifiers

The Old Age and Survivors' Insurance (OASI; German: AHV, French/Italian: AVS) number is a personal identifier assigned to all residents of Switzerland. Since the enactment of the CRA, Swiss cancer registries are required to record the OASI number for all newly registered cancers. The OASI number is pseudonymized and transmitted to the NACR, where it is essential for constructing and maintaining the National Cancer Dataset. This pseudonymized identifier enables:

1. Identification and removal of duplicate registrations by different cancer registries.
2. Accurate linkage and chronological ranking of cancer diagnoses when a patient moves between cantons (e.g., FPC diagnosed in Geneva, SPC diagnosed in Bern).

Before the CRA, the use of the OASI number by the cancer registries was not formally established. With the new law, registering the OASI number became mandatory. This also applies for cases predating the CRA. As a result, since 2003, over 95% of registered cases include this identifier. For cases diagnosed between 1990 and 2002, the proportion varies from 71% to 100% across years and cantons.

The individual cancer registries (CCRs and ChCR) maintain their own personal identifiers in addition to the OASI number. These can be used to uniquely idenitfy a person within a registry but not across the registries.

In order to distinguish between different cancer cases, a National Case Identifier (NCID) is created for each new diagnosis of a patient. The NCID is a unique number automatically generated at the point of registration by the cancer registration softwares and is linked to the OASI number of the patient. Since 2020, all NCIDs and the respective pseudonymized OASI number are included in a national information system (InSy). As a result, the same patients can be identified across the datasets of the CCRs and the ChCR. However, cancer cases registered in different registries (i.e., ChCR and CCR, or in two different CCRs) are assigned different

NCIDs. Therefore, when combining data on cases before 2020, a retrospective deduplication of cases is necessary (e.g., by comparing diagnostic information).

2.1.4 Vital status and follow-up information

Since 2020, all CCRs and the ChCR have been able to update the vital status information of registered persons semi-automatically through the Central Compensation Office (CCO) based on the individual's OASI number. This update applies not only to newly diagnosed cancer cases but also to earlier cases, provided an OASI number has been registered. Before 2020 (and still today CCRs were also able to update the vital status information by accessing cantonal inhabitant control data (active follow-up). Individuals whose vital status could not be determined at time of follow-up - e.g., because their OASI number was not registered or was incorrect (automatic vital status update through CCO not possible), or they moved out of the CCR region (active vital status update not possible) - are classified as "lost to follow-up", with the last known date alive retained. The proportion lost to follow-up is around 2% among adult cancer cases in the National Cancer Dataset between 2004 and 2018, and another 2% have an incomplete follow-up information meaning that active follow-up is lacking. For the ChCR, the proportion of those lost to follow-up is less than 2% among patients diagnosed between 1988 and 2017.

2.2 Swiss population data

End-year data of the permanent resident Swiss population were obtained from the Federal Statistical Office (FSO) for each year between 1990 - 2019 by age (single years), canton and sex. This data was used for the calculation of the incidence of FPC in the general population (see section 7.1).

3 Definition and overview of cohorts

In the Health Report on Cancer 2024, we examine the risk of SPC in three cohorts of cancer patients (groups of persons considered), categorized by the age at diagnosis of their FPC:

I. **Childhood cohort**: FPC diagnosed between ages 0-14 years
II. **Adolescents and young adults (AYA) cohort**: FPC diagnosed between ages 15-39 years
III. **Later adulthood cohort**: FPC diagnosed at the age of 40 or older.

Conducting separate analyses for these three cohorts is appropriate, as they differ significantly in both the frequency and types of FPCs, as well as in the common treatment modalities used.

Based on previous studies, we anticipated that the relative risk for SPCs would be higher after a FPC diagnosed at a younger age compared to an older age.[4,5]

The cohorts were defined as all individuals with a registered diagnosis of a FPC in the respective age window between 1990 and 2019. We identified a SPC as a further registered primary cancer after the FPC (second in sequence) in the same person during the same 1990-2019 period. Thus, persons were followed up for a maximum of 30 years. We used all available cancer registry data to identify FPCs and SPCs at any given age. For individuals under 20 years at FPC diagnosis, we thus combined data from the ChCR and the National Cancer Dataset from the NACR, while for older individuals, only data from the NACR were used (**Table 1**). As the distribution of cancer types varies significantly with age, we applied age-adapted cancer classification systems to categorize FPCs in the 3 cohorts (**Table 1**). The subsequent chapters provide detailed descriptions of data preparation steps, including cancer classification, the definition of SPC, and the creation of a harmonized dataset integrating all cancers from the ChCR and the NACR to facilitate the identification of FPCs and the SPCs in the 3 cohorts.

Table 1. Age ranges, source of registry data and cancer classification system used for FPCs and SPCs in the 3 cohorts.

Cohort	**Childhood**	**AYA**		**Later adulthood**
Age at FPC	0-14	15-19	20-39	≥ 40
Age at SPC (FU period)	0-44	15-69		≥ 40
Data source	ChCR/NACR	ChCR/NACR	NACR	NACR
Cancer classification FPC	ICCC-3	ICD-O-3 & Barr et al. 2020		ICD-10 & SCR 2021
Cancer classification SPC	ICD-10 & SCR 2021			

Abbreviations: AYA adolescents and young adults; **FPC** first primary cancer; **SPC** second primary cancer; **FU** follow-up; **ChCR** Childhood Cancer registry; **NACR** National Agency for Cancer registration; **ICCC-3** International Classification of Childhood Cancer, 3rd edition; **ICD-10** International Classification of Diseases, 10th revision; **SCR** Swiss Cancer Report.

4 Classification of cancers

The Swiss cancer registries (CCRs and ChCR) adhere to the recommendations of the International Agency for Research on Cancer (IARC), the International Association of Cancer Registries (IACR), and the European Network of Cancer Registries (ENCR) for the registration and coding of cancer cases. In accordance with these recommendations, the following

classification systems have been used to code cancer cases included in this Health Report: the 10th revision of the International Classification of Diseases (ICD-10),[6] the International Classification of Diseases for Oncology (ICD-O) versions 3.0-3.2,[7] and the 3rd version of the International Classification of Childhood Cancer (ICCC-3).[8]

4.1 ICD-10

The International Classification of Diseases (ICD) is an internationally standardised classification developed and maintained by the World Health Organization (WHO). It includes known health problems, diseases, injuries, symptoms, and other factors related to health status, grouping them into major categories. Since the first international classification at the end of the 19th century (1893), the ICD has been revised regularly to reflect advances in medicine. The 10th revision was released in May 1990 and has been officially used in Switzerland since 1995.[9] The classification of neoplasms according to ICD-10 is mainly based on their localization and behavior.[6]

4.2 ICD-O-3

The International Classification of Diseases for Oncology (ICD-O) is a specialized classification system used specifically to code cancers of any behavior (i.e. benign, uncertain/unknown, in-situ, or malignant). Currently the 3rd version (ICD-O-3) is in use worldwide. Unlike ICD-10, ICD-O has a two-axis structure, focusing on the topography (localization) of the cancer on one hand and the morphology (cellular characteristics) of the cancer on the other. The morphology code consists of six digits: the first four digits indicate the morphological type of the cancer, the fifth digit indicates its behavior, and the sixth digit classifies the differentiation grade for solid tumors and specifies the cell type for lymphomas and leukemias.[7]

4.3 ICCC-3

The ICCC-3 is a classification system tailored to childhood and adolescent cancers and is based on ICD-O-3. Children and adolescents experience different types of cancer than adults. The most frequent adult cancers are carcinomas (malignant neoplasms of epithelial origin) such as breast, lung, prostate and colon cancer, and are primarily classified according to their localization (see ICD-10 above). In contrast, carcinomas are rare in children and adolescents. Their tumors mostly arise from immature embryonic tissue and, as a result, primarily classified by their morphology rather than their localization.[8]

4.4 Classification of first and second primary cancers in this report

In this Health Report, only malignant SPCs were considered. These were coded according to ICD-10 and classified into 26 cancer localization groups following the Swiss Cancer Report 2021[9]: cancer of the oral cavity and pharynx, oesophageal cancer, stomach cancer, colorectal cancer, liver cancer, cancer of the biliary tract, pancreatic cancer, laryngeal cancer, lung cancer, pleura mesothelioma, melanoma of skin, breast cancer, cervical cancer, cancer of the corpus uteri, ovarian cancer, prostate cancer, testicular cancer, kidney cancer, bladder cancer, brain and Central Nervous System (CNS) cancer, thyroid cancer, Hodgkin lymphoma, Non-Hodgkin lymphoma, multiple myeloma, leukemia, other and unspecified cancers. The latter groups include ICD-10 codes ranging between C00-C43 and C45-C97.

FPCs in the childhood cohort were coded according to ICCC-3 and classified into the 12 main ICCC-3 cancer groups.[8] FPCs in the AYA cohort were coded according to ICD-O versions 3.0-3.2[7] and classified according to the tables 2-8 provided in Barr et al. (2020).[10] The classification by Barr et al. is tailored to the spectrum of cancers typically seen in the AYA age group. In both the childhood and AYA cohorts, certain non-malignant cancers were included as FPCs, as they are an integral part of these classifications. These were primarily non-malignant CNS tumors (main ICCC-3 group 3 for the childhood cohort, Barr's categories 3, 6.1, and 7.3 for the AYA cohort).

For the later adulthood cohort, only malignant cancers were included as FPCs. These were coded using ICD-10 and classified into the 26 cancer localization groups in line with the Swiss Cancer Report 2021[9] as was done for the SPCs. Analyses including FPCs of all age groups together included only malignant FPCs, incorporating the 26 cancer localization groups as in the Swiss Cancer Report 2021.[9]

5 Definition of second primary cancers (SPCs)

5.1 Distinguishing primary from non-primary cancers

A SPC is a newly developed cancer with different localization or morphology than the FPC. SPCs must thus be distinguished from recurrences, progressions, or metastases of the FPC. To identify SPCs, we used the International Rules for Multiple Primary Cancers (ICD-O Third Edition). These rules resulted from the joined work of the IARC, the WHO, the IACR, and the ENCR and are implemented in the software IARCcrgTools_v2.13. According to the report "Cancer Incidence in Five Continents"[11], the majority of cancer registries worldwide use these

rules to define multiple primary cancers. We are aware that this choice might cause comparability issues with reports and studies using different rules, e.g., with studies from the USA where multiple primary cancers are defined according to Surveillance Epidemiology and End Results (SEER) rules.[12–14] Such comparability issues have been examined and discussed in detail elsewhere.[5,15,16]

According to the International Rules for Multiple Primary Cancers (ICD-O Third Edition), a colorectal cancer diagnosed at the age of 70 years after a first primary breast cancer at the age of 50 years is a SPC. On the other hand, metastases diagnosed in the bones after a first primary breast cancer do not denote a SPC. A breast cancer diagnosed on the other side of the breast as the first primary breast cancer is not a SPC, if it is of the same morphology as the FPC.

5.2 Synchronicity period

SPCs can be either synchronous or metachronous cancers. Synchronous cancers occur shortly, e.g., within 4 months, after diagnosis of FPC, whereas metachronous cancers occur later. Synchronous cancers are usually, but not always, asymptomatic cancers detected during the thorough examination a patient undergoes as investigation for the FPC and might have not been detected otherwise.[5] To control the potential presence of detection bias, synchronous cancers are excluded in studies estimating the risk of SPC.[11,12,17–20]

There is no universally defined cut-off that separates synchronous and metachronous cancers. Cut-offs ranging from 1 up to 12 months after the FPC have been used in studies,[11,12,17–20] while the study by Baicry et al proposed a 4-month cut-off as the most appropriate, based on the instantaneous incidence rate of SPC within one year after the FPC diagnosis.[21] For this Health Report on Cancer 2024, we followed Baicry et al.'s recommendation and used a 4-month cut-off to separate synchronous and metachronous cancers.[21]

6 Creating a unified analysis dataset

6.1 Inclusion and exclusion criteria

We included all persons with a registered diagnosis of cancer based on the combined data obtained from the NACR and the ChCR, grouping them into the 3 cohorts based on age at diagnosis of FPC as outlined above.

We excluded in situ cancers, and non-melanotic skin cancer (ICD-10 code: C44), as they were not systematically registered by many registries before 2020. Non-melanotic skin cancer is

relatively common and usually nonfatal. Cancer registries around the world have different practices for recording non-melanotic skin cancer. In the interest of making international comparison of cancer more valid it is good practice to exclude non- melanotic skin cancer from the analyses.[22] Additionally, we excluded cancers of unknown sex, topography and morphology, and cancers where the combination of those variables were not plausible (e.g. cancer located in testis diagnosed in a woman).

As mentioned in section 4.4, the International Rules for Multiple Primary Cancers (ICD-O Third Edition) were used to identify SPCs and to exclude subsequent primary cancers from recurrences, progressions, or metastases of the FPC. Synchronous cancers, i.e., SPCs that occurred within the first 4 months after FPC diagnosis, were excluded (see section 5.2). By choosing a 4-month cut-off to separate synchronous and metachronous cancers, we assumed that the time at risk for a SPC begins 4 months after FPC diagnosis and lasts until the date of SPC diagnosis, the date of death, the latest date of available follow-up, or the end of follow-up (31st Dec. 2019), whichever came first. This means that patients with less than 4 months of follow-up were also excluded.

6.2 Combining NACR and ChCR datasets and identification of SPCs

To facilitate the identification of SPCs occurring in adulthood after a FPC in childhood or adolescence and apply the inclusion and exclusion criteria consistently across cohorts, the cancer records from the ChCR and from the NACR were combined into a unified dataset. This essentially involved two steps:

1) **Deduplication of cancer cases among persons aged < 20 years at FPC**: For these persons, we included cancer records from both the NACR and the ChCR, prioritizing ChCR data in cases of duplicate records. Despite the large overlap, these sources differ in their cancer records for the years prior to enactment of the CRA. This is because the NACR data are based on data registered in the CCRs (see section 2.1.3), which recorded cancers in this age group separately to the ChCR, but did not have nationwide coverage. On the other hand, the ChCR missed some diagnoses, which in turn were registered by the CCR and available in the NACR data set. To deduplicate the cancer cases from these two sources and identify any cancer cases contained in the NACR dataset but not in the ChCR, we obtained a list of NCIDs from InSy of all cancer cases recorded in a CCR that corresponded to a person (based on the pseudonymized OASI number) with a cancer record in the ChCR. We then used demographic (sex, month and year of birth) and diagnostic information

(month and year of diagnosis, canton of residence at diagnosis, diagnostic codes) from the NACR on each of these cases to decide whether the cases were already contained in the ChCR or needed to be additionally included. The ChCR data were then combined with these additional cases to form a deduplicated dataset including cancers in persons with a FPC at age < 20 years. These were combined with all other cancer cases from the NACR (FPC at age ≥ 20 years) into a unified dataset.

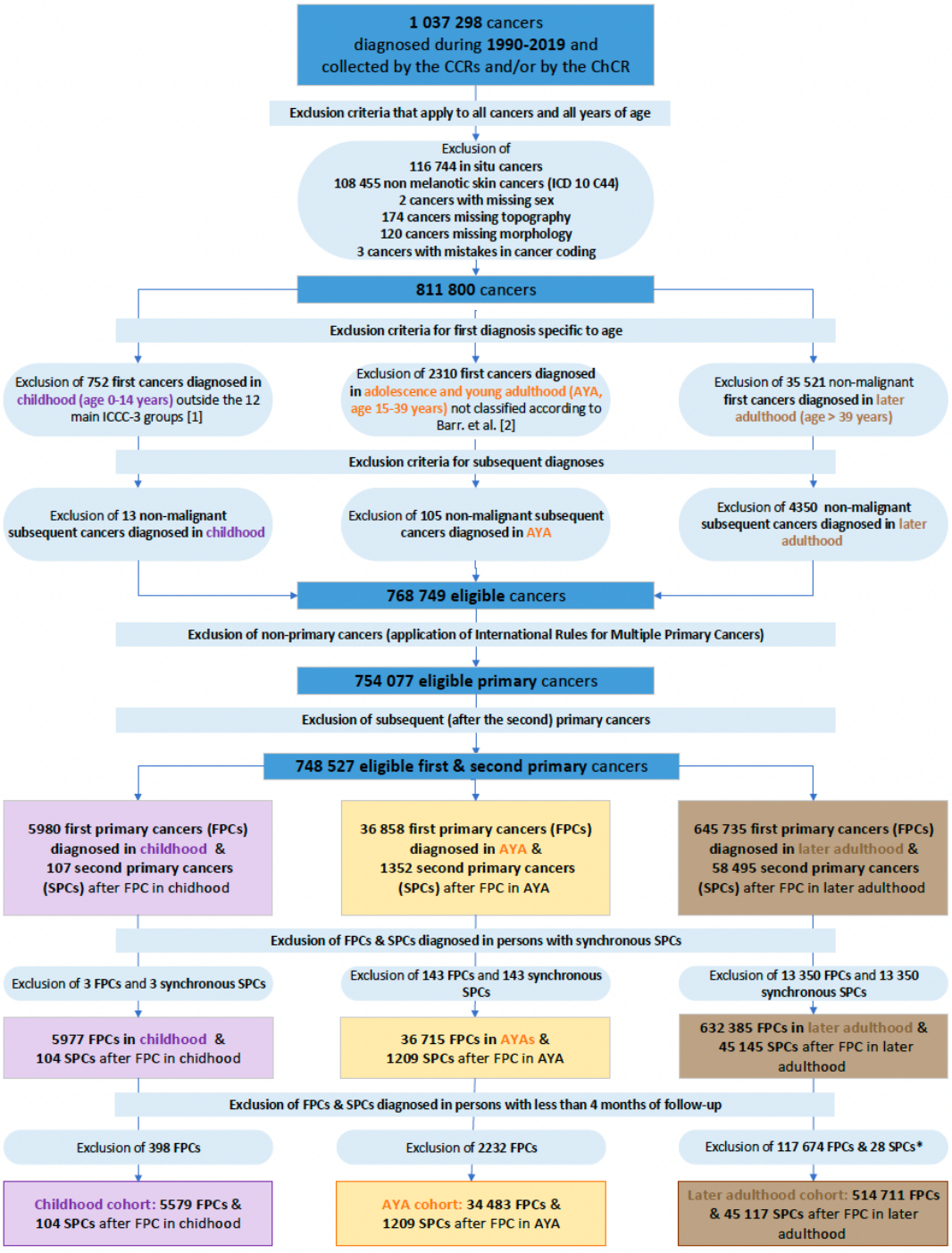
2) **Identification of SPCs**: In this unified dataset, we identified cancer diagnoses belonging to the same persons using the OASI number (based on the pseudonymized OASI available to NACR or the InSy NCIDs list described in step 1). For cancer cases not linkable via OASI number (i.e., with missing OASI number), we used registry specific personal identifiers. In accordance with the previous chapter, we then used the IARCcrgTools_v2.13 software to distinguish primary from non-primary cancers and ranked primary cancers excluding synchronous cancers to identify SPCs.

6.3 Number of FPCs and SPCs included

Of 1 037 298 cancers included in the unified NACR and ChCR dataset, we identified 754 077 eligible primary cancers and finally included a total of 554 773 FPCs and 46 430 SPCs in the analyses. A detailed account of the inclusion and exclusion steps for each of the 3 cohorts is given in the flowchart below (

Figure 1).

Figure 1. Flowchart of included first and second primary cancers.



[1] Steliarova-Foucher E, Stiller C, Lacour B, Kaatsch P. International Classification of Childhood Cancer, Third Edition. *Cancer* 2005;103:1457-67.

[2] Barr, R.D., Ries, L.A., Trama, A., Gatta, G., Steliarova-Foucher, E., Stiller, C.A. and Bleyer, W.A., 2019. A system for classifying cancers diagnosed in adolescents and young adults. *Available at SSRN 3502376*.

* 28 additional SPCs dropped either due to wrong date of SPC diagnosis or due to wrong date of maximum follow-up.

6.4 Data limitations

The cancer data used for this Health Report is subject to the same general data quality limitations inherent to cancer registry data. These include incomplete case ascertainment, coding inaccuracies, and other potential data errors. Nevertheless, case ascertainment in Swiss cancer registries has been estimated to be high,[1,3] suggesting that over 90% of eligible primary cancers diagnosed during the study period of this report (1990 - 2019) were likely captured.

The validity of our results hinges on accurately identifying multiple cancers in the same individual. Having a national person identifier, as the OASI number, is critical for achieving this with high accuracy. However, not all cases have a registered OASI number. For cases diagnosed since 2003, over 95% have a recorded OASI number (see Annual Data Quality Reports by the NACR), but for diagnoses between 1990 and 2002, the proportion was less than 80% for some years and cantons.

The absence of OASI numbers for some cases has two main implications: first, duplicate registrations in different cantons of the same diagnosis are likely to have gone undetected, causing the same cancer diagnosis to appear multiple times in the National Cancer Dataset. If the diagnosis was a FPC, it would likely have been recorded in duplicate as two FPCs. If the diagnosis was a SPC following a FPC registered in one of the cantons, then the SPC would have been correctly classified in that canton but misclassified as a new FPC in the other canton. Both scenarios result in inflated FPC counts. Second, two distinct cancers diagnosed in the same patient but registered by different CCRs would not have been linked. Consequently, a person who should have appeared once with a FPC and a SPC may instead have appeared as two separate individuals, each with one FPC diagnosis.

Another problem is that of missed FPCs: a FPC that occurred prior to the start of cancer registration in the patient's canton of residence at diagnosis would have been missed, because there was no CCR to register it. A SPC occurring in the same patient after the start of cancer registration, i.e. after the foundation of the CCR, would consequently be registered and included in our analysis as a FPC.

These issues may have affected the estimation of SPC risk. Duplicate registrations and missed FPCs could inflate the proportion of individuals with FPCs but no SPC, while unlinked cancers could lead to an under-ascertainment of SPCs. Together, these factors likely introduce to a downward bias in SPC risk estimates. However, as linkage via the OASI number was possible

for the vast majority of patients and these events are expected to involve only a small subset of remaining cases, we consider this bias to be minor.

7 Statistical methods

We use standardised incidence ratios (SIRs) as a measure or relative SPC risks and the cumulative incidence function to estimate the absolute risk of developing a SPC up to a given time after diagnosis of a FPC. Below, we provide definitions and interpretations of these measures, and outline procedures used for their estimation.

All data preparation steps and statistical analyses were performed in the R programming language version 4.4.0 (2024-04-24),[23] using packages dplyr,[24] popEpi,[25] and survival. The analyses' scripts can be requested from the project team.

7.1 Standardised incidence ratio (SIR)

Definition: The SIR is a measure used to compare the incidence of a disease in a population to that in a reference population, while controlling for other factors related to the disease (typically age and sex). It is calculated as the ratio of the observed number of cases in the population of interest to the number expected based on incidence rates in a reference population:

$$SIR = \frac{Observed\ cases}{Expected\ cases}$$

In our situation:

- **Observed cases** refer to the number of SPCs identified in the cohort under study.
- **Expected cases** are estimated by multiplying age-, sex-, calendar period (5-year periods during 1990-2019) specific cancer incidence rates for developing FPCs in the general population to the person-time at risk in the cohort.

Interpretation: An SIR of 1 means that patients in the cohort experience the same cancer risk as the persons in the general population. An SIR greater than 1 indicates that the incidence of SPCs in the cohort is higher than expected based incidence of FPCs in the general population, while an SIR less than 1 suggests a lower incidence than expected.

Methodology: To calculate SIRs, we transformed the unified national dataset into a table containing multiple rows per person. Each row corresponding to an individual was specific to a

combination of calendar year j and age k (single years). We calculated the time at risk t_{ikj} contributed by individual i to each combination of year and age, starting at FPC diagnosis and ending time at SPC diagnosis, the end of 2019, or last date of available follow-up, whichever came first. Only year and age combinations with non-zero person time at risk ($t_{kj} > 0$) were included as rows. We also defined an event indicator d_{ikj} taking on the value 1 if a SPC was diagnosed in individual i in year j and age k. The table also contained patient characteristics needed to define strata for stratified SIR analyses (cohort, FPC cancer type, FPC cancer treatment). We also calculated age-, sex, and calendar period specific incidence rates p_{skj} for FPCs in the general population using standard methods for routine cancer statistics in Switzerland (see Statistical Methods for Cancer reporting in Switzerland, pages 9-13) and merged these into the table by age, sex, and calendar period.

For a given combination of patient characteristics S (defines a stratum of individuals $i \in S$, e.g. all persons in a specific cohort, or persons in a specific cohort treated with chemotherapy only for their FPC), we calculated the stratum-specific SIR as

$$SIR_S = \frac{\sum_{i \in S} \sum_{kj} d_{ikj}}{\sum_{i \in S} \sum_{kj} t_{ikj} \cdot p_{s_i kj}}$$

where s_i denotes the sex of individual i and the numerator and denominator correspond to observed and expected number of SPCs in stratum S. Here summations over kj are over all combinations of years and ages with positive person year contributions. When computing SIRs for specific SPC types, incidence rates for corresponding FPC types in the general population were plugged in as p_{skjc}, with c denoting the cancer type.

Confidence intervals (CIs) for SIRs were calculated assuming that the observed number of SPCs (numerator of SIR, O_S) is Poisson distributed and ignoring the variation in the number of expected SPCs (denominator of SIR). Exact 95% CIs were first obtained for O_S using the relationship between Poisson and Chi-square distribution[28],

$$\frac{1}{2} X^2_{2 \cdot O_S, 0.025} \leq O_S \leq \frac{1}{2} X^2_{2 \cdot O_S + 2, 0.975},$$

where $X^2_{f,l}$ is the lth quantile of the Chi-Square distribution with f degrees of freedom. To obtained CIs for SIRs, these boundaries were divided by the denominator of the SIR.

7.2 Cumulative incidence function (CIF)

In the Health Report on Cancer 2024, we also present estimates of the absolute risk of developing a second primary cancer (SPC) within a specified time frame, up to a maximum of 30 years after the diagnosis of a first primary cancer (FPC). These estimates account for other potential events that could prevent the occurrence of an SPC - commonly referred to as competing risks. Such risks are appropriately addressed using the framework of cumulative incidence functions (CIF). For the Health Report on Cancer 2024, we consider death prior to SPC as the only competing event.

Definition: The CIF quantifies the event-specific risk up to time t and is defined as the joint probability of the event time T and the type of event X_T, i.e.

$$CIF_j(t) = \Pr\left[T \leq t, X_T = j\right]$$

with j being either SPC or death, whichever came first.

Interpretation: The CIF for SPC represents the proportion of individuals diagnosed with a FPC who are expected to develop an SPC within a specified time frame following their FPC diagnosis.

Methodology: We estimated the CIF using the Aalen-Johansen estimator[29], which is the competing risks analog to the Kaplan-Meier estimator for observed survival:

$$CIF_j(t) = \sum_{q:\, t_q \leq t} \left[\prod_{q:\, t_q \leq t} \left(1 - \frac{d_q}{n_q}\right)\right] \frac{d_{q_j}}{n_q},$$

where $0 < t_1 < t_2 < \cdots < t_l$ are the distinct observed event times, d_{q_j} is the number of observed jth events at time t_q, d_q is the number of total observed events at t_q, and n_q is the number of subjects at risk of experiencing one of the events $\mathrm{j} = 1, 2$ (here death or SPC) just prior to t_q, with $\mathrm{q} = 1, 2, \ldots, \mathrm{l}$.

8 Bibliography

1. Lorez, M. *et al.* Evaluation of completeness of case ascertainment in Swiss cancer registration. *Eur. J. Cancer Prev. Off. J. Eur. Cancer Prev. Organ. ECP* **26 Joining forces for better cancer registration in Europe**, S139–S146 (2017).
2. Michel, G. *et al.* The Swiss Childhood Cancer Registry: rationale, organisation and results for the years 2001-2005. *Swiss Med. Wkly.* **137**, 502–509 (2007).
3. Schindler, M. *et al.* Death certificate notifications in the Swiss Childhood Cancer Registry: assessing completeness and registration procedures. *Swiss Med. Wkly.* (2015) doi:10.4414/smw.2015.14225.
4. Feller, A. *et al.* The relative risk of second primary cancers in Switzerland: a population-based retrospective cohort study. *BMC Cancer* **20**, 51 (2020).
5. AIRTUM Working Group. Italian cancer figures, report 2013: Multiple tumours. *Epidemiol. Prev.* **37**, 1–152 (2013).
6. *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. (World Health Organization, Geneva, 1992).
7. *International Classification of Diseases for Oncology: ICD-O*. (World Health Organization, Geneva, 2000).
8. Steliarova-Foucher, E., Stiller, C., Lacour, B. & Kaatsch, P. International Classification of Childhood Cancer, third edition. *Cancer* **103**, 1457–1467 (2005).
9. Bundesamt für Statistik. *Schweizerischer Krebsbericht 2021*. https://www.bfs.admin.ch/asset/de/19305696 (2021).
10. Barr, R. D. *et al.* A system for classifying cancers diagnosed in adolescents and young adults. *Cancer* **126**, 4634–4659 (2020).
11. *Cancer Incidence in Five Continents.* vol. 9 (International Agency for Research on Cancer, Lyon, 2008).
12. Curtis, RE. *et al.* New malignancies among cancer survivors: SEER cancer registries, 1973-2000. *J. Epidemiol. Community Health* **62**, 375–376 (2008).
13. American Cancer Society. *Cancer Facts and Figures 2009. Special Section: Multiple Primary Cancers*. (2009).
14. Qiao, B. *et al.* Multiple Primary Cancers in the United States. *J. Regist. Manag.* **47**, 60–66 (2020).
15. Ess, S. *et al.* Geographic variation in breast cancer care in Switzerland. *Cancer Epidemiol.* **34**, 116–121 (2010).
16. Weir, H. K., Johnson, C. J., Ward, K. C. & Coleman, M. P. The effect of multiple primary rules on cancer incidence rates and trends. *Cancer Causes Control CCC* **27**, 377–390 (2016).
17. Teppo, L., Pukkala, E. & Saxén, E. Multiple cancer--an epidemiologic exercise in Finland. *J. Natl. Cancer Inst.* **75**, 207–217 (1985).
18. Storm, H. H. *et al.* Summary: multiple primary cancers in Denmark, 1943-80. *Natl. Cancer Inst. Monogr.* **68**, 411–430 (1985).
19. Coleman, M. P. Multiple primary malignant neoplasms in England and Wales, 1971-1981. *Yale J. Biol. Med.* **59**, 517–531 (1986).
20. Levi, F. *et al.* Multiple primary cancers in the Vaud Cancer Registry, Switzerland, 1974-89. *Br. J. Cancer* **67**, 391–395 (1993).
21. Baicry, F. *et al.* What is the most appropriate period to define synchronous cancers? *Cancer Epidemiol.* **71**, 101900 (2021).

22. Why are data for non-melanoma skin cancer sometimes excluded? | National Cancer Registry Ireland. https://www.ncri.ie/faqs/interpret-use-our-data/why-are-data-non-melanoma-skin-cancer-sometimes-excluded.
23. R Core Team. R: A Language and Environment for Statistical Computing. (2019).
24. Wickham H, François R, Henry L, Müller K, & Vaughan D. _dplyr: A Grammar of Data Manipulation_. R package version 1.1.4. (2023).
25. Miettinen J & Rantanen M. _popEpi: Functions for Epidemiological Analysis using Population Data_. R package version 0.4.12. (2024).
26. Therneau, T. M. & Grambsch, P. M. *Modeling Survival Data: Extending the Cox Model*. (Springer, New York, 2000).
27. Therneau T. A Package for Survival Analysis in R. (2024).
28. Sahai, H. & Khurshid, A. Confidence Intervals for the Mean of a Poisson Distribution: A Review. *Biom. J.* **35**, 857–867 (1993).
29. Aalen, O. & Johansen, S. An Empirical Transition Matrix for Non-homogeneous Markov Chains Based on Censored Observations. *Scand. J. Stat.* 141–150 (1978).