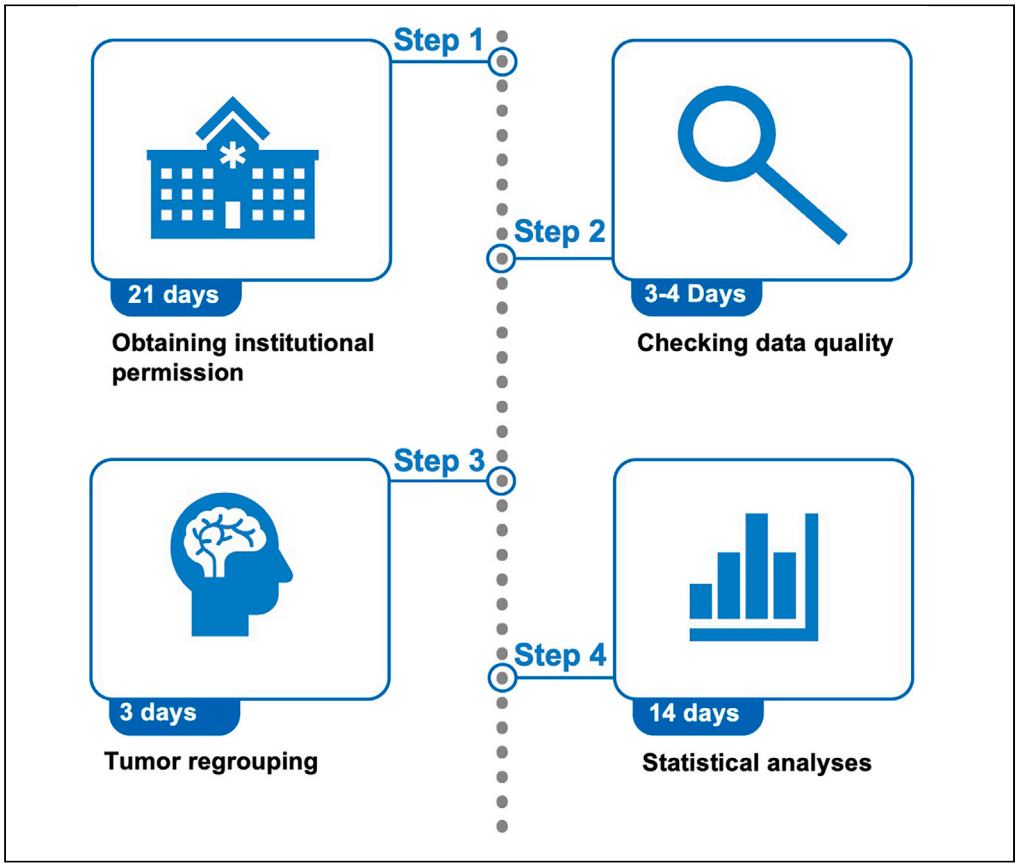
STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for investigating data quality and reporting outcomes of pediatric gliomas in population-based cancer registry research



Cancer registry data on pediatric gliomas come with inherent limitations as inclusion criteria and registration practices of these tumors differ between registries due to specific guidelines that are lacking. These limitations can lead to biased estimates in incidence and survival outcomes. Here, we present a protocol to investigate data quality and comparability for retrospective population-based pediatric glioma studies. We describe steps for obtaining institutional permissions, dealing with data quality issues, regrouping tumors, and reporting tumors in a clinically relevant manner.

Publisher's note: Undertaking any experimental protocol requires adherence to local institutional guidelines for laboratory safety and ethics.

Raoull Hoogendijk, Jasper van der Lugt, Mariëtte E.G. Kranendonk, ..., Otto Visser, Dannis G. van Vuurden, Henrike Karim-Kos

r.hoogendijk@prinsesmaximacentrum.nl (R.H.)
h.e.karim-kos@prinsesmaximacentrum.nl (H.K.-K.)

Highlights

Steps on recognizing data quality issues in registry data on pediatric gliomas

Steps described for grouping pediatric gliomas

Steps on how to deal with misclassified tumors

Hoogendijk et al., STAR Protocols 5, 102905
March 15, 2024
https://doi.org/10.1016/j.xpro.2024.102905

STAR Protocols

CellPress
OPEN ACCESS

Protocol

Protocol for investigating data quality and reporting outcomes of pediatric gliomas in population-based cancer registry research

Raoull Hoogendijk,[1,8,9,*] Jasper van der Lugt,[1] Mariëtte E.G. Kranendonk,[1] Gemma Gatta,[2] Riccardo Capocaccia,[3] Eelco W. Hoving,[1,4] Pieter Wesseling,[1,5] Otto Visser,[6] Dannis G. van Vuurden,[1,7] and Henrike Karim-Kos[1,6,7,*]

[1]Princess Máxima Center for Pediatric Oncology, Utrecht, the Netherlands

[2]Evaluative Epidemiology Unit, Department of Epidemiology and Data Science, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[3]Editorial Board, Epidemiol Prev, Milan, Italy

[4]Department of Neurosurgery, University Medical Center Utrecht, Utrecht, the Netherlands

[5]Department of Pathology, Amsterdam University Medical Centers, Amsterdam, the Netherlands

[6]Department of Research and Innovation, Netherlands Comprehensive Cancer Organization (IKNL), Utrecht, The Netherlands

[7]Senior author

[8]Technical contact

[9]Lead contact

*Correspondence: r.hoogendijk@prinsesmaximacentrum.nl (R.H.), h.e.karim-kos@prinsesmaximacentrum.nl (H.K.-K.)
https://doi.org/10.1016/j.xpro.2024.102905

SUMMARY

Cancer registry data on pediatric gliomas come with inherent limitations as inclusion criteria and registration practices of these tumors differ between registries due to specific guidelines that are lacking. These limitations can lead to biased estimates in incidence and survival outcomes. Here, we present a protocol to investigate data quality and comparability for retrospective population-based pediatric glioma studies. We describe steps for obtaining institutional permissions, dealing with data quality issues, regrouping tumors, and reporting tumors in a clinically relevant manner.

For complete details on the use and execution of this protocol, please refer to Hoogendijk et al.[1]

BEFORE YOU BEGIN

Over the last decade, molecular diagnostics for central nervous system (CNS) tumors has advanced rapidly.[2,3] For gliomas the first histo-molecular tumor types were introduced in the revised fourth edition of the WHO Classification of Tumors of the CNS (WHO-CNS4R) with diffuse midline glioma, H3 K27M–mutant and RELA fusion–positive ependymoma being of special interest for the pediatric population.[4] With the fifth edition of the World Health Organization Classification of Tumors of the Central Nervous System (WHO-CNS5), the classification system further expanded as pediatric diffuse gliomas were subdivided to pediatric-type diffuse low-grade gliomas and pediatric-type diffuse high-grade gliomas with both groups containing four unique histo-molecular tumor types.[5] Comparable developments occurred for ependymomas where histo-molecular characteristics are combined with the tumor location (i.e., supratentorial, posterior fossa and spinal).

Cancer registries use the International Classification of Diseases for Oncology, 3rd Edition (ICD-O-3), a multi-axial classification system, for coding tumor location (topography) and the histo-molecular diagnosis (morphology) of tumors.[6] The ICD-O-3 provides a standardized terminology for the documentation, reporting, and surveillance of diseases. However, the ICD-O-3 comes with inherent limitations. For example, the ICD-O-3 is lagging behind the WHO classification system making it impossible for registrars to code pediatric brain tumors accurately due to a lack of specific morphology codes. Additionally, the ICD-O-3 aggregates detailed anatomical locations into larger general groups. For example, the ICD-O-3 topography code brain stem (C71.7) includes several anatomic structures such as cerebral peduncle, fourth ventricle, pons and medulla oblongata. These different locations can impact the prognostic profile within the same tumor type. Therefore, analyzing data based on ICD-O-3 without detailed information on morphology and location can lead to clinically invalid outcomes.

In addition to the limitations of current registration practices, data on pediatric brain tumors in cancer registries can be incomplete. For example, registries that are dependent on pathological notification will miss brain tumors where resection is not feasible, especially those tumors located at the optic nerve or the pons. This is especially notable in low resource settings. Moreover, non-pathological-confirmed tumors are not registered in a consistent manner across registries. For example, optic nerve tumors are most commonly diagnosed as a pilocytic astrocytoma (ICD-O Morphology (ICD-O-M) 9421/1) but in our experience are also registered as malignant, glioma Not Otherwise Specified (NOS) (ICD-O-M 9380/3), neurofibromatosis (ICD-O-M9540/1) or neurofibroma (ICD-O-M9540/0). In contrast, the most likely high-grade diagnosis for brain stem tumors in children is a diffuse midline glioma, but these tumors are prone to be (mis)classified by registrars as a low-grade tumor due to the lack of pathological confirmation.

The incompleteness of non-malignant brain tumors for some cancer registries adds another layer of complexity.[7] Non-malignancy is based on the behavior code. The behavior code is included as the fifth digit in the ICD-O morphology code. For CNS tumors only /0 benign, /1 borderline, together classified as non-malignant, and /3 malignant are used. Despite the European Network of Cancer Registries (ENCR) recommending to include all CNS tumors in cancer registries regardless of their behavior since 1998, several authors have reported variation in these registration practices.[8–10]

To tackle some of the aforementioned limitations we recently proposed the implementation of the ICD-11 as a dynamical coding substitute for the ICD-O-3.[11] But limited financial resources and personnel availability of cancer registries hinder rapid implementation.

Therefore, in this STAR protocol we provide a detailed step-by-step guidance on how to recognize data quality issues and how to present outcomes of pediatric gliomas that are clinically relevant based on readily available data in population-based cancer registries.

Institutional permissions

The aforementioned study on sex differences in pediatric high-grade gliomas was conducted in accordance with the Dutch Medical Treatment Contracts Act (WGBO), General Data Protection Regulation (GDPR) and the Code of Conduct for Health Research.[1] We requested a waiver of consent and assent from our institutional review board (IRB) as this study involved collection of retrospective data from existing medical records and the study is assigned no more than a minimal risk. Moreover, the majority of patients who were diagnosed with high-grade brain tumors are no longer alive or are not receiving subsequent therapy at the treating institution.

When analyzing cancer registry data, please review national guidelines if IRB approval or a waiver is needed.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Clinical data	This study[1]	N/A
Software and algorithms		
R (version 4.2.2)	R Foundation	https://www.r-project.org/
Tidyverse	Wickham et al.[12]	https://www.tidyverse.org/
Survminer	N/A	https://github.com/kassambara/survminer
Survival	N/A	https://github.com/therneau/survival
SAS/STAT software	SAS Institute, Inc.	N/A
Stata statistical software	College Station, TX: StataCorp LLC	
Streg	College Station, TX: StataCorp LLC	https://www.stata.com/manuals/ststreg.pdf
Stcox	College Station, TX: StataCorp LLC	https://www.stata.com/manuals/ststcox.pdf
SEER*Stat software	Surveillance Research Program, National Cancer Institute	https://seer.cancer.gov/seerstat/

STEP-BY-STEP METHOD DETAILS

Data quality

Timing: 3–4 days

Before starting the analyses, the data must be checked for completeness and possible misclassification of the tumors included. Although strict criteria would be preferred when analyzing population-based data, in practice this is not feasible due to intrinsic heterogeneity and small sample sizes for some registries. Therefore, these steps need to be treated as a guideline and knowledge about the registry, like notification procedures for newly diagnosed patients, should be incorporated in the decision making.

1. Check the proportion of all CNS tumors that are classified with a non-malignant behavior code.

   Note: For complete registries with a sufficient sample size we expect a non-malignant proportion of about 50%–60% for the age group 0–19 years. Lower proportions suggest incomplete registration of non-malignant CNS tumors or difficult access to pathological reports or medical records.[13,14]

   Note: Dependent of the period of diagnosis pilocytic astrocytomas can be classified as ICD-O-M9421/1 or M9421/3. The latter should be reclassified as ICD-O-M9421/1 to adhere to the most recent registration guidelines and reduce the bias on incidence and survival estimates.[13,15,16] The proportion of non-malignant CNS tumors without pilocytic astrocytomas is expected to be within the range of 40%–50%.

2. Check the proportion of unspecified intracranial and intraspinal neoplasms (ICD-O-M 8000/0, 8000/1, 8000/3, 8002/3) for all CNS tumors.

   Note: For complete registries we expect a proportion of unspecified intracranial and intraspinal neoplasms of 15%. Higher proportions suggest errors in classification or difficult access to pathological reports or medical records.

3. Check the proportion of optic nerve tumors classified as malignant for all CNS tumors.

Note: For complete registries we expect a proportion of 2% of the total population for optic nerve tumors classified as malignant. Higher proportions suggest errors in classification or difficult access to pathological reports or medical records.

4. Check the proportion of tumors classified as malignant, glioma NOS (ICD-O-M9380/3).

   Note: For complete registries we expect a proportion of 5%–10% of the total population for malignant, glioma NOS (ICD-O-M9380/3). Higher proportions suggest errors in classification or difficult access to pathological reports or medical records.

5. Check the proportion of brain stem tumors classified as non-malignant.

   Note: For complete registries we expect a proportion of 3% of the total population for brain stem tumors classified as non-malignant. Higher proportions suggest errors in classification or difficult access to pathological reports or medical records.

6. Check the 5-year survival of gliomas with a dismal prognosis: anaplastic astrocytoma (M-9401/3), anaplastic oligodendroglioma (M9451/3) and glioblastoma (M-9440/3–9442/3). Higher 5-year survival rates in each of these tumor types as mentioned below suggest errors in follow-up.

   Note: Anaplastic astrocytoma, anaplastic oligodendroglioma, and glioblastoma are not recognized as unique tumor types in the WHO-CNS5. Therefore, these tumor types need to been seen in the context of retrospective cancer registry data.

   Note: Expected 5-year survival for anaplastic astrocytoma is 21% (95% CI 16-26).[8]

   Note: Expected 5-year survival for anaplastic oligodendroglioma is 30% (95% CI 20-40).[8]

   Note: Expected 5-year survival for glioblastoma is 14% (95% CI 11-18).[8]

7. Check the 5 year survival of brain stem tumors classified as malignant glioma, NOS (M-9380/3) excl. optic nerve tumors. Higher survival rates than average are suggestive of errors in classification. Higher 5-year survival as mentioned below suggests errors in follow-up or misclassification.

   Note: Expected 5-year survival for malignant glioma, NOS located at the brain stem is 31%.[17]

   ⚠ CRITICAL: All proportions without a reference were based on complete registries that provided data on pediatric brain tumors for the EUROCARE-6 project.[18] As previously mentioned, when checking data quality issues the results need to be balanced to the sample size, specific details of the registry and the pediatric brain tumor population in that country or region. For example, diagnostic options may be limited in low- and middle-income countries inflating the proportion of unspecified CNS tumors. Additionally, novel treatment modalities may increase survival outcomes for lethal pediatric brain tumors notably.[19,20] It is therefore up to the investigators discretion to make a final choice on the data quality. Nevertheless, the authors are of the opinion that completeness of non-malignant CNS tumors remains crucial when analyzing population-based cancer registry data. Lastly, the quality checks should in any case be published alongside the results, so that it becomes more transparent on what kind of information the results are based.

Grouping of pediatric low-grade gliomas

Timing: 1 day

In the WHO-CNS5 pediatric-type diffuse low-grade tumors consist out of diffuse astrocytoma, *MYB*- or *MYBL1*-altered, angiocentric glioma, polymorphous low-grade neuroepithelial tumor of the young and diffuse low-grade glioma, *MAPK* pathway-altered. Additionally, circumscribed astrocytic gliomas like pilocytic astrocytomas are also classified as low-grade tumors. Reporting on a unique tumor type level is preferred. However, coding practices in the past do not allow for up-to-date reporting. Therefore, we propose the following grouping and reporting criteria for standardization purposes and optimal clinical translation.

8. Combine ICD-O-M9421/1, 9421/3, 9425/3, 9410/3, 9411/3, 9420/3, 9400/3, 9384/1, 9424/3, 9442/1, 9450/3, 9382/3, and 9431/1 to form the ***pediatric low-grade gliomas group.*** Troubleshooting.

   Note: Gliofibroma ICD-O-M9442/1 is not a recognized unique tumor type in the more recent WHO classifications. These tumors can be classified as WHO CNS grade 1-4. For these tumors it is important to check the WHO CNS grade to make a final decision on classifying these tumors either in the ***pediatric low-grade gliomas group*** or ***pediatric high-grade gliomas group***.

9. Report outcomes according to the definition scheme below.
   a. Sex, the preferable definition of sex in the context of pediatric brain tumors is the sex assigned at birth. Please refer to the SAGER guidelines for additional information on sex and gender reporting in research.[21]
   b. Age groups that is 0–4, 5-9, 10-14, 15–19 years.
   c. CNS WHO grade 1 and 2; Arabic numerals are preferred above roman numericals to reflect the WHO grade assigned to tumors at initial diagnosis in line with the WHO-CNS5.[5] Troubleshooting.
   d. Glioma, NOS (ICD-O-M9380/3) Troubleshooting.
   e. Topography, that is brain stem (C71.1) and optic nerve (C72.3).

Grouping of pediatric high-grade gliomas

Timing: 1 day

In the WHO-CNS5 pediatric-type diffuse high-grade gliomas consist out of diffuse midline glioma, H3 K27-altered; diffuse hemispheric glioma, H3 G34-mutant; diffuse pediatric-type high-grade glioma, H3-wild type and IDH-wild type; and infant-type hemispheric glioma. Additionally, circumscribed astrocytic gliomas like pleomorphic xanthoastrocytoma grade 3 are also classified as high-grade tumors. Reporting on a unique tumor type level is preferred. However, past coding practices do not allow for up-to-date reporting. Therefore, we propose the following grouping and reporting criteria for standardization purposes and optimal clinical translation.

10. Combine ICD-O-M9401/3, 9381/3, 9442/1, 9440-9442/3, 9451/3, 9382/3, and 9385/3 to form the ***pediatric high-grade gliomas group.*** Troubleshooting.

    Note: Gliofibroma ICD-O-M9442/1 is not a recognized unique tumor type in the more recent WHO classifications. These tumors can be classified as WHO CNS grade 1-4. For these tumors it is important to check the WHO CNS grade to make a final decision on classifying these tumors either in the ***pediatric low-grade gliomas group*** or ***pediatric high-grade gliomas group***.

11. Report outcomes according to the definition scheme below.
    a. Sex, the preferable definition of sex in the context of pediatric brain tumors is the sex assigned at birth. Please refer to the SAGER guidelines for additional information on sex and gender reporting in research.[21]
    b. Age groups that is 0–4, 5-9, 10-14, 15–19 years.

Note: Although infant-type hemispheric glioma is limited to patients <1 year of age.[22] Limiting the age group to 0–1 year can lead to spurious outcomes with wide confidence intervals due to the rarity of these tumors. However, if a large sample size is available it can be considered to further stratify these patients.

c. CNS WHO grade 3 and 4; Arabic numerals are preferred above roman numericals to reflect the WHO grade assigned to tumors at initial diagnosis in line with the WHO-CNS5.[5] Troubleshooting.
d. Glioma, NOS (ICD-O-M9380/3) Troubleshooting.
e. Topography, that is brain stem (C71.1) and optic nerve (C72.3).

Optional: ICD-O-3 topography codes such as cerebrum (71.0) are non-specific as they include tumors in the cerebral cortex but also the thalamus. This limits clinical translation for reported outcomes. If the possibility exists to review the imaging and/or pathology report it would be preferred to regroup tumors to midline and hemispheric tumor locations. Midline tumors are defined as having their primary location in the thalamus, brain stem (e.g. pons, mesencephalon), medulla oblongata, cerebellum and spinal cord. All other tumors can in turn be classified as hemispheric.

Grouping of ependymomas

Timing: 1 day

In the WHO-CNS5 ependymal tumors consist of tumor types combining tumor location and histomolecular characteristics resulting in unique tumor types such as supratentorial ependymoma, *ZFTA* fusion-positive, supratentorial ependymoma, *YAP1* fusion-positive; posterior fossa ependymoma, group PFA; posterior fossa ependymoma, group PFB; and spinal ependymoma, *MYCN*-amplified. Additionally, it contains solely histology-based tumor types like myxopapillary ependymoma and subependymoma. Reporting on a unique tumor type level is preferred. However, past coding practices do not allow for up-to-date reporting. Therefore, we propose the following grouping and reporting criteria for standardization purposes and optimal clinical translation.

12. Group ICD-O-M9383/1, 9391/3, 9392/3, 9393/3, 9394/1, and 9396/3 to form the ***ependymoma group.***
13. Report outcomes according to the definition scheme below.
    a. Sex, the preferable definition of sex in the context of pediatric brain tumors is the sex assigned at birth. Please refer to the SAGER guidelines for additional information on sex and gender reporting in research.[21]
    b. Age groups that is 0–4, 5-9, 10-14, 15–19 years.
    c. CNS WHO grade 1 and a combination of 2 and 3; Arabic numerals are preferred above roman numericals to reflect the WHO grade assigned to tumors at initial diagnosis in line with the WHO-CNS5.[5] Troubleshooting.

Note: WHO grading in ependymomas is surrounded by controversy due to comparable prognostic profiles for WHO CNS grade 2 and 3 across studies.[23] Therefore, we propose to combine and report grade 2, 3 ependymomas as one group.

    d. Topography, that is Supratentorial (C70.0, C71.0, C71.1, C71.2, C71.3, C71.4), Posterior Fossa (C71.5, C71.6, C71.7, C75.3), Spinal (C72.0, C72.1) and Unknown (C71.9, C71.8).

EXPECTED OUTCOMES

The present STAR protocol in population-based cancer registry research on pediatric gliomas provides a structured way of checking the data quality and reporting outcomes in a clinically relevant

manner. Structuring the data in the proposed manner is needed as ICD-O coding as used in cancer registries comes with inherent limitations. We previously reported that the topography code C71.7 brain stem is an umbrella term that contains divers locations such as the pons, medulla oblongata, and the fourth ventricle. When compared with non-brain stem tumors comparable outcomes were found (median survival of 9.7 months for brain stem tumors and 9.8 months for non-brain stem tumors (P = 0.6)).[11] However, when we grouped the same patients into more clinically relevant categories, that is midline and hemispheric tumors, we observed a clinically relevant difference with a median survival for midline tumors at 9 months, significantly differing from hemispheric tumors with a median survival of 14.2 months (P = 0.01).[1] This STAR protocol will facilitate accurate and clinically relevant epidemiological outcomes for pediatric gliomas.

QUANTIFICATION AND STATISTICAL ANALYSIS

The details below provide a description of the most commonly used epidemiological outcomes in the context of pediatric gliomas.

Incidence rates

As cancer in children is rare, it is preferred to calculate incidence rates as the average annual number of cases per million person-years, using the mid-year population size. For the Netherlands these can be obtained from Statistics Netherlands (CBS).[24] Age-adjusted rates for international comparison purposes can be calculated according to the Segi world standard population.[25] It is preferred to calculate age-specific rates for the age groups 0-4, 5-9, 10-14, and 15–19 years. Differences in incidence rates between groups can be expressed as the standardized rate ratio (SRR). SRRs can be calculated according to Rothman et al.[26] or Fay.[27]

Survival rates

As competing causes of death are rare in childhood, Overall Survival (OS) can be calculated instead of Relative Survival (RS).[28] OS is defined as the time from date of diagnosis until death from any cause (i.e., event), date of emigration (i.e., censored) or to study endpoint. Progression Free Survival (PFS) is defined as the time period from date of diagnosis until progression, death from any cause (i.e., event), date of emigration (i.e., censored) or study endpoint. If progression date is not available, an alternative can be using the start date of second line therapy. In cases where no second line therapy was provided, the date of death from any cause can be used. Median, OS and PFS are preferably estimated by Kaplan-Meier method. Log rank testing or cox-proportional hazard models can be used to test for differences in OS and PFS. It is important to note that for most registries reporting of PFS will not be feasible as the date of progression is not a standard item within most population-based cancer registries. Additionally, even when included the determination of PFS comes with limitations as clear clinical and radiographic definitions of progression for patients with a glioma are lacking and treatment related changes may result in differences in interpretation.[29]

Long-term mortality and cure fraction

Follow-up of clinical and population-based cohorts of children diagnosed with CNS tumors have shown a persisting long-term excess mortality with respect to cancer-free children, generally attributable to adverse effects of treatments or the association with genetic diseases or malformations.[18,30,31] The higher mortality can be estimated from the analysis of reliable cause of death information and also, or in absence of it, by statistical models. This information is useful to assess the treatment toxicity, and also allows to correctly estimate the cure fraction, defined as the proportion of patients not bound to die for the progression of the diagnosed cancer.

The 2 main types of cure fraction models are the mixture dependent cure fraction model and the non-mixture cure fraction model. The choice of a model is largely dependent on the assumptions and the underlying study population. The most commonly used model when analyzing cure fraction in population based research is the standard mixture cure fraction model.

The basic formula for a standard mixture cure fraction model is,

$$S(t) = \pi + (1 - \pi)Su(t)$$

where π is the proportion cured and $S_u(t)$ is the survival function for the uncured individuals. $S_u(t)$ is usually a standard parametric survival curve function like the Weibull or lognormal.[32] Both models can be extended to include the non-cancer excess mortality associated with cancer diagnosis and treatment, and the background mortality of cancer-free children. However, a detailed discussion on cure fraction models is beyond the scope of this article. For more details and a recent example on the use of cure fraction models in pediatric oncology we refer to more extensive texts on this subject.[18,32–34]

Prevalence

Prevalence is another outcome related to long term survival. It refers to the number of individuals at a point in time within a population who have been previously diagnosed with the disease and are still alive. Prevalence provides information on the impact of the disease in the population including fatal and cured cancer cases. It is therefore an important indicator of the different needs of the pediatric glioma population. Three main methods are available for assessing prevalence; point prevalence, period prevalence and lifetime prevalence. Differentiation is based on the time frame considered for the estimation. Their respective formulas can be found below.

Formula point prevalence:

$$\frac{Number\ of\ cases\ with\ the\ characteristic\ at\ a\ specific\ point\ in\ time}{Total\ population} \times 100$$

Formula period prevalence:

$$\frac{Number\ of\ cases\ with\ the\ disease\ during\ a\ specific\ time\ period}{Total\ population} \times 100$$

Formula lifetime prevalence:

$$\frac{Number\ of\ cases\ that\ ever\ had\ the\ disease\ at\ any\ point\ in\ life}{Total\ population} \times 100$$

Descriptive, prevalence, incidence and survival analyses can be performed by using R: A language and environment for statistical computing using the tidyverse,[12] survival and survminer packages or the Surveillance Research Program, National Cancer Institute SEER*Stat software. Survival outcomes can also be analyzed via the streg and stcox functions in Stata Statistical Software, College Station, TX: StataCorp LLC. Analyses on incidence rates can be performed with SAS/STAT software.

LIMITATIONS

Although the goal of this protocol is to regroup pediatric gliomas in a clinically relevant manner with data that is readily available in cancer registries we also provide solutions for rising issues that need additional clinical information. Unfortunately, not all cancer registries have direct access to the electronic health records making it hard to accurately group unspecified tumors, like malignant glioma, NOS.

TROUBLESHOOTING

Problem 1

Dependent of the period of diagnosis pilocytic astrocytomas can be classified as ICD-O-M9421/1 or M9421/3 (related to Step 8).

Potential solution

Pilocytic astrocytomas classified as ICD-O-M9421/3 should be reclassified as ICD-O-M9421/1 to adhere to the most recent registration guidelines and reduce the bias on incidence and survival

estimates.[13,15,16] For an extensive discussion on the impact of misclassifying pilocytic astrocytomas see Ostrom et al.[13]

Problem 2

Pilomyxoid astrocytomas (ICD-O-M9425/3) were already debated in the WHO-CNS4R due to their molecular resemblance and comparable prognostic profile with pilocytic astrocytomas, and although these tumors were formally assigned a CNS WHO grade II it was decided in the WHO-CNS4R to suppress the WHO CNS grade until further studies clarify their behavior.[4] More importantly, pilomyxoid astrocytomas (ICD-O-M9425/3) are considered a subtype of pilocytic astrocytomas (related to Step 8).[5]

Potential solution

Although classification and registration practices for these tumors may differ across countries and can possible lead to biased estimates of survival for pilocytic astrocytomas, their impact is expected to be limited due to their rarity. Therefore, we are in favor to report pilomyxoid astrocytomas (ICD-O-M9425/3) as a separate tumor subtype. This can provide insight in their prognostic profile on a population based level. Additionally, a reclassification of pilomyxoid astrocytomas (ICD-O-M9425/3) to pilocytic astrocytoma (ICD-O-M9421/1) can be considered in some cases where the sample size is limited. Although the behavior of pilomyxoid astrocytomas (ICD-O-M9425/3) is still not clarified upon the publication of this protocol we propose to assign CNS WHO grade I to these tumors in further analyses in line with the CNS WHO grade of pilocytic astrocytomas.

Problem 3

Most glial CNS tumor type have a unique CNS WHO grade that can be extracted from the WHO-CNS5 based on the ICD-O morphology code.[5] For example, subependymomas (ICD-O-M9383/1) and subependymal giant cell astrocytoma (ICD-O-M9384/1) are per definition classified as CNS WHO grade 1. However, for some tumor types, i.e., oligoastrocytic tumors (ICD-O-M9382/3) or oligodendrogliomas (ICD-O-M9451/3) the appropriate CNS WHO grade (2 versus 3) is not clear due to a lack of sufficient detail in the ICD-O morphology code (related to Steps 11 and 13).

Potential solution

Assigning the CNS WHO grade to oligoastrocytic tumors (ICD-O-M9382/3) is challenging in the current molecular era as these tumors are nowadays usually classified as either an oligodendroglioma or an astrocytoma (e.g., astrocytoma IDH-mutant). However, it is important to note that these two tumor entities hardly occur in children.[35] Additionally, analyzing oligodendrogliomas (ICD-O-M9451/3) in children on a molecular level can change the diagnosis to a supratentorial ependymoma or a glioneuronal tumor such as a diffuse glioneuronal tumor with oligodendroglioma-like features and nuclear clusters (DGONC).[35] However, in most cases it will not feasible to reclassify these tumors molecularly to the most recent classification. Therefore, in line with ICD-O coding oligoastrocytic tumors and oligodendrogliomas should be assigned a CNS WHO grade 2 or 3. If the CNS WHO grade is available this should be the primary choice for assigning the CNS WHO grade. If the CNS WHO grade is not provided by the cancer registry the preferred CNS WHO grade should be based on the behavior code, thus CNS WHO grade 3. These choices should be specified in the method section of the manuscript.

Problem 4

ICD-O coding mistakes together with ambiguous coding practices are not common but unavoidable when analyzing pediatric glial tumors (related to steps 8 and 10).

Potential solution

Below we address some ICD-O coding issues we encountered when analyzing European data on glial tumors and provide potential solutions. If the CNS WHO grade is available this should be the primary choice for assigning the CNS WHO grade.

ICD-O-M9400/1
This seems to be a misclassification of the behavior code and the most likely diagnosis is a diffuse low-grade astrocytoma (ICD-O-M9400/3). The CNS WHO grade should therefore be classified as grade 2.

ICD-O-M9401/1
This seems to be a misclassification of the behavior code and the most likely diagnosis is an anaplastic astrocytoma (ICD-O-M9401/3). The CNS WHO grade should therefore be classified as grade 3.

ICD O-M9440/1
This seems to be a misclassification of the behavior code and the most likely diagnosis is a glioblastoma (ICD-O-M9440/3). The CNS WHO grade should therefore be classified as grade 4.

The underlying reasoning should be specified in the method section of the manuscript.

Problem 5
Unspecified tumors like malignant gliomas, NOS (9380/3) are a challenge for accurate reporting of cancer registration data on pediatric gliomas. Misclassification of unspecified tumors can lead to biased estimates of incidence and survival (related to steps 8 and 10).

Potential solution
The ENCR recently released a revised version of their recommendations on basis of diagnosis. These recommendations provide guidance on how to register the underlying diagnostic modalities. It also offers direction on diagnosing CNS tumors based solely on clinical investigations.[16] For instance, malignant glioma, NOS (ICD-O-M9380/3), can now be more precisely classified by adding a sixth digit onto the morphology code. Resulting in malignant glioma, NOS (ICD-O-M9380/39), low-grade glioma (ICD-O-M9380/32), and high-grade glioma (ICD-O-M9380/33). If this information is not available, data for unspecified tumors like malignant gliomas, NOS should be checked for their diagnoses based on imaging. If there is a high probability of a low-grade tumor (WHO grade 1-2) these tumors should be included in the low-grade astrocytomas and gliomas group. However, radiologically seemingly low grade, diffuse infiltrative tumors located in the pons should be included in the high-grade astrocytomas and gliomas group, as the vast majority of these are nowadays known to represent the highly malignant diffuse midline gliomas, H3 K27-altered.

RESOURCE AVAILABILITY

Lead contact
Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Raoull Hoogendijk (r.hoogendijk@prinsesmaximacentrum.nl).

Technical contact
Questions about the technical specifics of performing the protocol should be directed to and will be answered by the technical contact, Raoull Hoogendijk (r.hoogendijk@prinsesmaximacentrum.nl).

Materials availability
This study did not generate new unique materials.

Data and code availability
Data reported in this paper will be shared by the lead contact upon reasonable request. This paper does not report original code. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

ACKNOWLEDGMENTS

The authors would like to thank the registration team of the Netherlands Comprehensive Cancer Organization (IKNL) for the collection of data for the NCR and this project.

AUTHOR CONTRIBUTIONS

Conceptualization, R.H., H.K.-K., J.v.d.L., and D.G.v.V.; methodology, R.H., H.K.-K., J.v.d.L., and D.G.v.V.; writing – original draft, R.H.; writing – review and editing, all authors; supervision, H.K.-K., J.v.d.L., D.G.v.V., P.W., and E.W.H.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

1. Hoogendijk, R., van der Lugt, J., Baugh, J., Kline, C., Kranendonk, M., Hoving, E., Kremer, L., Wesseling, P., Karim-Kos, H., and van Vuurden, D. (2023). Sex-related incidence and survival differences in pediatric high-grade glioma subtypes: a population-based cohort study. iScience *26*, 107957. https://doi.org/10.1016/j.isci.2023.107957.

2. Schwartzentruber, J., Korshunov, A., Liu, X.Y., Jones, D.T.W., Pfaff, E., Jacob, K., Sturm, D., Fontebasso, A.M., Quang, D.A.K., Tönjes, M., et al. (2012). Driver mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma. Nature *482*, 226–231. https://doi.org/10.1038/nature10833.

3. Capper, D., Jones, D.T.W., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., Reuss, D.E., et al. (2018). DNA methylation-based classification of central nervous system tumours. Nature *555*, 469–474. https://doi.org/10.1038/nature26000.

4. Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., and Ellison, D.W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. Acta Neuropathol. *131*, 803–820. https://doi.org/10.1007/s00401-016-1545-1.

5. Louis, D.N., Perry, A., Wesseling, P., Brat, D.J., Cree, I.A., Figarella-Branger, D., Hawkins, C., Ng, H.K., Pfister, S.M., Reifenberger, G., et al. (2021). The 2021 WHO Classification of Tumors of the Central Nervous System: a summary. Neuro Oncol. *23*, 1231–1251. https://doi.org/10.1093/neuonc/noab106.

6. World Health Organization (2013). International Classification of Diseases for Oncology (ICD-O), 3rd edition. 1st revision, 3rd ed.. Edition.

7. Hoogendijk, R., van der Lugt, J., van Vuurden, D.G., Visser, O., and Karim-Kos, H.E. (2022). Reporting pediatric brain tumors according to their behavior code can result in biased survival estimates-A European perspective to Girardi et al. Neuro Oncol. *24*, 1205–1207. https://doi.org/10.1093/neuonc/noac039.

8. Gatta, G., Peris-Bonet, R., Visser, O., Stiller, C., Marcos-Gragera, R., Sánchez, M.J., Lacour, B., Kaatsch, P., Berrino, F., Rutkowski, S., et al. (2017). Geographical variability in survival of European children with central nervous system tumours. Eur. J. Cancer *82*, 137–148. https://doi.org/10.1016/j.ejca.2017.05.028.

9. Girardi, F., Rous, B., Stiller, C.A., Gatta, G., Fersht, N., Storm, H.H., Rodrigues, J.R., Herrmann, C., Marcos-Gragera, R., Peris-Bonet, R., et al. (2021). The histology of brain tumors for 67 331 children and 671 085 adults diagnosed in 60 countries during 2000-2014: a global, population-based study (CONCORD-3). Neuro Oncol. *23*, 1765–1776. https://doi.org/10.1093/neuonc/noab067.

10. Recommendations for coding tumours of the brain and central nervous system. https://encr.eu/sites/default/files/pdf/braincns.pdf.

11. Hoogendijk, R., van der Lugt, J., Hoving, E., Kremer, L., Visser, O., Wesseling, P., van Vuurden, D., and Karim-Kos, H. (2022). The 5th edition of the World Health Organization Classification of Tumors of the Central Nervous System: Implications for cancer registries. Neuro Oncol. *24*, 1811–1814. https://doi.org/10.1093/neuonc/noac176.

12. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686. https://doi.org/10.21105/joss.01686.

13. Ostrom, Q.T., Kruchko, C., and Barnholtz-Sloan, J.S. (2020). Pilocytic astrocytomas: where do they belong in cancer reporting? Neuro Oncol. *22*, 298–300. https://doi.org/10.1093/neuonc/noz202.

14. Ostrom, Q.T., Cioffi, G., Waite, K., Kruchko, C., and Barnholtz-Sloan, J.S. (2021). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2014-2018. Neuro Oncol. *23*, iii1–iii105. https://doi.org/10.1093/neuonc/noab200.

15. National Cancer Institute. https://seer.cancer.gov/tools/solidtumor/2023/Non_malignant_CNS_STM.pdf.

16. Visser, O., Caetano dos Santos, F.L., Cuccaro, F., Forjaz, G., Irmina Michalek, M.M., Sulkowska, U., Martos, C., and Giusti, F. (2022). ENCR Recommendation - Basis of Diagnosis.

17. Hoogendijk, R., van der Lugt, J., van Vuurden, D., Kremer, L., Wesseling, P., Hoving, E., and Karim-Kos, H.E. (2022). Survival rates of children and young adolescents with CNS tumors improved in the Netherlands since 1990: A population-based study. Neurooncol. Adv. *4*, vdab183. https://doi.org/10.1093/noajnl/vdab183.

18. Botta, L., Gatta, G., Capocaccia, R., Stiller, C., Canete, A., Dal Maso, L., Innos, K., Mihor, A., Erdmann, F., Spix, C., et al. (2022). Long-term survival and cure fraction estimates for childhood cancer in Europe (EUROCARE-6): results from a population-based study. Lancet Oncol. *23*, 1525–1536. https://doi.org/10.1016/S1470-2045(22)00637-4.

19. Gallego Perez-Larraya, J., Garcia-Moure, M., Labiano, S., Patino-Garcia, A., Dobbs, J., Gonzalez-Huarriz, M., Zalacain, M., Marrodan, L., Martinez-Velez, N., Puigdelloses, M., et al. (2022). Oncolytic DNX-2401 Virus for Pediatric Diffuse Intrinsic Pontine Glioma. N. Engl. J. Med. *386*, 2471–2481. https://doi.org/10.1056/NEJMoa2202028.

20. Kline, C., Franson, A., Banerjee, A., Reddy, A.T., Raber, S., Hoffman, C., Minturn, J., van der Lugt, J., de Vos-Kerkhof, E., van Vuurden, D., et al. (2023). CTNI-84. PNOC022: A COMBINATION THERAPY TRIAL USING AN ADAPTIVE PLATFORM DESIGN FOR PATIENTS WITH DIFFUSE MIDLINE GLIOMAS (DMGS) AT INITIAL DIAGNOSIS, POST-RADIATION THERAPY AND AT TIME OF PROGRESSION. Neuro Oncol. *25*, v97. https://doi.org/10.1093/neuonc/noad179.0365.

21. Heidari, S., Babor, T.F., De Castro, P., Tort, S., and Curno, M. (2016). Sex and Gender Equity in Research: rationale for the SAGER guidelines and recommended use. Res. Integr. Peer Rev. *1*, 2. https://doi.org/10.1186/s41073-016-0007-6.

22. Chiang, J., Bagchi, A., Li, X., Dhanda, S.K., Huang, J., Pinto, S.N., Sioson, E., Dalton, J., Tatevossian, R.G., Jia, S., et al. (2024). High-grade glioma in infants and young children is histologically, molecularly, and clinically diverse-Results from the SJYC07 trial and institutional experience. Neuro Oncol. *26*, 178–190. https://doi.org/10.1093/neuonc/noad130.

23. Ellison, D.W., Kocak, M., Figarella-Branger, D., Felice, G., Catherine, G., Pietsch, T., Frappaz, D., Massimino, M., Grill, J., Boyett, J.M., and Grundy, R.G. (2011). Histopathological grading of pediatric ependymoma: reproducibility and clinical relevance in European trial cohorts. J. Negat. Results Biomed. *10*, 7. https://doi.org/10.1186/1477-5751-10-7.

24. Statline. https://opendata.cbs.nl/#/CBS/nl/.

25. Ahmad, O.B., B, P.C., Lopez, A.D., Murray, C.J.L., Lozano, R., and M, I. (2001). Age Standardization of Rates: A New WHO Standard.

26. Rothman KJ, G.S. (2008). Modern Epidemiology, 3rd Edition (Lippincott Williams & Wilkin).

27. Fay, M.P. (1999). Approximate confidence intervals for rate ratios from directly standardized rates with sparse data. Commun. Stat. Theor. Methods *28*, 2141–2160. https://doi.org/10.1080/03610929908832411.

28. Sankila, R., Martos Jiménez, M.C., Miljus, D., Pritchard-Jones, K., Steliarova-Foucher, E., and Stiller, C. (2006). Geographical comparison of cancer survival in European children (1988-1997): report from the Automated Childhood Cancer Information System project. Eur. J. Cancer *42*, 1972–1980. https://doi.org/10.1016/j.ejca.2006.05.013.

29. Cooney, T., Lane, A., Bartels, U., Bouffet, E., Goldman, S., Leary, S.E.S., Foreman, N.K., Packer, R.J., Broniscer, A., Minturn, J.E., et al. (2017). Contemporary survival endpoints: an International Diffuse Intrinsic Pontine Glioma Registry study. Neuro Oncol. *19*, 1279–1280. https://doi.org/10.1093/neuonc/nox107.

30. Byrne, J., Schmidtmann, I., Rashid, H., Hagberg, O., Bagnasco, F., Bardi, E., De Vathaire, F., Essiaf, S., Winther, J.F., Frey, E., et al. (2022). Impact of era of diagnosis on cause-specific late mortality among 77 423 five-year European survivors of childhood and adolescent cancer: The PanCareSurFup consortium. Int. J. Cancer *150*, 406–419. https://doi.org/10.1002/ijc.33817.

31. Mertens, A.C., Liu, Q., Neglia, J.P., Wasilewski, K., Leisenring, W., Armstrong, G.T., Robison, L.L., and Yasui, Y. (2008). Cause-specific late mortality among 5-year survivors of childhood cancer: the Childhood Cancer Survivor Study. J. Natl. Cancer Inst. *100*, 1368–1379. https://doi.org/10.1093/jnci/djn310.

32. Lambert, P.C., Thompson, J.R., Weston, C.L., and Dickman, P.W. (2007). Estimating and modeling the cure fraction in population-based cancer survival analysis. Biostatistics *8*, 576–594. https://doi.org/10.1093/biostatistics/kxl030.

33. Sposto, R. (2002). Cure model analysis in cancer: an application to data from the Children's Cancer Group. Stat. Med. *21*, 293–312. https://doi.org/10.1002/sim.987.

34. Felizzi, F., Paracha, N., Pöhlmann, J., and Ray, J. (2021). Mixture Cure Models in Oncology: A Tutorial and Practical Guidance. Pharmacoecon. Open *5*, 143–155. https://doi.org/10.1007/s41669-021-00260-z.

35. Wesseling, P., and Capper, D. (2018). WHO 2016 Classification of gliomas. Neuropathol. Appl. Neurobiol. *44*, 139–150. https://doi.org/10.1111/nan.12432.