Cancer Epidemiology 78 (2022) 102145

Contents lists available at ScienceDirect

Cancer Epidemiology

journal homepage: www.elsevier.com/locate/canep

Neuroblastoma in Spain: Linking the national clinical database and epidemiological registries – A study by the Joint Action on Rare Cancers☆

Adela Cañete [a,b,*], Rafael Peris-Bonet [a], Riccardo Capocaccia [c], Elena Pardo-Romaguera [a], Vanessa Segura [d], Ana Muñoz-López [a], Ana Fernández-Teijeiro [e,f], Jaume Galceran-Padros [g], Gemma Gatta [h], Spanish Neuroblastoma Linkage Working Group (SpNbLinkWG)[1]

[a] Spanish Registry of Childhood Tumours (RETI-SEHOP), University of Valencia, Valencia, Spain
[b] Paediatric Oncohematology Unit, Hospital La Fe; and Department of Paediatrics, University of Valencia, Valencia, Spain
[c] Direzione scientifica, Epidemiologia & Prevenzione, Milano, Italy
[d] Clinical Translational Cancer Research Group – Paediatric Oncology, IIS La Fe, Valencia, Spain
[e] Spanish Society of Paediatric Haematology and Oncology (SEHOP), Spain
[f] Paediatric Oncohematology Unit, Hospital Universitario Virgen Macarena, University of Sevilla, Sevilla, Spain
[g] Spanish Network of Cancer Registries (REDECAN), Spain
[h] Department of Epidemiological Research and Molecular Medicine, Evaluative Epidemiology Unit, Fondazione IRCCS Istituto Nazionale dei Tumori di Milano, Italy

ARTICLE INFO

Keywords:
Neuroblastoma
Childhood cancer
Cancer registry
Clinical registry
Record linkage
Completeness
Incidence
Survival
European Reference Network
Joint Action on Rare Cancers

ABSTRACT

Purpose: Linkage between clinical databases and population-based cancer registries may serve to evaluate European Reference Networks' (ERNs) activity, by monitoring the proportion of patients benefiting from these and their impact on survival at a population level. To test this, a study targeting neuroblastoma (Nb) was conducted in Spain by the European Joint Action on Rare Cancers.

Material and methods: Subjects: Nb cases, incident 1999–2017, aged < 15 years. Linkage included: Spanish Neuroblastoma Clinical Database (NbCDB) (1217 cases); Spanish Registry of Childhood Tumours (RETI) (1514 cases); and 10 regional population-based registries (RPBCRs) which cover 33% of the childhood population (332 cases). Linkage was semiautomatic. We estimated completeness, incidence, contribution, deficit, and 5-year survival in the databases and specific subsets.

Results: National completeness estimates for RETI and NbCDB were 91% and 72% respectively, using the Spanish RPBCRs on International Incidence of Childhood Cancer (https://iicc.iarc.fr/) as reference. RPBCRs' specific contribution was 1.6%. Linkage required manual crossover in 54% of the semiautomatic matches. Five-year survival was 74% (0–14 years) and 90% (0–18 months).

Conclusions: All three databases were incomplete as regards Spain as a whole and should therefore be combined to achieve full childhood cancer registration. A unique personal patient identifier could facilitate such linkage. Most children have access to Nb clinical trials. Consolidated interconnections between the national registry and clinical registries (including ERNs and paediatric oncology clinical groups) should be established to evaluate outcomes.

1. Introduction

Childhood cancer incidence in Europe is 141 cases/1,000,000 children aged 0–14 years [1]. According to the published threshold for rare cancers [2], all childhood cancers are life-threatening rare diseases, facing huge challenges (lack of correct diagnosis and treatment, and hence worse survival). Rare cancer survival rates vary across Europe[3,4], reflecting these deficiencies in detection, diagnosis and treatment, and disparities among regions/countries [5]. To overcome this, joint expertise and networked healthcare are required. European Reference Networks (ERNs) were born as *virtual networks involving healthcare*

☆ and the SpNbLinkWG.

* Correspondence to: POG Department, University of Valencia, Paediatric Oncohematology Unit, Torre G-21, Hospital La Fe, Avda Fernando Abril Martorell 106, Valencia 46026, Spain.
E-mail address: adela.canete@uv.es (A. Cañete).

[1] Paediatric oncology units and cancer registries that contributed data to the linkage are included

https://doi.org/10.1016/j.canep.2022.102145
Received 13 October 2021; Received in revised form 8 February 2022; Accepted 11 March 2022
Available online 25 March 2022

Nomenclature

ERNs	European Reference Networks
CDB	Clinical database
JARC	Joint Action on Rare Cancers
NbCDB	Neuroblastoma clinical database
RETI	Spanish Registry of Childhood Tumours
RPBCRs	Regional population-based cancer registries.
SEHOP	Spanish Society of Paediatric Haematology and Oncology
CC	Childhood cancers
ASRw	Age-standardised incidence rate
ICCC-3	International Classification of Childhood Cancer
INES	Infant Neuroblastoma European Study.
EUNS	European Unresectable Neuroblastoma Study.
SIOPEN	Neuroblastoma Group of the European Society of Paediatric Oncology

providers across Europe, aimed at tackling complex or rare diseases and conditions that require highly specialised treatment and concentration of knowledge and resources [6]. PAEDCAN is the ERN for paediatric cancer [7].

The ERNs' capacity to work with patients and the final impact on the real population must be monitored by current cancer information systems. Clinical databases (CDBs) and population-based cancer registries (PBCRs) can be linked to monitor the proportion of rare cancer patients benefiting from the ERNs, and thereby contribute to evaluate ERNs at a population level.

To this end, we conducted a linkage study in Spain, within the European Joint Action on Rare Cancers (JARC) [8], targeting neuroblastoma (Nb), the most frequent paediatric extracranial solid tumour (incidence in Spain: 14.9 cases/1,000,000 children aged 0–14 years [9]). Nb has a heterogeneous clinical presentation and outcome. Trials have long shown that infants (children younger than 366 days) have better outcomes. London et al. [10] showed that age is a continuous variable for risk stratification, and the International Neuroblastoma Risk Group [11] confirmed this, setting the cut-off point at 18 months (547 days).

We linked the Spanish Neuroblastoma Clinical Database (NbCDB), the Spanish Registry of Childhood Tumours (RETI) (both national) and regional population-based cancer registries (RPBCRs), and report our results in terms of linkage, feasibility, completeness and survival.

2. Material and methods

We studied children aged 0–14 years diagnosed with Nb (Group IVa, International Classification of Childhood Cancer [12]) nationwide across the period 1999–2017. The average Spanish national population in this age group during the study period totalled 6,585,544 (Spanish National Statistics Institute [13]). 3063 records from three independent sources were included in the study.

2.1. Sources

2.1.1. Spanish Neuroblastoma Clinical Database (NbCDB)

The NbCDB [14] includes all cases in clinical trials/studies. The Spanish Neuroblastoma Group was created in 1987 by the Spanish Society of Paediatric Oncology (*SEHOP*) to help clinicians with diagnosis, treatment, and trial enrolment. A procedure is available which facilitates biological, histopathological studies and trial assignment for centres that participate voluntarily. All patients (or guardians) are required to give informed consent. Cases are confirmed centrally and followed up according to clinical guidelines. Valencia's La Fe Hospital database stores data on 1400 patients from 38 hospitals, up until 2017. Hospital coverage remained stable across the study period. Second opinions and non-incident cases were excluded from the study. Cases contributed: 1217.

2.1.2. Spanish Registry of Childhood Tumours (RETI)

The RETI [15] coordinates a nationwide network including all the 48 paediatric oncology units distributed across all regions of Spain. After obtaining informed consent, these units report all new cases and follow them up on request. The main goal is survival estimation as an indicator of results. RETI's estimated average completeness for all of Spain is currently 95%, and approaches to be virtually complete for children aged 0–14 years in five regions (Aragon, Basque Country, Catalonia, Madrid, Navarre) [15,16], covering 37% of Spain [13] (Supplementary Table 1). These regions constitute the RETI's "high-completeness area" included in the European Network of Cancer Registries[17] and population-based international projects [1,4,9,18]. Cases contributed: 1514.

2.1.3. Regional population-based cancer registries [19]

Sixteen RPBCRs were active during the study period, covering 41% of children: ten participated in this study (Albacete, Asturias, Canary Islands, Castile-Leon, Ciudad Real, Girona, Majorca, Murcia, Tarragona and Valencian Region), covering 33% of Spanish children [13] (Supplementary Table 1). Five registries could not participate. All regional registries are members of the European Network [17] and the Spanish Network of Cancer Registries [19]. They have participated in international projects which require and evaluate population-base completeness [1,4,9,20]. Cases contributed: 332.

All participating registries and databases had regulatory data-exchange agreements with RETI. See variables in Annex-1.

2.2. Incidence

Incidence rates were calculated as number of cases per million child-years, and age-standardised rates adjusted to the world standard population (ASRw) [21]. 95% confidence intervals (95%CI) were estimated by Poisson approximation. The population at risk in the participating RPBCRs was calculated according to their geographical areas and participation periods (Supplementary Table 2). In the RETI, NbCDB, RETI&NbCDB and (RETI&NbCDB)&RPBCRs sets (see linkage below), the population at risk was the entire Spanish population aged 0–14 years in 1999–2017 [13].

2.3. Completeness

Completeness of registration by NbCDB, RETI, RETI&NbCDB and (RETI&NbCDB)&RPBCRs was estimated by the ratio of observed-to-expected cases [22]. 95%CI was estimated using Vandenbroucke's method [23]. Expected cases were estimated from the 12 Spanish RPBCRs included in the International Incidence of Childhood Cancer (IICC-3) [9], excluding RETI (hereinafter RPBCRs-IICC-3) and the corresponding population for the whole of Spain in 1999–2017 [13]. The IICC-3 was chosen as the most recent world account of childhood cancer incidence by country and registry. The RPBCRs-IICC-3 covered the period 1990–2013 and 35% of Spanish children (Supplementary Table 1).

Since the incidence rate for ages 0–18 months was not available in the RPBCRs-IICC-3, and the population for this age group was not in the official population figures[13], the incidence rate for this age group was estimated using data from the RETI's high-completeness area in 2000–2016. Cases were those aged 0–18 months, and the population at risk was approximated by adding half of all children aged 1 year to the 0-year population [13]. Estimated rate 0–18 months = 66.4 cases/1,000,000 (95%CI: 58.9–73.8).

Since RPBCRs are population-based registries previously evaluated

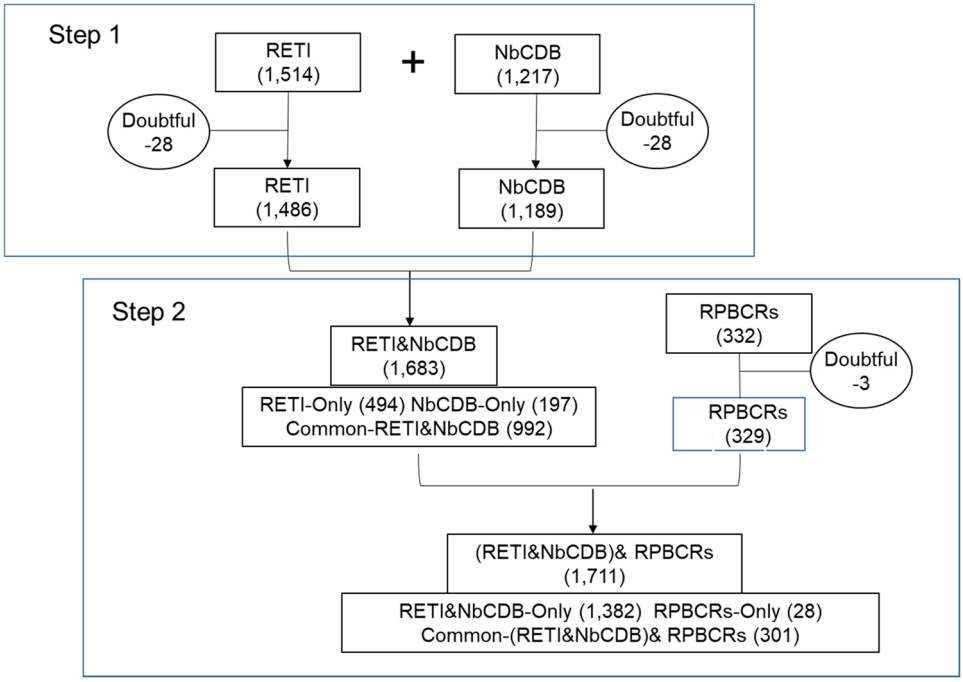

Fig. 1. Simplified diagram of the linkage.

in international studies, their completeness was assumed to be virtually complete by current standards.

2.4. Linkage

All cases were included in linkage. Matching was semi-automatic using Dalink v4.2 software [24]. Since there was no unique personal identifier common to all files, personal details and shared variables were used for linkage. Weights were assigned to each item as follows: 1st surname (25%); 2nd surname (25%); name (24%); and birth and incidence dates (13% each). A similarity probability was obtained for each matching. A human decision was needed for probabilities < 60% (due to name variability, abbreviations, spelling, etc.). Complementary information was consulted, when necessary. Nevertheless, a matching decision was still impossible for some record pairs, which coincided in some details while differing in others, e.g., some records had the same first and last names, but different birth or incidence dates. Where complementary information did not permit a decision, records were classified as 'doubtful matchings' and the patients excluded from all analyses.

The linkage procedure was as follows (see also Fig. 1):

1st Step. Linkage between RETI and NbCDB (all Spain, 0–14 years, 1999–2017).

The result was the RETI&NbCDB file, with cases present only in RETI and not in NbCDB (RETI-only), only in NbCDB and not in RETI (NbCDB-only), or in both (common-RETI&NbCDB). RETI&NbCDB was the final set.

2nd Step. Linkage between RETI&NbCDB and RPBCRs.

The result was the (RETI&NbCDB)&RPBCRs file, with RETI&NbCDB-only cases, RPBCR-only cases and common-(RETI&NbCDB)&RPBCRs cases. (RETI&NbCDB)&RPBCRs was the final set.

2.5. Contribution and deficit

Contribution and deficit were evaluated using five indicators [25]. For simplicity, only two files (A and B) are considered here:

2.5.1. Total contribution

Cases in A.

$$\text{Co\% A} = \frac{\text{Cases in final set without repetitions}}{\text{Cases in final set}} \times 100.$$

2.5.2. Specific contribution

Cases in A, but not B.

$$\text{SCo\% A} = \frac{\text{Cases in final set without repetitions}}{\text{Cases in final set}} \times 100.$$

2.5.3. Underreporting

URT% A = 100 – Co% file A.

2.5.4. Overlapping between two files

Cases common to A and B, without repetitions.

$$\text{Ov\% A (in B)} = \frac{\text{Cases in final set without repetitions}}{\text{Cases in A}} \times 100.$$

2.5.5. Overlapping among all files

Cases common to all files, without repetitions.

$$\text{OvT\%} = \frac{\text{Cases in final set without repetitions}}{\text{Cases in final set}} \times 100.$$

2.6. Survival

Observed 5-year survival was estimated using the Kaplan–Meier method[26]. To allow for complete 5-year follow-up, estimates were made for the 1999–2011, 1999–2005, and 2006–2011 periods. Cases were followed up until the 5th anniversary of the incidence date, using the three sources and the National Death Index where necessary. Without a common identifier, searches in the Death Index had to be performed using personal identification details. Patients lost to

follow-up were those who were alive at the last documented date of follow-up before the 5th anniversary. No estimates were made in < 15 cases. Survival differences were tested using the Log-Rank[27] for ages 0–14 years and 0–18 months, period 1999–2011. All calculations were performed using the IBM SPSS Statistics 24.0 programme.

3. Results

The three sources supplied 3063 records in total. Linkage required manual crossover in 54% of matches. Fifty-nine doubtful matching patients were excluded (RETI, 28; NbCDB, 28; RPBCRs, 3), leaving 3004 patients (RETI, 1486; NbCDB, 1189; RPBCRs, 329) for analysis purposes.

Table 1
Linkage of Nb in Spain study. Number of cases in each file entering the linkage (RETI, NbCDB, RPBCRs) and linked sets (RETI&NbCDB, (RETI&NbCDB)&RPBCs), 1999–2017, and the reference set of registries (RPBCRs-IICC-3), 1990–2013, by age group and incidence rate.

Files/sets	N			Incidence rates per 1,000,000 person-years			
	0–18 months	0–4 years	0–14 years	0–18 months	0–4 years	0–14 years	0–14 years ASRw (95% CI)
RETI	803	1304	1486	63.5	31.0	11.9	13.4 (12.7–14.1)
NbCDB	654	1047	1189	51.7	24.9	9.5	10.7 (10.1–11.3)
RPBCRs	186	295	329	N/A	30.8	11.6	13.1 (11.7–14.5)
RETI&NbCDB	915	1478	1683	72.4	35.1	13.5	15.1 (14.4–15.9)
(RETI&NbCDB) &RPBCRs	931	1502	1711	73.7	35.7	13.7	15.4 (14.7–16.1)
RPBCRs-IICC-3	N/A	500	561	N/A	35.0	12.3	14.9 (13.8–16.0)

NOTES
-Nb: Neuroblastoma
-RETI: File of the Spanish Registry of Childhood Tumours
-NbCDB: File of the Spanish Neuroblastoma Clinical Database
-RPBCRs: File of the regional population-based cancer registries participating in the study (Supplementary Table 2). The difference between the total number of cases in this row and the total in Supplementary Table 2 is due to the exclusion of three patients with doubtful matching, as stated in the Results section
-RETI&NbCDB: Set resulting from linkage of the RETI and NbCDB files
-(RETI&NbCDB)&RPBCRs: Set resulting from linkage of the RETI&NbCDB and RPBCR files
-RPBCRs-IICC-3: Set of Spanish RPBCRs in the IICC-3[9], with the RETI excluded, used as a reference. Period: 1990–2013 (see Supplementary Table 1)
-ASRw: Age standardised rate based on the world standard population[21]
-N/A: Not available
-Populations at risk:
-RETI, NbCDB, RETI&NbCDB and (RET&NbCDB)&RPBCRs in all Spain (1999–2017): 0–14 years of age, 125,125,334 person-years; 0–18 months, 12,640,259 person-years
-RPBCRs, in the geographical area covered by the registries and their period of participation (see Supplementary Table 2) (0–14 years of age): 28,407,194 person-years; 0–18 months, not available
-RPBCRs-IICC-3 (1990–2013): 0–14 years, 45,529,309; 0–18 months, not available

Table 2
Linkage of Nb in Spain study. Number of cases and completeness of the files entering the linkage (RETI, NbCDB, RPBCRs) and linked sets (RETI&NbCDB, RETI&NbCDB)&RPBCs), 1999–2017, and the reference set of registries (RPBCRs-IICC-3), 1990–2013, by age group.

Sources/sets	N 0–14 years	Completeness for Spain % (CI 95%)		
		0–18 months[a]	0–4 years[b]	0–14 years[b]
RETI	1486	95.7 (89.2–102.4)	88.5 (83.8–93.4)	90.6 (86.1–95.3)
NbCDB	1189	77.9 (72.1–84.0)	71.1 (66.8–75.4)	72.5 (68.4–76.7)
RPBCRs	329	c		
RETI&NbCDB	1683	109.0 (102.1–116.2)	100.3 (95.3–105.5)	102.6 (97.8–107.6)
(RETI&NbCDB) &RPBCRs	1711	110.9 (103.9–118.2)	101.9 (96.9–107.2)	104.3 (99.5–109.4)
RPBCRs-IICC-3	561	c		

NOTES
-Nb: Neuroblastoma
-RETI: File of the Spanish Registry of Childhood Tumours
-NbCDB: File of the Spanish Neuroblastoma Clinical Database
-RPBCRs: File of the regional population-based cancer registries participating in the study (Supplementary Table 2). The difference between the total number of cases in this row and the total in Supplementary Table 2 is due to the exclusion of three patients with doubtful matching, as stated in the Results section
-RETI&NbCDB: Set resulting from linkage of the RETI and NbCDB files
-(RETI&NbCDB)&RPBCRs: Set resulting from linkage of the RETI&NbCDB and RPBCRs files
-RPBCRs-IICC-3: Set of Spanish RPBCRs in the IICC-3[9], with the RETI excluded, used as a reference. Period: 1990–2013 (see Supplementary Table 1)
-Completeness:
a: The expected number of Nb cases to calculate completeness for children 0–18 months of age is based on the rate estimated from the population-based area of RETI for 0–18 months.
b: The expected number of Nb cases to calculate completeness for children aged 0–4 and 0–14 years is based on the rates from the RPBCR-IICC-3 set for the same groups of age.
c: Completeness of RPBCRs assumed to be virtually complete in their geographical area of coverage.
-Population at risk
-RETI, NbCDB, RETI&NbCDB and (RET&NbCDB)&RPBCRs in all Spain (1999–2017): 0–14 years of age, 125,125,334 person-years; 0–18 months, 12,640,259 person-years
-RPBCRs, in the geographical area covered by the registries and their period of participation (see Supplementary Table 2) (0–14 years of age): 28,407,194 person-years; 0–18 months, not available
-RPBCRs-IICC-3 (1990–2013): 0–14 years, 45,529,309; 0–18 months, not available

The RETI incidence rates in Spain were similar to those of the RPBCRs set, while the NbCDB rates were slightly lower (Table 1). Taking the RPBCRs-IICC-3 as reference, both RETI and NbCDB failed to achieve total completeness (Table 2).

The linked RETI&NbCDB file had 1683 individual patients (992 common, 494 RETI-only and 197 NbCDB-only; Supplementary Table 3). Accordingly, the RETI's and NbCDB's underreporting rates were 11.7% and 29.4% respectively (Table 3). In the second linkage step, 28 additional cases were identified in the RPBCRs (Supplementary Table 3). Since participating RPBCRs covered one-third of the Spanish population, about 100 cases are estimated to be missing from the merged RETI&NbCDB during the period 1999–2017. Hence, final completeness estimates of 84% and 67% were obtained for the RETI and NbCDB respectively, figures slightly lower than those obtained taking the RPBCRs-IICC-3 as reference (91% and 73% respectively; Table 2).

The linked files showed completeness of around 100% for age groups 0–4 and 0–14 years, for which the RPBCRs-IICC-3 set was the reference. For ages 0–18 months, a result > 100% was obtained (Table 2).

Due to the low number of cases from the RPBCRs, there were no significant differences by sex or age between this and the RETI, NbCDB, RETI&NbCDB or (RETI&NbCDB)&RPBCRs sets. The same applied to the RPBCRs-only, RETI-only, and NbCDB-only subsets. Table 4 shows similarities between the RETI, NbCDB, RETI&NbCDB and RPBCRs files.

Few cases were lost to follow-up. Overall survival was around 90% in children aged 0–18 months and about 74% in the 0–14 year age group. For children older than 18 months, survival in the set with a higher number of cases varied from 58% for children up to 4 years old (n = 389), to 53% for ages 5–9 years (n = 114) and 61% for the older group (n = 61). RPBCRs-only cases (n = 23) registered the highest survival. NbCDB-only cases showed the lowest survival at ages 0–18 months (n = 72) and the highest at ages 19 months to 4 years (n = 49) (Table 5). Survival was similar, without statistically significant differences in the sets with more than 800 cases. The most recently diagnosed patients survived slightly longer than did patients in the past (Supplementary Table 4).

Table 3
Linkage of Nb in Spain study. Indicators of contribution and deficit of each file entering the linkages by age group, 1999–2017.

Linkages		Co%	SCo%	URT%	Ov%	OvT%
1st Linkage Result: RETI&NbCDB						
0–14 years	RETI	88.3	29.4	11.7	66.8	58.9
	NbCDB	70.6	11.7	29.4	83.4	
0–18 months	RETI	87.8	28.5	12.2	67.5	59.2
	NbCDB	71.5	12.2	28.5	82.9	
2nd Linkage: Result: (RETI&NbCDB)&RPBCRs						
0–14 years	RETI&NbCDB	98.4	80.8	1.6	17.9	17.6
	RPBCRs	19.2	1.6	80.8	91.5	
0–18 months	RETI&NbCDB	98.3	80.0	1.7	18.6	18.3
	RPBCRs	20.0	1.7	80.0	91.4	

NOTES
-Nb: Neuroblastoma
-Co%: Contribution
-SCo%: Specific contribution
-URT%: Under-reporting of a file in relation to the total final set
-Ov%: overlapping between files
-OvT%: overlapping in total final set
-RETI: File of the Spanish Registry of Childhood Tumours
-NbCDB: File of the Spanish Neuroblastoma Clinical Database.
-RETI&NbCDB: Complete set resulting from the 1st linkage, RETI and NbCDB files
-RPBCRs: File of the regional population-based cancer registries participating in the study (see Supplementary Table 2)
-(RETI&NbCDB)&RPBCRs: Complete set resulting from the 2nd linkage, RETI&NbCDB and RPBCRs files

4. Discussion

The JARC [8] is a project supported by the European Union and Member States aimed at advancing the quality of care for and research on rare cancers. Task 4.4, WP4 aspires to evaluate the impact of ERNs, by linking clinical databases and population-based cancer registries to monitor the proportion of childhood cancers benefiting from the ERNs, and the coverage and completeness of epidemiological and clinical sources in the target population.

Nb was chosen because it has a dedicated national clinical database in Spain, international clinical research collaboration in Europe established under SIOPEN [28], and national clinical databases in most countries. As regards cancer registration, Spain has yet to complete its national childhood cancer registration. Two systems coexist: several RPBCRs (described above); and the RETI, which is striving for national completeness [16]. Additionally, every SEHOP clinical working group maintains its database for clinical and outcome evaluations. The NbCDB was developed to co-ordinate sample delivery, trial recruitment, second consultations, and clinical follow-up. Accordingly, the NbCDB, RETI and 10 RPBCRs were included in the study.

In Spain, children do not have a national identification number, whereas adults are required by law to have one for purely administrative purposes. Since there is no unique patient identifier in Spain for clinical, public health or research issues, linkage had to be semi-automatic, using personal identification details. Manual crossover was required in 54% of matches: as this is neither efficient nor sustainable, we recommend implementing a unique (national/European) personal health identifier capable of tracking patients in the health system and research environment, such as the European Unified Patient Identity Management (EUPID20) [29].

Completeness of expected Nb incidence in Spain was insufficient in both RETI and NbCDB. However, the merged RETI&NbCDB and

Table 4
Linkage of Nb in Spain study. Differences by age, sex, year of diagnosis and survival between the main files and sets resulting from the linkages.

	RETI	NbCDB	RETI&NbCDB	RPBCRs
N	1486	1189	1683	329
Age in years (mean (st err))	1.80 (0.065)	1.74 (0.070)	1.78 (0.061)	1.71 (0.137)
Age 0–18 months % respect 0–14years)	54.0%	55.0%	54.4%	56.5%
Sex (% female)	46.2%	44.9%	46.3%	46.8%
Year of diagnosis (mean (st err))	2008.52 (0.138)	2008.34 (0.141)	2008.40 (0.129)	2006.53 (0.250)
5-year survival (st err) Age: 0–18 months	90.2 (1.3)	90.1 (1.4)	89.2 (1.3)	90.4 (2.4)
5-year survival (st err) Age: 0–14 years	73.6 (1.4)	73.5 (1.5)	74.0 (1.3)	73.3 (2.7)

NOTES
-Nb: Neuroblastoma
-RETI: File of the Spanish Registry of Childhood Tumours
-NbCDB: File of the Spanish Neuroblastoma Clinical Database
-RETI&NbCDB: Set resulting from the linkage of files RETI and NbCDB
-RPBCRs: File of the regional population-based cancer registries participating in the study (see Supplementary Table 2)

Table 5
Linkage of Nb in Spain study. 5-year survival by age group in the main files, sets, and subsets resulting from the linkages, with percentage of cases lost to follow-up. Period of diagnosis: 1999–2011.

Files/sets/subsets	Group of age								0–14 years		
	0–18 months		19 months-4 years		5–9 years		10–14 years				
	No. of cases	5-year survival (95% CI)	No. of cases	5-year survival (95% CI)	No. of cases	5-year survival (95% CI)	No. of cases	5-year survival (95% CI)	**No. of cases**	**% Lost**	**5-year survival (95% CI)**
RETI	529	90 (88–93)	334	55 (50–60)	102	54 (44–64)	26	54 (35–73)	**991**	**0.0**	**74 (71–76)**
NbCDB	446	90 (87–93)	282	56 (50–62)	82	46 (36–57)	18	61 (39–84)	**828**	**0.4**	**74 (71–76)**
RPBCRs	156	90 (86–95)	95	53 (43–63)	26	46 (27–65)[a]			**277**	**0.0**	**73 (68–79)**
RETI&NbCDB	601	89 (87–92)	383	58 (53–63)	113	52 (43–61)	29	59 (41–76)	**1126**	**0.3**	**74 (71–77)**
Comm-RETI&NbCDB	374	92 (89–94)	233	52 (45–58)	71	48 (36–59)	15	53 (28–79)	**693**	**0.0**	**73 (70–76)**
(RETI&NbCDB)&RPBCRs	615	89 (87–92)	389	58 (53–63)	114	53 (43–62)	31	61 (44–78)	**1149**	**0.3**	**74 (72–77)**
RETI-only	155	87 (81–92)	101	62 (53–72)	42	64 (50–79)[a]			**298**	**0.0**	**75 (70–80)**
NbCDB-only	72	82 (73–91)	49	77 (66–89)	14	50 (24–76)[a]			**135**	**2.2**	**77 (70–84)**
RPBCRs-only	14	b	6	b	1	b	2	b	**23**	**0.0**	**87 (73–100)[c]**

NOTES
-Nb: Neuroblastoma
-RETI: File of the Spanish Registry of Childhood Tumours
-NbCDB: File of the Spanish Neuroblastoma Clinical Database
-RPBCRs: File of the regional population-based cancer registries participating in the study (see Supplementary Table 2)
-RETI&NbCDB: Set resulting from the linkage of files RETI and NbCDB
-Comm-(RETI&NbCDB): Subset of cases common to RETI and NbCDB
-(RETI&NbCDB)&RPBCRs: Set resulting from the linkage of files RETI&NbCDB and RPBCRs
-RETI-Only: Subset of cases present in RETI but not in NbCDB
-NbCDB-Only: Subset of cases present in NbCDB but not in RETI
-RPBCRs-Only: Subset of cases present in RPBCRs but not in (RETI&NbCDB)
-a: Group of age 5–14 years
-b: Survival not calculated because < 15 cases were available
-c: 95% CI truncated at 100

(RETI&NbCDB)&RPBCRs files approached full completeness when the RPBCRs-IICC-3 set was used as reference (0–4 and 0–14 age groups), though some overestimation cannot be ruled out (see limitations below). Nevertheless, the result for the 0–18 month age group, i.e., completeness > 100%, is illogical. The following explanation might at least partially account for this: since completeness for cases aged 0–18 months had to be calculated by taking the rate obtained from the RETI's high-completeness area as reference, the observed overestimate of completeness for the merged file could be the effect of an underestimate of expected cases in the 0–18 month age group. This underestimate would have been the result of a deficit in registration of 0–18 month cases in the RETI and its high-completeness area, which would have produced an erroneously low reference rate and thus a correspondingly low number of expected cases. Concurrently, clinical experience shows that there might be a small number of cases not registered in the RETI (low-stage cases diagnosed in neonatology or local hospitals and not referred to paediatric oncology units, due to spontaneous regression or only surgical treatment). Hence, the observed overestimate of completeness must be assumed to have highlighted a specific RETI limitation for patients aged 0–18 months. Furthermore, some artefacts may have spuriously increased the number of observed cases in all ages, including the undetected presence of cases not resident in Spain, which cannot be definitively ruled out (see limitations below).

The specific contribution of the participating RPBCRs to the final set (RETI&NbCDB)&RPBCRs was small (1.6%), as they were substantially overlapped by the RETI and NbCDB. All but 28 of the RPBCRs' Nb cases had been included in clinical studies and/or treated at Nb units, suggesting good access to specialised paediatric oncological care and complementarity of the databases, which could be improved by making them interoperable. Importantly for paediatric oncologists, missing cases were identified in the linkage of the NbCDB and RETI, helping clinician researchers to retrieve such patents, improve NbCDB completeness, and understand how these patients were treated and followed up.

This study has several limitations: (1) the NbCDB could not be restricted to regions covered by the RPBCRs for linkage purposes because residence data are not recorded in the NbCDB. We therefore recommend introducing patient place/province of residence in clinical databases; (2) some duplications were likely present in the RETI&NbCDB file, due to matching errors; (3) some cases not resident in Spain may be present in the linked files, thus contributing to an overestimate of completeness; (4) five RPBCRs accounting for 8% of coverage were absent from the analysis. Nevertheless, since all RPBCRs in Spain covered 41% of the total childhood population, even if these 5 registries had participated, total completeness for all of Spain would not have been achieved; (5) the RPBCRs participating in the study covered only 33% of the child population in Spain; (6) the reference rate used for the completeness estimation was based on the 12 RPBCRs included in IICC-3 [9], which together covered 35% of the child population (Supplementary Table 1); and (7) the time-span covered by the RPBCRs-IICC-3 was 1990–2013, whereas the study period was 1999–2017. This temporal discrepancy might have led to a slight overestimate of completeness, since childhood cancer incidence gradually increased in southern Europe from 1991 to 2010 [18].

Despite some incompleteness, the RETI&NbCDB file is currently the most complete available set of cases for studying Nb survival in Spain. Further analyses should be conducted using clinical criteria which were not the focus of this paper. Our 5-year survival results are consistent with those in Europe [4]. Age-specific survival comparisons are difficult because we applied clinical age stratification [10]. In general, there were no significant differences in 5-year survival. The discrepancies in survival between the NbCDB-only and common-RETI&NbCDB subsets are minimal and should be clarified by examining the clinical data, identifying the possible effects of a lack of follow-up, and analysing individual cases and registration practice from 1999 to 2011, when several trials (INES [30], EUNS [31]) and interim studies were ongoing.

From our standpoint, the absence of differences in survival among the larger subsets indicates good access to care, and confirms the good work and quality of clinical practices and evaluation by the different information systems (clinical/ epidemiological).

The small differences in incidence and survival reaffirm the complementarity of the two files combined in the RETI&NbCDB set. We therefore recommend integrating the NbCDB and RETI, together with RPBCRs, thereby approximating population-based levels, and enabling dynamic, updated survival analyses without placing constraints on the development of the clinical database.

To assess the extent to which the target population accesses appropriate healthcare for childhood cancer and its subsequent outcome, it is essential that Spain, like other European countries, achieve full national registration of childhood cancer [16]. According to JARC's Rare Cancer Agenda 2030 [32], registries must incorporate variables such as the Toronto Staging Guidelines [33,34]. Moreover, a stable interconnection between the national registry and the clinical databases of the relevant healthcare providers (including ERNs) and SEHOP clinical groups should be established for all childhood cancers. To this end, a unique personal patient identifier common to all databases, systematic cross-linkage, and appropriate IT tools are needed. European data-protection regulations [35] must be positively implemented to ensure that clinical and population-based databases are interoperable, and so allow for more efficient evaluation of outcomes.

Role of the funding source

The funding sources had no role in study conception or design, data collection, data analysis, data interpretation, writing of the paper or in the decision to submit for publication.

Funding

This work is part of the project "Joint Action on Rare Cancers"/JARC funded by the European Union's Health Programme (2014–2020) [Grant No. 724161]. The Spanish Registry of Childhood Tumours has the support of the Ministry of Health of Spain, the University of Valencia and the Spanish Society of Paediatric Oncology; as well as the collaboration of the Enriqueta Villavecchia Foundation.

CRediT authorship contribution statement

A.C.N., G.G., R.C., and R.P.B. conceived and designed the study. E.P.R. and V.S.C., in collaboration with A.F.T., J.G.P., and W.G. members acquired the data for this study. A.M.L., E.P.R., and V.S.C. performed the quality control. A.M.L. did the statistical analyses. A.C.N., A.M.L., E.P.R., G.G., R.C., and R.P.B. interpreted the results. A.C.N., G.G., R.C., and R.P.B. drafted the manuscript, which was reviewed and approved in its final version by all authors.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the Ministry of Health of Spain, the Spanish Society of Haematology and Paediatric Oncology, and the Spanish Network of Cancer Registries for endorsing this study. We also thank K. Pritchard Jones, A. Trama, M. Regaña, and J. Calabuig for their suggestions; V. Castel and D. Ramal for their work with the Spanish Neuroblastoma Clinical Database; J.M. Borras and J. Ribes for the DALINK program; and R. Fernández-Delgado and M.S. Barreda for their continuous help in the Spanish Registry of Childhood Tumours.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.canep.2022.102145.

References

[1] E. Steliarova-Foucher, M. Colombet, L.A.G. Ries, F. Moreno, A. Dolya, F. Bray, et al., International incidence of childhood cancer, 2001-10: a population-based registry study, Lancet Oncol. 18 (2017) 719e31, https://doi.org/10.1016/S1470-2045(17)30186-9.
[2] G. Gatta, A.P. Dei Tos, J.M. Van Der Zwan, P.G. Casali, S. Siesling, A. Tavilla, et al., Rare cancers are not so rare: the rare cancer burden in Europe, Eur. J. Cancer 47 (2011) 2493–2511, https://doi.org/10.1016/j.ejca.2011.08.008.
[3] G. Gatta, A. Trama, R. Capocaccia, RARECARENet Working Group. Epidemiology of rare cancers and inequalities in oncologic outcomes, Eur. J. Surg. Oncol. 45 (2019) 3–11, https://doi.org/10.1016/j.ejso.2017.08.018.
[4] G. Gatta, L. Botta, S. Rossi, T. Aareleid, M. Bielska-Lasota, J. Clavel, et al., Childhood cancer survival in Europe 1999-2007: results of EUROCARE-5-a population-based study, Lancet Oncol. 15 (2014) 35–47, https://doi.org/10.1016/S1470-2045(13)70548-5.
[5] S. Mathoulin-Pélissier, K. Pritchard-Jones, Evidence-based data and rare cancers: the need for a new methodological approach in research and investigation, Eur. J. Surg. Oncol. 45 (2019) 22–30, https://doi.org/10.1016/j.ejso.2018.02.015.
[6] European Reference Networks. Available from: ⟨https://ec.europa.eu/health/ern_en⟩ [accessed 23 November 2020].
[7] ERN PAEDCAN. European Reference Network for Paediatric Oncology. ⟨https://paedcan.ern-net.eu/⟩ [accessed 23 May 2021].
[8] Joint Action on Rare Cancers (JARC). ⟨https://jointactionrarecancers.eu/⟩ [accessed 23 November 2020].
[9] E. Steliarova-Foucher, M. Colombet, L.A.G. Ries, P. Hesseling, F. Moreno, H. Y. Shin, et al. (Eds.), International Incidence of Childhood Cancer, Volume III (electronic version), International Agency for Research on Cancer, Lyon, France, 2017. ⟨http://iicc.iarc.fr/results/⟩. accessed 23 November 2020.
[10] W.B. London, R.P. Castleberry, K.K. Matthay, A.T. Look, R.C. Seeger, H. Shimada, et al., Evidence for an age cutoff greater than 365 days for neuroblastoma risk group stratification in the Children's Oncology Group, J. Clin. Oncol. 23 (2005) 6459–6465, https://doi.org/10.1200/JCO.2005.05.571.
[11] S.L. Cohn, A.D. Pearson, W.B. London, T. Monclair, P.F. Ambros, G.M. Brodeur, et al., The International Neuroblastoma Risk Group (INRG) Classification System: An INRG Task Force Report, J. Clin. Oncol. 27 (2008) 289–297, https://doi.org/10.1200/JCO.2008.16.6785.
[12] E. Steliarova-Foucher, C. Stiller, B. Lacour, et al., International Classification of Childhood Cancer, Cancer 103 (2005) 1457–1467, https://doi.org/10.1002/cncr.20910.
[13] Instituto Nacional de Estadística (INE). ⟨http://www.ine.es⟩ [accessed 27 March 2019].
[14] V. Castel Sánchez, C. Melero Moreno, P. García-Miguel García-Rosados, A. Navajas Gutiérrez, J.I. Ruiz Jiménez, S. Navarro Fos, et al., Neuroblastoma en niños menores de 1 año, An Esp Pediat. 47 (1997) 584–590.
[15] A. Cañete Nieto, E. Pardo Romaguera, A. Muñoz López, S. Valero Poveda, S. Porta Cebolla, M.S. Barreda Reines, R. Peris Bonet, Cáncer infantil en España. Estadísticas 1980-2020. Registro Español de Tumores Infantiles ((RETI-SEHOP)), Universitat de València,, Valencia, 2021 (Available from), ⟨https://www.uv.es/rnti/informes.html⟩.
[16] E. Steliarova-Foucher, C. Stiller, M. Colombet, P. Kaatsch, R. Zanetti, R. Peris-Bonet, Registration of childhood cancer: moving towards pan-European coverage? Eur. J. Cancer 51 (2015) 1064–1079, https://doi.org/10.1016/j.ejca.2015.03.009.
[17] European Network of Cancer Registries (ENCR). Available from: ⟨https://www.encr.eu/encr-members-contact-list⟩ [accesed 3 January 2022].
[18] E. Steliarova-Foucher, M.M. Fidler, M. Colombet, B. Lacour, P. Kaatsch, M. Piñeros, et al., Changing geographical patterns and trends in cancer incidence in children and adolescents in Europe, 1991-2010 (Automated Childhood Cancer Information System): a population-based study, Lancet Oncol. 19 (2018) 1159–1169.
[19] REDECAN. Red Española de Registros de Cáncer. Available from: ⟨https://redecan.org/redecan.org/es/page04b2.html?id=14&title=quienes-somos⟩ [accessed 23 November 2020].
[20] Cancer Incidence in Five Continents (CI5C). Available from: ⟨https://ci5.iarc.fr⟩ [accesed 3 January 2022].
[21] M. Segi, Cancer mortality for selected sites in 24 countries (1950–57), Department of Public Health, Tohoku University of Medicine,, Sendai, Japan, 1960.
[22] J.A. Ross, R.K. Severson, B.H. Pollock, L.L. Robinson, Childhood cancer in the United States, Cancer 77 (1996) 201–207, https://doi.org/10.1002/(SICI)1097-0142(19960101)77:1<201::AID-CNCR32>3.0.CO;2-7.
[23] J.P. Vandenbroucke (Ed.), Am. J. Epidemiol. 115 (1982) 303–304.
[24] J. Gomez-Bao, J.L. Larriba-Pey, J. Ribes Puig, Record linkage performance for large data sets. PAVLAD '09 Proceedings of the ACM First International Workshop on Privacy and Anonymity for Very Large Databases, ACM,, New York, NY, USA, 2009, pp. 9–16, https://doi.org/10.1145/1651449.1651453.
[25] M.F. Abad Garcia, Investigación Evaluativa en Documentación. Aplicación a la Documentación Médica, Universitat de València, Valencia, 1997. ISBN 84-370-3305-5.

[26] E.L. Kaplan, P. Meier, Nonparametric estimation from incomplete observations, J. Am. Statist. Assoc. 53 (282) (1958) 457–481, https://doi.org/10.1080/01621459.1958.10501452.
[27] R. Peto, M.C. Pike, P. Armitage, N.E. Breslow, D.R. Cox, S.V. Howard, et al., Design and analysis of randomized clinical trials requiring prolonged observation of each patient II Analysis and examples, Br. J. Cancer 35 (1977) 1–39, https://doi.org/10.1038/bjc.1977.1.
[28] SIOPEN. Available from: ⟨https://www.siopen.net/⟩ [accessed 23 November 2020].
[29] H. Ebner, D. Hayn, M. Falgenhauer, M. Nitzlnader, G. Schleiermacher, R. Haupt, et al., Piloting the European Unified Patient Identity Management (EUPID) Concept to facilitate secondary use of neuroblastoma data from clinical trials and biobanking, Stud. Health Technol. Inform. 223 (2016) 31–38, https://doi.org/10.3233/978-1-61499-645-3-31.
[30] H. Rubie, B. De Bernardi, M. Gerrard, A. Canete, R. Ladenstein, J. Couturier, et al., Excellent outcome with reduced treatment in infants with non-metastatic and unresectable neuroblastoma without MYCN amplification: results of the prospective INES 99.1, J. Clin. Oncol. 29 (2011) 449–455, https://doi.org/10.1200/JCO.2010.29.5196.
[31] J.A. Kohler, H. Rubie, V. Castel, K. Beiske, K. Holmes, C. Gambini, F. Casale, C. Munzer, G. Erminio, S. Parodi, S. Navarro, C. Marquez, M. Peuchmaur, C. Cullinane, P. Brock, D. Valteau-Couanet, A. Garaventa, R. Haupt, Treatment of children over the age of one year with unresectable localised neuroblastoma without MYCN amplification: results of the SIOPEN study, Eur. J. Cancer 49 (2013) 3671–3679, https://doi.org/10.1016/j.ejca.2013.07.002.
[32] PG Casali, A. Trama (Eds), Rare Cancer Agenda 2030. Ten recommendations from the EU Joint Action on Rare Cancers. Lecce, It. Yucanprint, 2020 [ISBN 978–88-31642–71-2]. Available from: ⟨https://www.jointactionrarecancers.eu/attachments/article/265/Rare_Cancer_Agenda_2030.pdf⟩ [accessed 23 November 2020].
[33] J.F. Aitken, D.R. Youlden, A.S. Moore, P.D. Baade, L.J. Ward, V.J. Thursfield et al. Childhood cancer staging for population registries according to the Toronto Childhood Cancer Staging Guidelines. Cancer Council Queensland Australia. Brisbane, Australia, 2017. Available from: ⟨https://cancerqld.blob.core.windows.net/content/docs/childhood-cancer-staging-for-population-registries.pdf⟩ [accessed 23 November 2020].
[34] J.F. Aitken, D.R. Youlden, A.S. Moore, P.D. Baade, L.J. Ward, V.J. Thursfield, P.C. Valery, A.C. Green, S Gupta, AL Frazier, Childhood cancer staging for population registries according to the Toronto Childhood Cancer Stage Guidelines. Cancer Council Queensland and Cancer Australia: Brisbane, Australia; 2017. [Estadificación del cáncer infantil para registros de base poblacional. Traducida por el Grupo de Trabajo del RETI en la SEHOP, 2020. DOI: 10.7203/72724]. Available from: ⟨http://doi.org/10.7203/72724⟩ [accessed 23 November 2020].
[35] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Available from: ⟨http://data.europa.eu/eli/reg/2016/679/oj⟩ [accessed 23 November 2020].