BMJ Open

Health care utilisation preceding relapse or second malignant neoplasm after childhood acute lymphoblastic leukaemia: a population-based matched cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-050285
Article Type:	Original research
Date Submitted by the Author:	15-Feb-2021
Complete List of Authors:	Jensen, Karen; Aarhus University Hospital, Department of Paediatrics and Adolescent Medicine Klug Albertsen, Birgitte; Aarhus University Hospital, Department of Pediatrics and Adolescent Medicine Schrøder, Henrik; Aarhus University Hospital, Department of Paediatrics and Adolescent Medicine Zalounina Falborg, Alina; Aarhus University, Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health Schmiegelow, Kjeld; Copenhagen University Hospital, Department of Paediatrics and Adolescent Medicine; University of Copenhagen, Institute of Clinical Medicine, Faculty of Health and Medical Sciences Rosthøj, Steen ; Aalborg University Hospital, Department of Paediatrics and Adolescent Medicine Callesen, Michael Thude; Odense University Hospital, Hans Christian Andersen Children's Hospital Vedsted, Peter; Aarhus Universitet, Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health
Keywords:	EPIDEMIOLOGY, Leukaemia < HAEMATOLOGY, Paediatric oncology < PAEDIATRICS, PRIMARY CARE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Health care utilisation preceding relapse or second malignant neoplasm after childhood acute lymphoblastic leukaemia: a population-based matched cohort study

Karen Schow Jensen[1], Birgitte Klug Albertsen[1], Henrik Schrøder[1], Alina Zalounina Falborg[2], Kjeld Schmiegelow[3,4], Steen Rosthøj[5], Michael Thude Callesen[6] and Peter Vedsted[2]

[1]Department of Paediatrics and Adolescent Medicine, Aarhus University Hospital, Aarhus, Denmark.

[2]Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health, Aarhus University, Aarhus, Denmark.

[3]Department of Paediatrics and Adolescent Medicine, Copenhagen University Hospital, Copenhagen, Denmark.

[4]Institute of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

[5]Department of Paediatrics and Adolescent Medicine, Aalborg University Hospital, Aalborg, Denmark.

[6]Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark.

Corresponding author: Karen Schow Jensen, Department of Paediatrics and Adolescent Medicine, Aarhus University Hospital, Palle Juul-Jensens Blvd. 99, 8200 Aarhus N, Denmark.

E-mail: kascje@rm.dk

Abstract: 293 words

Main text: 1731 words, one text box, one table and four figures

Supplementary data: Two tables

Abstract

Objectives To investigate health care utilisation including both primary and secondary health care six months before the diagnosis of a relapse or a second malignant neoplasm (SMN) in survivors of childhood acute lymphoblastic leukaemia (ALL).

Design and setting A Danish population-based matched cohort study linking multiple nationwide registries.

Participants A total of 622 childhood ALL 2.5-year event-free survivors diagnosed between 1994 and 2015. Cases were survivors developing a relapse or an SMN and references were survivors still in

first remission. Each case was matched with five references on age, sex, treatment protocol and risk group.

Primary outcome measures Consultations in general practice and hospital the last six months before relapse or SMN. Cases and references were compared with monthly incidence rate ratios (IRR) from negative binomial regression models.

Results Of the 622 childhood ALL survivors, 60 (9.6%) developed a relapse or an SMN. Health care utilisation in general practice increased among cases the last month before the event compared with references with an IRR of 2.71 (95% CI 1.71-4.28). Data showed a bimodal structure with a significantly increased number of visits four, five and six months before the event. Hospital health care utilisation increased two months before the event in cases with an IRR of 5.01 (95% CI 3.78-6.63) the last month before the event and an IRR of 1.94 (95% CI 1.32-2.85) the second-last month comparing cases and references.

Conclusions Survivors of childhood ALL developing a relapse or an SMN have a short period of increased health care utilisation before diagnosis. At hospital, this might be explained by pre-diagnostic examinations. In general practice, data suggest a bimodal structure with children later developing a relapse having more contacts also half a year before the relapse, which suggests that there could be early warnings.

Keywords

Leukaemia, Paediatric oncology, Epidemiology, Primary care

Strengths and limitations of this study

- The first study to investigate healthcare utilisation before a relapse or an SMN in survivors of childhood ALL.
- Use of complete nationwide registries with nearly no loss to follow-up linked on an individual level ensured that the study was population-based thus limiting selection bias.
- Outcome data are collected routinely and uniformly in the Danish healthcare system and potential misclassification is thus expected to be non-differential.
- A relatively small case group, leading to low statistical precision.
- Unmeasured confounding could be present.

Introduction

Five-year survival from childhood acute lymphoblastic leukaemia (ALL) now exceeds 90% with an event-free five-year survival of around 85%.[1] With increased survival rates, more survivors need

scheduled surveillance programmes for detection of possible late effects as well as screening for relapse of ALL or second malignant neoplasm (SMN). ALL survivors are known to have more chronic conditions (late effects) than their general population peers and to have increased use of both primary and secondary health care services after end of treatment.[2-14] Studies examining the occurrence of late effects have contributed with important knowledge to follow-up programmes. However, to the best of our knowledge, no studies have investigated the use of health care before a relapse or an SMN in survivors of childhood ALL.

Studies of health care use before a primary diagnosis of childhood ALL have revealed increased health care use 2-3 months before the diagnosis, thus reflecting a short period of symptoms.[15, 16] Adolescents and young adults are found to have a longer interval with increased primary health care use for 5-6 months before primary diagnosis.[17] Earlier studies indicate that the increased primary health care use could have a bimodal structure with the first peak 10-12 months before the primary diagnosis.[15]

Health care utilisation may reflect both the duration of symptoms before the diagnosis is established and the sectorial distribution of utilised care associated with these symptoms. Considerable focus is devoted to follow-up strategies for this group, and knowledge about the duration of increased health

care use and the sectorial distribution of patients' help-seeking behaviour is therefore highly relevant. To address this knowledge gap, we aimed to analyse health care utilisation in general practice and hospital during the six-month period preceding a relapse or an SMN in survivors of childhood ALL.

Methods

Study design and setting

This study is a nationwide, population-based, matched cohort study linking information from several Danish registries. We followed the RECORD guidelines for reporting of studies conducted using observational, routinely collected health data.[18] (Supplementary Table S1).

In Denmark, the health care system is tax-financed and available to all residents (population 5.8 million). All Danish citizens are assigned a unique identifier, the Civil Personal Registration (CPR) number. The CPR number follows every resident from birth to death; data extracted from Danish public registries were linked on an individual level using the CPR number.

Participants

Eligible subjects were patients (1.0-17.9 years) diagnosed with non-infant Philadelphia chromosome-negative B-cell precursor or T-lineage ALL between 1994 and 2015, identified in the Danish part of the Nordic Society of Paediatric Haematology and Oncology (NOPHO) ALL database. Cases were defined as childhood ALL survivors having a relapse or an SMN as the first event 2.5 years or more after primary diagnosis and before December 2017. Cases were matched 1:5 with childhood ALL survivors still in first remission with the same sex, age group (under 10 years or 10 years or more), NOPHO treatment protocol (ALL1992, ALL2000 or ALL2008) and risk group (high-risk or non-high risk) (see flow chart, Figure 1). Matching was based on incidence density sampling using the STATA command, sttocc. Due to the population-based design, the study sample size was determined by the number of cases in the area during the study period and no sample size calculation was performed.

Data sources and variables

Data were extracted from national registries (Text box) and hosted by Statistics Denmark. Authors had access to a de-identified data output. Data on health care utilisation were extracted for the period 1 January 1997 to 31 December 2017.

Text box. Data sources and variables

	Registries	Variables
Exposures	NOPHO ALL Registry*[1, 19]	Relapse of ALL
		First remission
	Danish Cancer Registry†[20]	Second malignant neoplasm
Outcomes		
Primary health care	National Health Insurance Service Register‡[21]	Daytime contacts to general practice:
		Daytime face-to-face contacts
		Email consultations
		Daytime telephone consultations
		Daytime home visits
		Out-of-hours contacts:
		Out-of-hours face-to-face contacts
		Out-of-hours telephone consultations
		Out-of-hours home visits
		Diagnostic procedures in general practice:
		Blood test
		Urine test
		Streptococcus throat test
		Pulmonary functions test
		Electrocardiogram
Secondary health care	Danish National Patient Registry§[22]	Contacts to public and private hospitals:
		Inpatient hospitalisations
		Outpatient visits

Covariates	Danish Civil Registration System[23]	Sex
		Age
		Vital status
		Immigration
		Emigration
	NOPHO ALL Registry	Diagnosis of childhood ALL
		Treatment protocol (ALL1992, ALL2000 or ALL2008)
		Risk group (high-risk or non-high risk)
		Immunophenotype (B-precursor ALL or T-ALL)

*NOPHO ALL Registry, Nordic Society of Paediatric Haematology and Oncology ALL Registry. The registry holds data on all children aged 1.0-14.9 years in Denmark diagnosed with ALL since 1992. From 2008 and onwards, the ALL Registry was extended to include children and adolescents aged 1.0-17.9 years.

†The Danish Cancer Registry holds information on all new cases of cancer in Denmark.

‡The National Health Insurance Service Register holds information on all contacts to general practice in Denmark. The following contacts were excluded: preventive health examination of children, vaccinations, screening for cervical cancer and pregnancy care. For a complete list of codes, see Supplementary data Table S2.

§The Danish National Patient Registry holds information on all contacts to public and private hospitals. The following contacts were excluded: visits to the accident and emergency department.

Statistical methods

The index date was the date of event (relapse or SMN) for cases. The corresponding index date for references was defined as the date with the same interval from the primary diagnosis as for the

case. For all included individuals, follow-up started no earlier than 2.5 years after diagnosis to ensure that treatment had ended and remission reached. Health care utilisation was assessed from six months before the index date/event.

The monthly rates for primary health care contacts (daytime contacts, out-of-hours contacts and diagnostic procedures) and hospital contacts (inpatient hospitalisations and hospital outpatient contact) were calculated as crude estimates for each of the six months preceding the index date. Negative binominal regression models were used to calculate incidence rate ratios (IRRs) to compare monthly rates of contacts between cases and references. Cluster robust variance estimation was applied to account for possible cluster effects at patient level. This was relevant as measurements on the same person were repeated monthly.

Estimates of IRRs were adjusted for sex, age and time since diagnosis. To adjust for age and time since diagnosis, we used restricted cubic splines with six knots to allow for a non-linear relationship. Furthermore, we performed analyses restricted to cases developing a relapse and to their references. All estimates are presented with 95% confidence intervals (CIs). All tests were two-sided and a P-value ≤0.05 was considered statistically significant. Data were analysed using the statistical software Stata 16.1 (StataCorp LLC, TX, USA).

Patient and public involvement

The study included no patient and public involvement.

Results

Patient characteristics

The study included 60 cases and 295 references; 49 (81.7%) of the 60 cases suffered a relapse and 11 (18.3%) an SMN (Table 1). In two cases, there were fewer than five matching references.

Table 1. Characteristics of the study population

Characteristic	Cases* N = 60	References† N = 295
Sex, n (%)		
Male	38 (63.3)	190 (64.4)
Female	22 (36.7)	105 (35.6)
Median age at index date‡, (IQI)	11.3 (8.4-16.1)	11.1 (7.7-15.7)
Age group at index date, n (%)		
Age < 10 years	21 (35.0)	130 (44.1)

Age ≥ 10 years	39 (65.0)	165 (55.9)
Treatment protocol, n (%)		
NOPHO ALL1992	24 (40.0)	120 (40.7)
NOPHO ALL2000	22 (36.7)	105 (35.6)
NOPHO ALL2008	14 (23.3)	70 (23.7)
Cell line, n (%)		
B-precursor ALL	55 (91.7)	253 (85.8)
T-ALL	5 (8.3)	42 (14.2)
Risk group, n (%)		
Non-high-risk	46 (76.7)	230 (78.0)
High-risk	14 (23.3)	65 (22.0)
Median time from diagnosis to index date (years, IQI)	3.8 (3.2-5.1)	3.8 (3.2-5.1)
Type of event, n (%)		
Relapse	49 (81.7)	-
SMN	11 (18.3)	-

*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.

†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

‡Index date, the date of event for cases and the corresponding date for references.

SMN, second malignant neoplasm; IQI, interquartile interval; NOPHO, Nordic Society of Paediatric Haematology and Oncology; ALL, acute lymphoblastic leukaemia.

Health care utilisation

We found a mean of 0.73 (95% CI: 0.53-1.02) daytime general practice visits during the month before the event in cases corresponding to an IRR of 2.71 (95% CI: 1.71-4.28). For the month before the event, we found an IRR of 8.12 (95% CI 3.01-21.86) for general practice out-of-hours contacts and an IRR of 5.89 (95% CI 2.44-14.21) for diagnostic procedures in general practice (Figure 2). For daytime general practice visits, data suggest a possible bimodal structure with increased IRRs during 4-6 months before the event.

For cases, hospital utilisation was 3.42 (95% CI 2.83-4.12) contacts in the last month before the event compared to 0.72 (95% CI 0.61-0.85) contacts for references, corresponding to an IRR of 5.01 (95% CI 3.78-6.63) the month before the event. For the second-last month before the event, we found an IRR of 1.94 (95% CI 1.32-2.85) (Figure 3).

In analyses restricted to cases developing a relapse, hospital utilisation also increased two months before the event (significantly increased only one month before the event). In general practice, data continued to suggest a bimodal structure (Figure 4).

Discussion

The present national, population-based matched cohort study shows that utilisation of general practice and hospital services increased significantly two months before the diagnosis of a relapse or an SMN compared to references still in first remission. This indicates that the diagnoses were made within two months. Our data showed a bimodal structure for daytime consultations in general practice in general and for cases developing a relapse in particular, with increased utilisation 5-6 months before relapse. This indicates that there could be early warnings. The increased use of hospital health care services the last month before relapse is most likely explained by the diagnostic workup.

Strengths and limitations

The population-based design with use of nationwide registries linked on an individual level is a strength. This ensured optimal completeness of data and follow-up. However, a relapse diagnosis is not registered in the Danish Cancer Registry. Therefore, data on relapses were collected from the NOPHO ALL registry.[1, 19] The NOPHO ALL registry is a very robust data source as it is updated regularly by research nurses and paediatric oncologists. Nevertheless, the registry might not contain data on all relapses that occur after patients leave a paediatric department. Children with a relapse

that was unregistered would belong to the reference group, which could lead to bias towards underestimating relapse frequency and the differences in use of health care services.

Electronic outcome data are collected routinely and uniformly in the Danish healthcare system. Data were collected for remuneration and not for the purpose of the present study. Potential misclassification of outcomes is expected to be equally distributed among cases and references, and any such misclassification is expected to be non-differential.[24]

The relatively small case group in our study is a limitation, leading to a low statistical precision with broad confidence intervals.

We compared periods with the same interval from diagnosis in cases and references as previous research has shown that time since diagnosis affects utilisation of health care.[2, 5, 8, 9] We made an effort to reduce confounding by age, gender, calendar period and treatment regime by matching cases with references. We were not able to adjust for sociodemographic factors and unmeasured confounding could thus be present. We expect potential bias to be negligible, and we believe that our findings can be generalised to other countries with comparable healthcare systems.

Comparison with existing literature

Previous studies on health care utilisation in ALL survivors have found increased use of primary and secondary heath care after end of treatment.[2-14] However, previous studies did not evaluate health care use before a relapse or an SMN. Studies on health care utilisation before primary ALL diagnosis in childhood found increased use of health care 2-3 months before the primary diagnosis;[15, 16] and based on these findings, we expected a short duration of increased health care use. Furthermore, a bimodal structure for general practice health care use before the primary diagnosis is reported, but with the first peak 10-12 months before diagnosis.[15]

A recent study examining use of health care before a cancer recurrence or an SMN in adult cancer survivors reported increased use of health care up to a year before diagnosis among patients diagnosed with a wide range of solid tumours.[25] Based on knowledge on health care use before a primary cancer, it is expected that patients with solid tumours have a longer interval of increased health care utilisation.[15, 17]

Conclusions

Survivors of childhood ALL developing a relapse or an SMN when in remission had a higher use of general practice and hospital health care services compared with matched references, 1-2 months

before the event. There was a possible bimodal structure for daytime visits to general practice with increased visits also 4-6 months before the event. As health care utilisation may be seen as a proxy for morbidity, this indicates that there could be early warnings. To the best of our knowledge, this is the first study to investigate use of health care before a relapse or an SMN in survivors of childhood ALL in remission, and further research is needed.

Figure legends

Figure 1 Flow diagram of the study population

Children with relapse/SMN and matched references in first remission.
*Matching on age group, sex, risk group and treatment protocol.
†The number in brackets is the number of unique persons – the same child can serve as a control more than once.
Ph-negative BCR-ALL or T-ALL, Philadelphia chromosome-negative B-cell precursor or T-lineage acute lymphoblastic leukaemia; CPR number, civil personal registration number; HSCT in CR1, haematopoietic stem cell transplantation in first complete remission; SMN, second malignant neoplasm.

Figure 2 General practice health care utilisation

General practice utilisation by months before event for cases* (n=60) compared with references† (n=295). (A) Daytime. (B) Out-of-hours. (C) Diagnostic procedures.
Top panel: Contacts/diagnostic procedure mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.
*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.

†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
ALL, acute lymphoblastic leukaemia; SMN, Second malignant neoplasm.

Figure 3 Hospital health care utilization

Hospital health care utilisation by months before event for cases* (n=60) compared with references† (n=295).
Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.
*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.
†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

Figure 4 Health care utilisation

Health care utilisation by months before event for cases (n=49) compared with references* (n=243).
Cases are survivors of childhood ALL developing a relapse as the first event (cases developing an SMN are excluded in this analysis).
Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.
*References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

Acknowledgments

The authors wish to thank Kaare Rud Flarup for his assistance with data management.

Footnotes

Contributors: KJ designed the study, analysed and interpreted data and wrote the manuscript; BA interpreted data and edited the manuscript; HS designed the study and edited the manuscript; AF analysed and interpreted data and edited the manuscript; KS interpreted data and edited the manuscript; SR interpreted data and edited the manuscript; MC interpreted data and edited the manuscript; PV designed the study, interpreted data and edited the manuscript. All authors approved the final manuscript.

Funding: This work was supported by a grant from the Danish Cancer Society (Kræftens Bekæmpelse). Grant number R124-A7831-15-S2. The funders had no role in the design and conduct of the study.

Competing interests: Birgitte Klug Albertsen declares the following: sponsor for the investigator initiated NOR-GRASPALL 2016 study. Kjeld Schmiegelow declares the following: Speaker and/or Advisory Board Honoraria from Jazz Pharmaceuticals (2020) and Servier (2020); speaker fee from

Amgen (2020) and Medscape (2020); Educational grant from Servier (2020). The remaining authors declare that they have no competing interests.

Patient consent for publication: Not required.

Ethical approval: This study was approved by the Danish Data Protection Agency (ID 277). Medical ethical approval was not required according to Danish law.

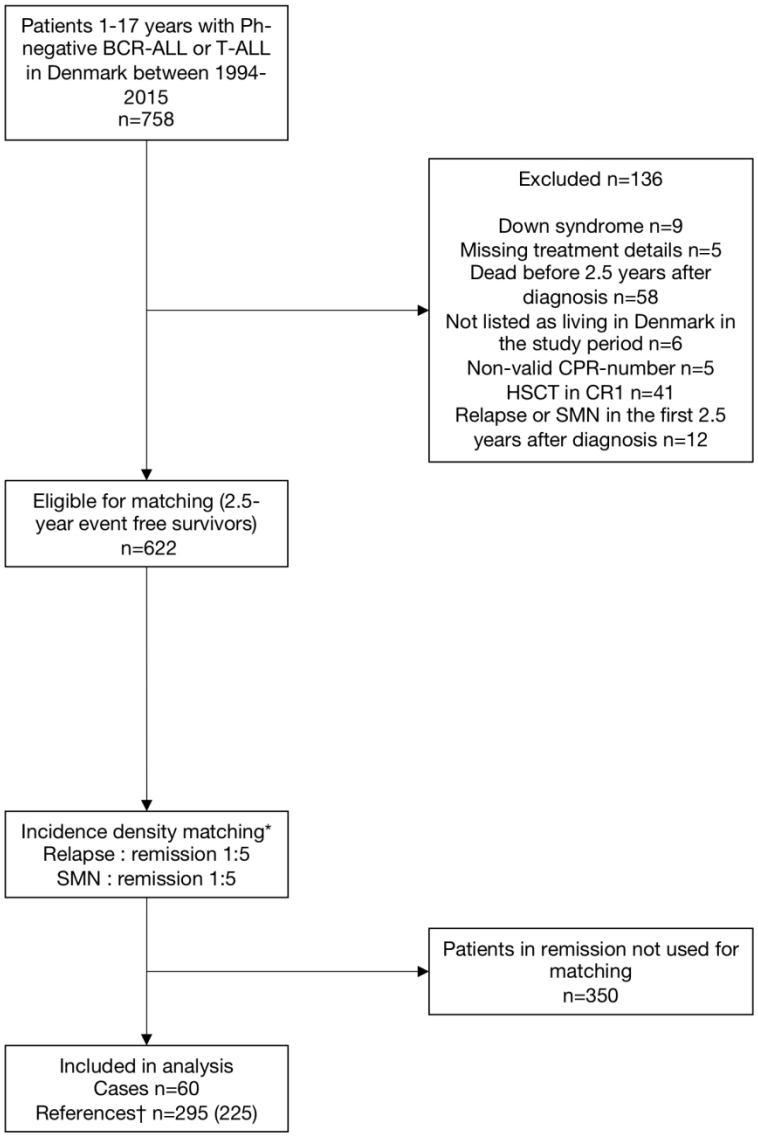
Data sharing statement: According to the data agreement with the data provider, we are not allowed to share our data. Data are stored and maintained electronically at Statistics Denmark.

References

1. Toft N, Birgens H, Abrahamsson J, et al. Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia. *Leukemia* 2018;32(3):606-15. doi: 10.1038/leu.2017.265 [published Online First: 2017/08/19]
2. de Fine Licht S, Rugbjerg K, Gudmundsdottir T, et al. Long-term inpatient disease burden in the Adult Life after Childhood Cancer in Scandinavia (ALiCCS) study: A cohort study of 21,297 childhood cancer survivors. *PLoS medicine* 2017;14(5):e1002296. doi: 10.1371/journal.pmed.1002296 [published Online First: 2017/05/10]
3. Font-Gonzalez A, Feijen E, Geskus RB, et al. Risk and associated risk factors of hospitalization for specific health problems over time in childhood cancer survivors: a medical record linkage study. *Cancer medicine* 2017;6(5):1123-34. doi: 10.1002/cam4.1057 [published Online First: 2017/04/06]
4. Sieswerda E, Font-Gonzalez A, Reitsma JB, et al. High Hospitalization Rates in Survivors of Childhood Cancer: A Longitudinal Follow-Up Study Using Medical Record Linkage. *PLoS One* 2016;11(7):e0159518. doi: 10.1371/journal.pone.0159518 [published Online First: 2016/07/21]
5. Sorensen GV, Winther JF, de Fine Licht S, et al. Long-Term Risk of Hospitalization Among Five-Year Survivors of Childhood Leukemia in the Nordic Countries. *Journal of the National Cancer Institute* 2019;111(9):943-51. doi: 10.1093/jnci/djz016 [published Online First: 2019/02/13]
6. Mulrooney DA, Hyun G, Ness KK, et al. The changing burden of long-term health outcomes in survivors of childhood acute lymphoblastic leukaemia: a retrospective analysis of the St Jude Lifetime Cohort Study. *The Lancet Haematology* 2019;6(6):e306-e16. doi: 10.1016/s2352-3026(19)30050-x [published Online First: 2019/05/13]
7. Bhakta N, Liu Q, Ness KK, et al. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). *Lancet (London, England)* 2017;390(10112):2569-82. doi: 10.1016/s0140-6736(17)31610-0 [published Online First: 2017/09/12]
8. Heins MJ, Lorenzi MF, Korevaar JC, et al. Non-oncology physician visits after diagnosis of cancer in adolescents and young adults. *Journal of cancer survivorship : research and practice* 2016;10(4):783-8. doi: 10.1007/s11764-016-0523-x [published Online First: 2016/02/15]
9. Heins MJ, Lorenzi MF, Korevaar JC, et al. Non-oncology physician visits after diagnosis of cancer in children. *BMC family practice* 2016;17:60. doi: 10.1186/s12875-016-0462-7 [published Online First: 2016/06/02]
10. Rebholz CE, Reulen RC, Toogood AA, et al. Health care use of long-term survivors of childhood cancer: the British Childhood Cancer Survivor Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011;29(31):4181-8. doi: 10.1200/jco.2011.36.5619 [published Online First: 2011/09/29]
11. Shaw AK, Pogany L, Speechley KN, et al. Use of health care services by survivors of childhood and adolescent cancer in Canada. *Cancer* 2006;106(8):1829-37. doi: 10.1002/cncr.21798 [published Online First: 2006/03/17]

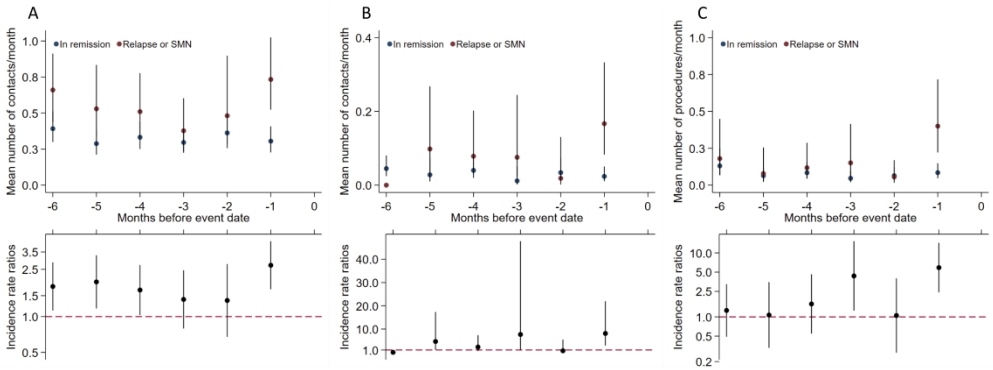
12. Oeffinger KC, Mertens AC, Hudson MM, et al. Health care of young adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *Annals of family medicine* 2004;2(1):61-70. [published Online First: 2004/04/01]
13. Streefkerk N, Heins MJ, Teepen JC, et al. The involvement of primary care physicians in care for childhood cancer survivors. *Pediatric blood & cancer* 2019;66(8):e27774. doi: 10.1002/pbc.27774 [published Online First: 2019/04/30]
14. McBride ML, Lorenzi MF, Page J, et al. Patterns of physician follow-up among young cancer survivors: report of the Childhood, Adolescent, and Young Adult Cancer Survivors (CAYACS) research program. *Canadian family physician Medecin de famille canadien* 2011;57(12):e482-90. [published Online First: 2011/12/16]
15. Ahrensberg JM, Fenger-Gron M, Vedsted P. Use of primary care during the year before childhood cancer diagnosis: a nationwide population-based matched comparative study. *PLoS One* 2013;8(3):e59098. doi: 10.1371/journal.pone.0059098 [published Online First: 2013/04/05]
16. Yang TO, Liu YL, Huang WT, et al. Specific and Non-specific Clinical Presentations in the Year Before the Diagnosis of Childhood Leukaemia. *Pediatric blood & cancer* 2016;63(8):1387-93. doi: 10.1002/pbc.26029 [published Online First: 2016/04/30]
17. Ahrensberg JM, Fenger-Gron M, Vedsted P. Primary Care Use before Cancer Diagnosis in Adolescents and Young Adults - A Nationwide Register Study. *PLoS One* 2016;11(5):e0155933. doi: 10.1371/journal.pone.0155933 [published Online First: 2016/05/21]
18. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS medicine* 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885 [published Online First: 2015/10/07]
19. Schmiegelow K, Forestier E, Hellebostad M, et al. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia* 2010;24(2):345-54. doi: 10.1038/leu.2009.251 [published Online First: 2009/12/17]
20. Gjerstorff ML. The Danish Cancer Registry. *Scandinavian journal of public health* 2011;39(7 Suppl):42-5. doi: 10.1177/1403494810393562 [published Online First: 2011/08/04]
21. Andersen JS, Olivarius Nde F, Krasnik A. The Danish National Health Service Register. *Scandinavian journal of public health* 2011;39(7 Suppl):34-7. doi: 10.1177/1403494810394718 [published Online First: 2011/08/04]
22. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scandinavian journal of public health* 2011;39(7 Suppl):30-3. doi: 10.1177/1403494811401482 [published Online First: 2011/08/04]
23. Schmidt M, Pedersen L, Sorensen HT. The Danish Civil Registration System as a tool in epidemiology. *European journal of epidemiology* 2014;29(8):541-9. doi: 10.1007/s10654-014-9930-3 [published Online First: 2014/06/27]
24. Erlangsen A, Fedyszyn I. Danish nationwide registers for public health and health-related research. *Scandinavian journal of public health* 2015;43(4):333-9. doi: 10.1177/1403494815575193 [published Online First: 2015/03/12]
25. Rasmussen LA, Jensen H, Virgilsen LF, et al. Healthcare utilisation in general practice and hospitals in the year preceding a diagnosis of cancer recurrence or second primary

cancer: a population-based register study. *BMC health services research* 2019;19(1):941. doi: 10.1186/s12913-019-4757-y [published Online First: 2019/12/07]



Flow diagram of the study population
Children with relapse/SMN and matched references in first remission.
*Matching on age group, sex, risk group and treatment protocol.
†The number in brackets is the number of unique persons – the same child can serve as a control more than once.
Ph-negative BCR-ALL or T-ALL, Philadelphia chromosome-negative B-cell precursor or T-lineage acute lymphoblastic leukaemia; CPR number, civil personal registration number; HSCT in CR1, haematopoietic stem cell transplantation in first complete remission; SMN, second malignant neoplasm.

136x190mm (300 x 300 DPI)



General practice health care utilisation

General practice utilisation by months before event for cases* (n=60) compared with references† (n=295). (A) Daytime. (B) Out-of-hours. (C) Diagnostic procedures.

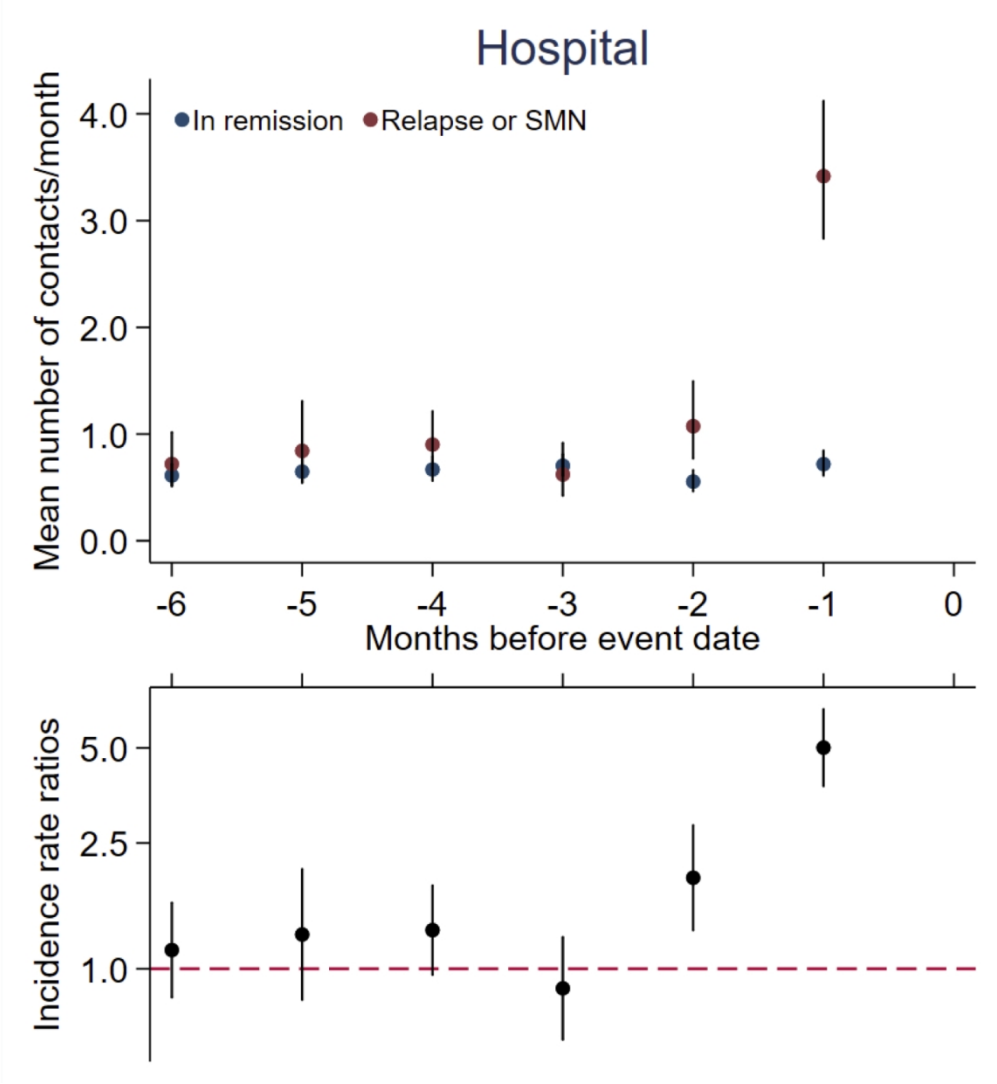
Top panel: Contacts/diagnostic procedure mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.

†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

ALL, acute lymphoblastic leukaemia; SMN, Second malignant neoplasm.

282x105mm (300 x 300 DPI)



Hospital health care utilization

Hospital health care utilisation by months before event for cases* (n=60) compared with references† (n=295).

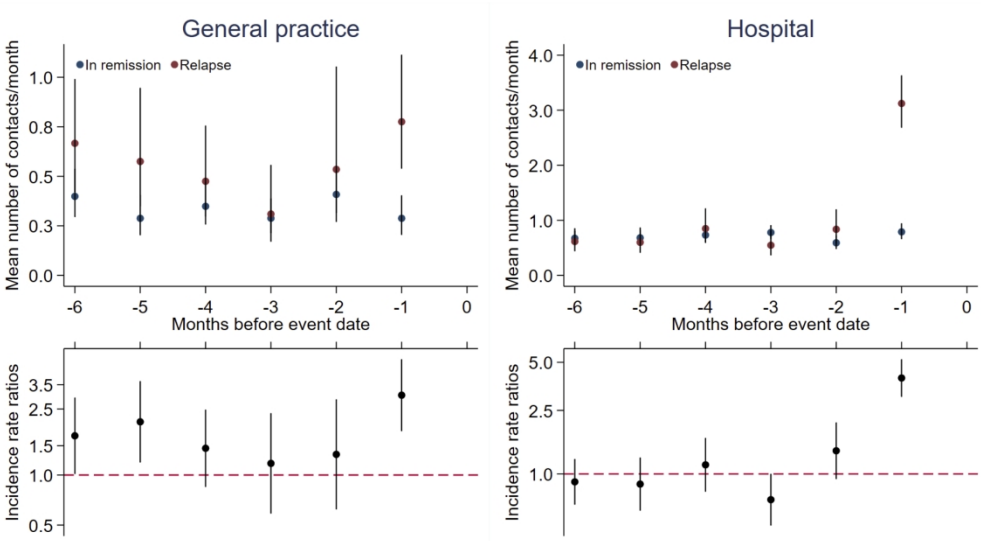
Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.

†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

133x144mm (300 x 300 DPI)



Health care utilisation

Health care utilisation by months before event for cases (n=49) compared with references* (n=243). Cases are survivors of childhood ALL developing a relapse as the first event (cases developing an SMN are excluded in this analysis).

Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

265x146mm (300 x 300 DPI)

Table S1

The RECORD statement – checklist of items extended from the STROBE statement to be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Pages 1 and 2	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Page 2 Page 2 Page 2
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Pages 3-4		
Objectives	3	State specific objectives, including any prespecified hypotheses	Page 4		
Methods					
Study Design	4	Present key elements of study design early in the paper	Page 4		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Pages 4-5		
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	Page 5	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not	Page 5 N/A

		Cross-sectional study - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case	Page 5	published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Figure 1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Pages 5-7	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Supplementary Table S2
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Text box		
Bias	9	Describe any efforts to address potential sources of bias	Pages 10-11		
Study size	10	Explain how the study size was arrived at	Page 5		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	N/A		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed	Page 7 Page 7 N/A N/A		

		Cross-sectional study - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Page 7		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Page 5 N/A
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Page 5
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	Figure 1 Figure 1 Figure 1	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Page 5 and Figure 1
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	Table 1 Table 1 Table 1		

Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure *Cross-sectional study* - Report numbers of outcome events or summary measures	Page 8		
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Pages 8-9 and Figures 2-3 Table 1 N/A		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	Page 9, Figure 4		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Page 10		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Pages 10-11	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Pages 10-11
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Page 11		
Generalisability	21	Discuss the generalisability (external validity) of the study results	Page 11		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study	Page 13		

		and, if applicable, for the original study on which the present article is based			
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Page 14

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

Table S2.

Information about contacts to general practice obtained from the Danish National Health Insurance Service Register

Contact type	Speciality code	Time code	Services
Daytime contacts	80	1	
			0101 0102 0105 0201 0411 0421 0431 0441 0451 0461 0491
Out-of-hours contacts	80	8 9	
			0101 0102 0471 0501
Diagnostic procedures	80	1 8 9	
Blood test			2101 2601 4309 4311 4312 4544 4611 7108 7110 7115 7120 7125 7126 7136 7150 7159 7168

			7177 7184 7186 7256 7263 7301 7302 7305 7309 7330 7403
Urine test			2132 4308 7101 7102 7122 7189
Strep throat test			4310 7109
Pulmonary functions test			4543 7113 7121 7183
Electrocardiogram			4313 7155 7156

BMJ Open

Health care utilisation preceding relapse or second malignant neoplasm after childhood acute lymphoblastic leukaemia: a population-based matched cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2021-050285.R1
Article Type:	Original research
Date Submitted by the Author:	15-Jun-2021
Complete List of Authors:	Jensen, Karen; Aarhus University Hospital, Department of Paediatrics and Adolescent Medicine Klug Albertsen, Birgitte; Aarhus University Hospital, Department of Pediatrics and Adolescent Medicine Schrøder, Henrik; Aarhus University Hospital, Department of Paediatrics and Adolescent Medicine Zalounina Falborg, Alina; Aarhus University, Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health Schmiegelow, Kjeld; Copenhagen University Hospital, Department of Paediatrics and Adolescent Medicine; University of Copenhagen, Institute of Clinical Medicine, Faculty of Health and Medical Sciences Rosthøj, Steen ; Aalborg University Hospital, Department of Paediatrics and Adolescent Medicine Callesen, Michael Thude; Odense University Hospital, Hans Christian Andersen Children's Hospital Vedsted, Peter; Aarhus Universitet, Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health
Primary Subject Heading:	Haematology (incl blood transfusion)
Secondary Subject Heading:	General practice / Family practice, Paediatrics
Keywords:	Leukaemia < HAEMATOLOGY, Paediatric oncology < PAEDIATRICS, PRIMARY CARE

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

Health care utilisation preceding relapse or second malignant neoplasm after childhood acute lymphoblastic leukaemia: a population-based matched cohort study

Karen Schow Jensen[1], Birgitte Klug Albertsen[1], Henrik Schrøder[1], Alina Zalounina Falborg[2], Kjeld Schmiegelow[3,4], Steen Rosthøj[5], Michael Thude Callesen[6] and Peter Vedsted[2]

[1]Department of Paediatrics and Adolescent Medicine, Aarhus University Hospital, Aarhus, Denmark.

[2]Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health, Aarhus University, Aarhus, Denmark.

[3]Department of Paediatrics and Adolescent Medicine, Copenhagen University Hospital, Copenhagen, Denmark.

[4]Institute of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.

[5]Department of Paediatrics and Adolescent Medicine, Aalborg University Hospital, Aalborg, Denmark.

[6]Hans Christian Andersen Children's Hospital, Odense University Hospital, Odense, Denmark.

Corresponding author: Karen Schow Jensen, Department of Paediatrics and Adolescent Medicine, Aarhus University Hospital, Palle Juul-Jensens Blvd. 99, 8200 Aarhus N, Denmark.

E-mail: kascje@rm.dk

Abstract: 300 words

Main text: 2015 words, one text box, one table and four figures

Supplementary data: Two tables

Abstract

Objectives To investigate health care utilisation including both primary and secondary health care six months before the diagnosis of a relapse or a second malignant neoplasm (SMN) in survivors of childhood acute lymphoblastic leukaemia (ALL).

Design and setting A Danish population-based matched cohort study linking multiple nationwide registries.

Participants Participants was recruited from a total of 622 childhood ALL 2.5-year event-free survivors diagnosed between 1994 and 2015. Cases were survivors developing a relapse or an

SMN and references were survivors still in first remission. Each case was matched with five references on age, sex, treatment protocol and risk group.

Primary outcome measures Consultations in general practice and hospital the last six months before relapse or SMN. Cases and references were compared with monthly incidence rate ratios (IRR) from negative binomial regression models.

Results Of the 622 childhood ALL survivors, 60 (9.6%) developed a relapse (49) or an SMN (11) and 295 matched references were identified. Health care utilisation in general practice increased among cases the last month before the event compared with references with an IRR of 2.71 (95% CI 1.71-4.28). Data showed a bimodal structure with a significantly increased number of visits four, five and six months before the event. Hospital health care utilisation increased two months before the event in cases with an IRR of 5.01 (3.78-6.63) the last month before the event and an IRR of 1.94 (1.32-2.85) the second-last month comparing cases and references.

Conclusions Survivors of childhood ALL developing a relapse or an SMN have a short period of increased health care utilisation before diagnosis. At hospital, this might be explained by pre-diagnostic examinations. In general practice, data suggest a bimodal structure with children later developing a relapse having more contacts also half a year before the relapse, suggesting that there could be early warnings.

Keywords

Leukaemia, Paediatric oncology, Primary care

Strengths and limitations of this study

- The first study to investigate healthcare utilisation before a relapse or an SMN in survivors of childhood ALL.
- Use of complete nationwide registries with nearly no loss to follow-up linked on an individual level ensured that the study was population-based thus limiting selection bias.
- Outcome data are collected routinely and uniformly in the Danish healthcare system and potential misclassification is thus expected to be non-differential.
- A small case group, leading to low statistical precision.
- Unmeasured confounding could be present.

Introduction

Five-year survival from childhood acute lymphoblastic leukaemia (ALL) now exceeds 90% with an event-free five-year survival of around 85%.[1] With increased survival rates, more survivors need

scheduled surveillance programmes for detection of possible late effects as well as screening for relapse of ALL or second malignant neoplasm (SMN). ALL survivors are known to have more chronic conditions (late effects) than their general population peers and to have increased use of both primary and secondary health care services after end of treatment.[2-14] Studies examining the occurrence of late effects have contributed with important knowledge to follow-up programmes. However, to the best of our knowledge, no studies have investigated the use of health care before a relapse or an SMN in survivors of childhood ALL.

Studies of health care use before a primary diagnosis of childhood ALL have revealed increased health care use 2-3 months before the diagnosis, thus reflecting a short period of symptoms.[15 16] Adolescents and young adults are found to have a longer interval with increased primary health care use for 5-6 months before primary diagnosis.[17] Earlier studies indicate that the increased primary health care use could have a bimodal structure with the first peak 10-12 months before the primary diagnosis.[15]

Health care utilisation may reflect both the duration of symptoms before the diagnosis is established and the sectorial distribution of utilised care associated with these symptoms. Considerable focus is devoted to follow-up strategies for this group, and knowledge about the duration of increased health

care use and the sectorial distribution of patients' help-seeking behaviour is therefore highly relevant. To address this knowledge gap, we aimed to analyse health care utilisation in general practice and hospital during the six-month period preceding a relapse or an SMN in survivors of childhood ALL.

Methods

Study design and setting

This study is a nationwide, population-based, matched cohort study linking information from several Danish registries. We followed the RECORD guidelines for reporting of studies conducted using observational, routinely collected health data.[18] (Supplementary Table S1).

In Denmark, the health care system is tax-financed and free and equally available to all residents (population 5.8 million). All children in Denmark developing ALL are treated in this tax-financed system ensuring that the study is population-based. After ALL treatment cessation, children in Denmark are followed in hospital-based outpatient surveillance programs; visits are scheduled 6-12 times the first year, 4-6 times the second year and 1-3 times a year the following years.[19] There are no scheduled visits in general practice.

All Danish citizens are assigned a unique identifier, the Civil Personal Registration (CPR) number. The CPR number follows every resident from birth to death; data extracted from Danish public registries were linked on an individual level using the CPR number.

Participants

Eligible subjects were patients (1.0-17.9 years) diagnosed with non-infant B-cell precursor or T-lineage ALL between 1994 and 2015, and treated according to three consecutive Nordic Society of Pediatric Hematology and Oncology (NOPHO) trials the ALL1992, ALL2000 and ALL2008 trials.[1 20] Participants were identified in the Danish part of the NOPHO ALL registry. Cases were defined as childhood ALL survivors having a relapse or an SMN as the first event 2.5 years or more after primary diagnosis and before December 2017. Cases were matched 1:5 with childhood ALL survivors still in first remission with the same sex, age group (under 10 years or 10 years or more), NOPHO treatment protocol (ALL1992, ALL2000 or ALL2008) and risk group (high-risk or non-high risk) (see flow chart, Figure 1). Matching was based on incidence density sampling using the STATA command, sttocc. Due to the population-based design, the study sample size was determined by the number of cases in the area during the study period and no sample size calculation was performed.

Data sources and variables

Data were extracted from national registries (Text box) and hosted by Statistics Denmark. Authors had access to a de-identified data output. Data on health care utilisation were extracted for the period 1 January 1997 to 31 December 2017. A relapse is defined as the reoccurrence of ALL after complete remission; a relapse can occur as an isolated bone marrow relapse, an isolated extramedullary relapse (e.g. the central nervous system or testis) or a combined bone marrow and extramedullary relapse. SMN is defined as the occurrence of a new malignant neoplasm. Survivors of ALL are at increased risk of developing a new malignant neoplasm compared to population peers; other haematological malignancies and tumours of the central nerves system are the most common types of SMNs.[21]

Text box. Data sources and variables

	Registries	Variables
Exposures	NOPHO ALL Registry*[1 20]	Relapse of ALL
		First remission
	Danish Cancer Registry†[22]	Second malignant neoplasm
Outcomes		
Primary health care	National Health Insurance Service Register‡[23]	Daytime contacts to general practice:
		Daytime face-to-face contacts

		Email consultations
		Daytime telephone consultations
		Daytime home visits
		Out-of-hours contacts:
		Out-of-hours face-to-face contacts
		Out-of-hours telephone consultations
		Out-of-hours home visits
		Diagnostic procedures in general practice:
		Blood test
		Urine test
		Streptococcus throat test
		Pulmonary functions test
		Electrocardiogram
Secondary health care	Danish National Patient Registry§[24]	Contacts to public and private hospitals:
		Inpatient hospitalisations
		Outpatient visits
Covariates	Danish Civil Registration System[25]	Sex
		Age
		Vital status
		Immigration
		Emigration
	NOPHO ALL Registry	Diagnosis of childhood ALL

	Treatment protocol (ALL1992, ALL2000 or ALL2008)
	Risk group (high-risk or non-high risk)
	Immunophenotype (B-precursor ALL or T-ALL)

*NOPHO ALL Registry, Nordic Society of Paediatric Haematology and Oncology ALL Registry. The registry holds data on all children aged 1.0-14.9 years in Denmark diagnosed with ALL since 1992. From 2008 and onwards, the ALL Registry was extended to include children and adolescents aged 1.0-17.9 years.
†The Danish Cancer Registry holds information on all new cases of cancer in Denmark.
‡The National Health Insurance Service Register holds information on all contacts to general practice in Denmark. The following contacts were excluded: preventive health examination of children, vaccinations, screening for cervical cancer and pregnancy care. For a complete list of codes, see Supplementary data Table S2.
§The Danish National Patient Registry holds information on all contacts to public and private hospitals. The following contacts were excluded: visits to the accident and emergency department.

Statistical methods

The index date was the date of event (relapse or SMN) for cases. The corresponding index date for references was defined as the date with the same interval from the primary diagnosis as for the case. For all included individuals, follow-up started no earlier than 2.5 years after diagnosis to ensure that treatment had ended and remission reached. Health care utilisation was assessed from six months before the index date/event.

The monthly rates for primary health care contacts (daytime contacts, out-of-hours contacts and diagnostic procedures) and hospital contacts (inpatient hospitalisations and hospital outpatient contact) were calculated as crude estimates for each of the six months preceding the index date. Negative binominal regression models were used to calculate incidence rate ratios (IRRs) to compare monthly rates of contacts between cases and references. Cluster robust variance estimation was applied to account for possible cluster effects at patient level. This was relevant as measurements on the same person were repeated monthly.

Estimates of IRRs were adjusted for sex, age and time since diagnosis. To adjust for age and time since diagnosis, we used restricted cubic splines with six knots to allow for a non-linear relationship. Furthermore, we performed analyses restricted to cases developing a relapse and to their references. All estimates are presented with 95% confidence intervals (CIs). All tests were two-sided and a P-value ≤0.05 was considered statistically significant. Data were analysed using the statistical software Stata 16.1 (StataCorp LLC, TX, USA).

Patient and public involvement

The study included no patient and public involvement.

Results

Patient characteristics

The study included 60 cases and 295 references; 49 (81.7%) of the 60 cases suffered a relapse and 11 (18.3%) an SMN (Table 1). In two cases, there were fewer than five matching references.

Table 1. Characteristics of the study population

Characteristic	Cases* N = 60	References† N = 295
Sex, n (%)		
Male	38 (63.3)	190 (64.4)
Female	22 (36.7)	105 (35.6)
Median age at index date‡, (IQI)	11.3 (8.4-16.1)	11.1 (7.7-15.7)
Age group at index date, n (%)		
Age < 10 years	21 (35.0)	130 (44.1)
Age ≥ 10 years	39 (65.0)	165 (55.9)
Treatment protocol, n (%)		
NOPHO ALL1992	24 (40.0)	120 (40.7)
NOPHO ALL2000	22 (36.7)	105 (35.6)
NOPHO ALL2008	14 (23.3)	70 (23.7)

Cell line, n (%)		
B-precursor ALL	55 (91.7)	253 (85.8)
T-ALL	5 (8.3)	42 (14.2)
Risk group, n (%)		
Non-high-risk	46 (76.7)	230 (78.0)
High-risk	14 (23.3)	65 (22.0)
Median time from diagnosis to index date (years, IQI)	3.8 (3.2-5.1)	3.8 (3.2-5.1)
Type of event, n (%)		
Relapse	49 (81.7)	-
SMN	11 (18.3)	-

*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.
†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
‡Index date, the date of event for cases and the corresponding date for references.
SMN, second malignant neoplasm; IQI, interquartile interval; NOPHO, Nordic Society of Paediatric Haematology and Oncology; ALL, acute lymphoblastic leukaemia.

Health care utilisation

We found a mean of 0.73 (95% CI: 0.53-1.02) daytime general practice visits during the month before the event in cases corresponding to an IRR of 2.71 (95% CI: 1.71-4.28). For the month before the event, we found an IRR of 8.12 (95% CI 3.01-21.86) for general practice out-of-hours contacts and an IRR of 5.89 (95% CI 2.44-14.21) for diagnostic procedures in general practice (Figure 2). For

daytime general practice visits, data suggest a possible bimodal structure with increased IRRs during 4-6 months before the event.

For cases, hospital utilisation was 3.42 (95% CI 2.83-4.12) contacts in the last month before the event compared to 0.72 (95% CI 0.61-0.85) contacts for references, corresponding to an IRR of 5.01 (95% CI 3.78-6.63) the month before the event. For the second-last month before the event, we found an IRR of 1.94 (95% CI 1.32-2.85) (Figure 3).

In analyses restricted to cases developing a relapse, hospital utilisation also increased two months before the event (significantly increased only one month before the event). In general practice, data continued to suggest a bimodal structure (Figure 4).

Discussion

The present national, population-based matched cohort study shows that utilisation of general practice and hospital services increased significantly two months before the diagnosis of a relapse or an SMN compared to references still in first remission. Our data showed a bimodal structure for daytime consultations in general practice in general and for cases developing a relapse more

pronounced, with increased utilisation 5-6 months before relapse. This indicates that there could be early warnings. The increased use of hospital health care services the last month before relapse is most likely explained by the diagnostic workup.

Strengths and limitations

The population-based design with use of nationwide registries linked on an individual level is a strength. This ensured optimal completeness of data and follow-up. However, a relapse diagnosis is not registered in the Danish Cancer Registry. Therefore, data on relapses were collected from the NOPHO ALL registry.[1 20] The NOPHO ALL registry is a very robust data source as it is updated regularly by research nurses and paediatric oncologists. Nevertheless, the registry might not contain data on all relapses that occur after patients leave a paediatric department. Children with a relapse that was unregistered would belong to the reference group, which could lead to bias towards underestimating relapse frequency and the differences in use of health care services.

Electronic outcome data are collected routinely and uniformly in the Danish healthcare system. Data were collected for remuneration and not for the purpose of the present study. Potential

misclassification of outcomes is expected to be equally distributed among cases and references, and any such misclassification is expected to be non-differential.[26]

The relatively small case group in our study is a limitation, leading to a low statistical precision with broad confidence intervals. Another limitation is the absence of information regarding the motivations for contacts to the healthcare system as this information is not available in the National Health Insurance Service Register.[23]

We compared periods with the same interval from diagnosis in cases and references as previous research has shown that time since diagnosis affects utilisation of health care.[2 5 8 9] We made an effort to reduce confounding by age, gender, calendar period and treatment regime by matching cases with references. We had no information on the amount and type of late effects and we were thus not able to match by late effects. However, previous studies suggest that the types of late effects have changed over calendar time making it relevant to match on treatment era (protocol).[6] We were not able to adjust for sociodemographic factors and unmeasured confounding could thus be present. We expect potential bias to be negligible, and we believe that our findings can be generalised to other countries with comparable healthcare systems.

Comparison with existing literature

Previous studies on health care utilisation in ALL survivors have found increased use of primary and secondary heath care after end of treatment.[2-14] However, previous studies did not evaluate health care use before a relapse or an SMN. Studies on health care utilisation before primary ALL diagnosis in childhood found increased use of health care 2-3 months before the primary diagnosis;[15 16] and based on these findings, we expected a short duration of increased health care use. Furthermore, a bimodal structure for general practice health care use before the primary diagnosis is reported, but with the first peak 10-12 months before diagnosis.[15]

A recent study examining use of health care before a cancer recurrence or an SMN in adult cancer survivors reported increased use of health care up to a year before diagnosis among patients diagnosed with a wide range of solid tumours.[27] Based on knowledge on health care use before a primary cancer, it is expected that patients with solid tumours have a longer interval of increased health care utilisation.[15 17]

Conclusions

Survivors of childhood ALL developing a relapse or an SMN when in remission had a higher use of general practice and hospital health care services compared with matched references, 1-2 months

before the event. There was a possible bimodal structure for daytime visits to general practice with increased visits also 4-6 months before the event. As health care utilisation may be seen as a proxy for morbidity, this indicates that there could be early warnings. To the best of our knowledge, this is the first study to investigate use of health care before a relapse or an SMN in survivors of childhood ALL in remission, and further research is needed. If an increased use of general practice services up to 6 months before the diagnosis of a relapse or an SMN is confirmed in future research, there may be a window for earlier diagnosis. An increased knowledge of the patient pathway to relapse/SMN diagnosis is important to ensure optimal organisation of surveillance programmes.

Figure legends

Figure 1 Flow diagram of the study population

Children with relapse/SMN and matched references in first remission.
*Matching on age group, sex, risk group and treatment protocol.
†The number in brackets is the number of unique persons – the same child can serve as a control more than once and controls can later become cases.
BCR-ALL or T-ALL, B-cell precursor or T-lineage acute lymphoblastic leukaemia; CPR number, civil personal registration number; HSCT in CR1, haematopoietic stem cell transplantation in first complete remission; SMN, second malignant neoplasm.

Figure 2 General practice health care utilisation

General practice utilisation by months before event for cases* (n=60) compared with references† (n=295). (A) Daytime. (B) Out-of-hours. (C) Diagnostic procedures.
Top panel: Contacts/diagnostic procedure mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.
*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.
†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
ALL, acute lymphoblastic leukaemia; SMN, Second malignant neoplasm.

Figure 3 Hospital health care utilization

Hospital health care utilisation by months before event for cases* (n=60) compared with references† (n=295).
Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.
*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.
†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

Figure 4 Health care utilisation

Health care utilisation by months before event for cases (n=49) compared with references* (n=243). Cases are survivors of childhood ALL developing a relapse as the first event (cases developing an SMN are excluded in this analysis).
Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.
ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

Acknowledgments

The authors wish to thank Kaare Rud Flarup for his assistance with data management.

Footnotes

Contributors: KJ designed the study, analysed and interpreted data and wrote the manuscript; BA interpreted data and edited the manuscript; HS designed the study and edited the manuscript; AF analysed and interpreted data and edited the manuscript; KS interpreted data and edited the manuscript; SR interpreted data and edited the manuscript; MC interpreted data and edited the manuscript; PV designed the study, interpreted data and edited the manuscript. All authors approved the final manuscript.

Funding: This work was supported by a grant from the Danish Cancer Society (Kræftens Bekæmpelse). Grant number R124-A7831-15-S2. The funders had no role in the design and conduct of the study.

Competing interests: Birgitte Klug Albertsen declares the following: sponsor for the investigator initiated NOR-GRASPALL 2016 study. Kjeld Schmiegelow declares the following: Speaker and/or Advisory Board Honoraria from Jazz Pharmaceuticals (2020) and Servier (2020); speaker fee from Amgen (2020) and Medscape (2020); Educational grant from Servier (2020). The remaining authors declare that they have no competing interests.

Patient consent for publication: Not required.

Ethical approval: This study was approved by the Danish Data Protection Agency (ID 277). Medical ethical approval was not required according to Danish law.

Data sharing statement: According to the data agreement with the data provider, we are not allowed to share our data. Data are stored and maintained electronically at Statistics Denmark.

References

1. Toft N, Birgens H, Abrahamsson J, et al. Results of NOPHO ALL2008 treatment for patients aged 1-45 years with acute lymphoblastic leukemia. *Leukemia* 2018;32(3):606-15. doi: 10.1038/leu.2017.265 [published Online First: 2017/08/19]
2. de Fine Licht S, Rugbjerg K, Gudmundsdottir T, et al. Long-term inpatient disease burden in the Adult Life after Childhood Cancer in Scandinavia (ALiCCS) study: A cohort study of 21,297 childhood cancer survivors. *PLoS medicine* 2017;14(5):e1002296. doi: 10.1371/journal.pmed.1002296 [published Online First: 2017/05/10]
3. Font-Gonzalez A, Feijen E, Geskus RB, et al. Risk and associated risk factors of hospitalization for specific health problems over time in childhood cancer survivors: a medical record linkage study. *Cancer medicine* 2017;6(5):1123-34. doi: 10.1002/cam4.1057 [published Online First: 2017/04/06]
4. Sieswerda E, Font-Gonzalez A, Reitsma JB, et al. High Hospitalization Rates in Survivors of Childhood Cancer: A Longitudinal Follow-Up Study Using Medical Record Linkage. *PLoS One* 2016;11(7):e0159518. doi: 10.1371/journal.pone.0159518 [published Online First: 2016/07/21]
5. Sorensen GV, Winther JF, de Fine Licht S, et al. Long-Term Risk of Hospitalization Among Five-Year Survivors of Childhood Leukemia in the Nordic Countries. *Journal of the National Cancer Institute* 2019;111(9):943-51. doi: 10.1093/jnci/djz016 [published Online First: 2019/02/13]
6. Mulrooney DA, Hyun G, Ness KK, et al. The changing burden of long-term health outcomes in survivors of childhood acute lymphoblastic leukaemia: a retrospective analysis of the St Jude Lifetime Cohort Study. *The Lancet Haematology* 2019;6(6):e306-e16. doi: 10.1016/s2352-3026(19)30050-x [published Online First: 2019/05/13]
7. Bhakta N, Liu Q, Ness KK, et al. The cumulative burden of surviving childhood cancer: an initial report from the St Jude Lifetime Cohort Study (SJLIFE). *Lancet (London, England)* 2017;390(10112):2569-82. doi: 10.1016/s0140-6736(17)31610-0 [published Online First: 2017/09/12]
8. Heins MJ, Lorenzi MF, Korevaar JC, et al. Non-oncology physician visits after diagnosis of cancer in adolescents and young adults. *Journal of cancer survivorship : research and practice* 2016;10(4):783-8. doi: 10.1007/s11764-016-0523-x [published Online First: 2016/02/15]
9. Heins MJ, Lorenzi MF, Korevaar JC, et al. Non-oncology physician visits after diagnosis of cancer in children. *BMC family practice* 2016;17:60. doi: 10.1186/s12875-016-0462-7 [published Online First: 2016/06/02]
10. Rebholz CE, Reulen RC, Toogood AA, et al. Health care use of long-term survivors of childhood cancer: the British Childhood Cancer Survivor Study. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2011;29(31):4181-8. doi: 10.1200/jco.2011.36.5619 [published Online First: 2011/09/29]
11. Shaw AK, Pogany L, Speechley KN, et al. Use of health care services by survivors of childhood and adolescent cancer in Canada. *Cancer* 2006;106(8):1829-37. doi: 10.1002/cncr.21798 [published Online First: 2006/03/17]

12. Oeffinger KC, Mertens AC, Hudson MM, et al. Health care of young adult survivors of childhood cancer: a report from the Childhood Cancer Survivor Study. *Annals of family medicine* 2004;2(1):61-70. [published Online First: 2004/04/01]
13. Streefkerk N, Heins MJ, Teepen JC, et al. The involvement of primary care physicians in care for childhood cancer survivors. *Pediatric blood & cancer* 2019;66(8):e27774. doi: 10.1002/pbc.27774 [published Online First: 2019/04/30]
14. McBride ML, Lorenzi MF, Page J, et al. Patterns of physician follow-up among young cancer survivors: report of the Childhood, Adolescent, and Young Adult Cancer Survivors (CAYACS) research program. *Canadian family physician Medecin de famille canadien* 2011;57(12):e482-90. [published Online First: 2011/12/16]
15. Ahrensberg JM, Fenger-Gron M, Vedsted P. Use of primary care during the year before childhood cancer diagnosis: a nationwide population-based matched comparative study. *PLoS One* 2013;8(3):e59098. doi: 10.1371/journal.pone.0059098 [published Online First: 2013/04/05]
16. Yang TO, Liu YL, Huang WT, et al. Specific and Non-specific Clinical Presentations in the Year Before the Diagnosis of Childhood Leukaemia. *Pediatric blood & cancer* 2016;63(8):1387-93. doi: 10.1002/pbc.26029 [published Online First: 2016/04/30]
17. Ahrensberg JM, Fenger-Gron M, Vedsted P. Primary Care Use before Cancer Diagnosis in Adolescents and Young Adults - A Nationwide Register Study. *PLoS One* 2016;11(5):e0155933. doi: 10.1371/journal.pone.0155933 [published Online First: 2016/05/21]
18. Benchimol EI, Smeeth L, Guttmann A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS medicine* 2015;12(10):e1001885. doi: 10.1371/journal.pmed.1001885 [published Online First: 2015/10/07]
19. Jensen KS, Oskarsson T, Lähteenmäki PM, et al. Detection mode of childhood acute lymphoblastic leukaemia relapse and its effect on survival: a Nordic population-based cohort study. *Br J Haematol* 2021 doi: 10.1111/bjh.17555 [published Online First: 2021/05/28]
20. Schmiegelow K, Forestier E, Hellebostad M, et al. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia* 2010;24(2):345-54. doi: 10.1038/leu.2009.251 [published Online First: 2009/12/17]
21. Schmiegelow K, Levinsen MF, Attarbaschi A, et al. Second malignant neoplasms after treatment of childhood acute lymphoblastic leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* 2013;31(19):2469-76. doi: 10.1200/jco.2012.47.0500 [published Online First: 2013/05/22]
22. Gjerstorff ML. The Danish Cancer Registry. *Scandinavian journal of public health* 2011;39(7 Suppl):42-5. doi: 10.1177/1403494810393562 [published Online First: 2011/08/04]
23. Andersen JS, Olivarius Nde F, Krasnik A. The Danish National Health Service Register. *Scandinavian journal of public health* 2011;39(7 Suppl):34-7. doi: 10.1177/1403494810394718 [published Online First: 2011/08/04]
24. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scandinavian journal of public health* 2011;39(7 Suppl):30-3. doi: 10.1177/1403494811401482 [published Online First: 2011/08/04]

25. Schmidt M, Pedersen L, Sorensen HT. The Danish Civil Registration System as a tool in epidemiology. *European journal of epidemiology* 2014;29(8):541-9. doi: 10.1007/s10654-014-9930-3 [published Online First: 2014/06/27]
26. Erlangsen A, Fedyszyn I. Danish nationwide registers for public health and health-related research. *Scandinavian journal of public health* 2015;43(4):333-9. doi: 10.1177/1403494815575193 [published Online First: 2015/03/12]
27. Rasmussen LA, Jensen H, Virgilsen LF, et al. Healthcare utilisation in general practice and hospitals in the year preceding a diagnosis of cancer recurrence or second primary cancer: a population-based register study. *BMC health services research* 2019;19(1):941. doi: 10.1186/s12913-019-4757-y [published Online First: 2019/12/07]

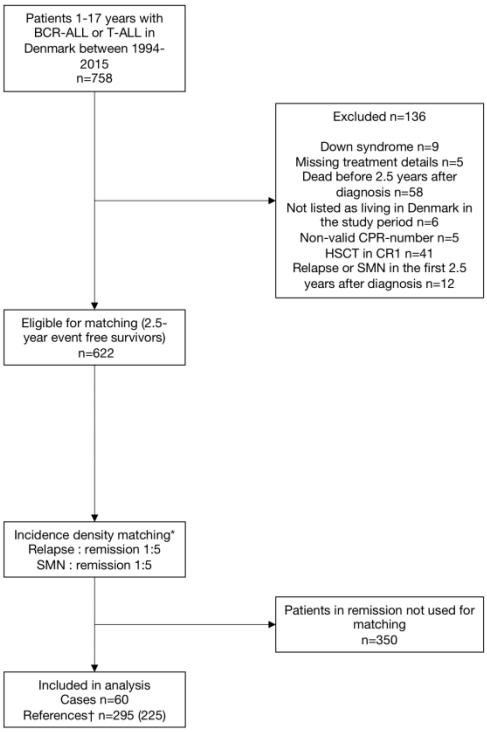
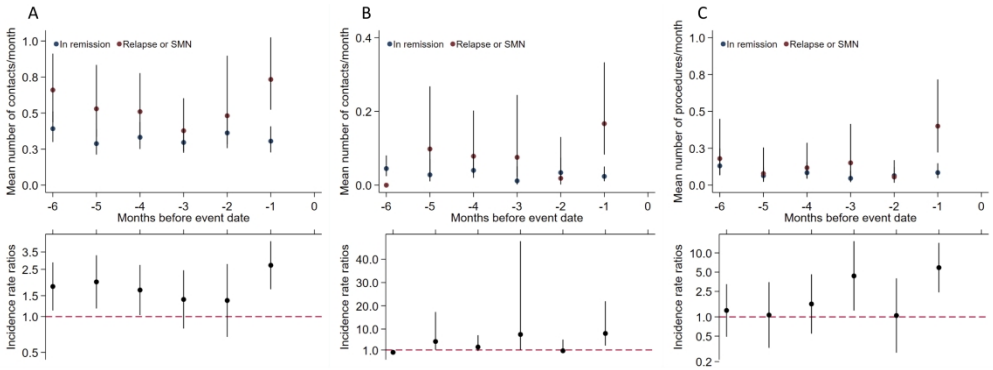


Figure 1 Flow diagram of the study population
Children with relapse/SMN and matched references in first remission.
*Matching on age group, sex, risk group and treatment protocol.
†The number in brackets is the number of unique persons – the same child can serve as a control more than once and controls can later become cases.
BCR-ALL or T-ALL, B-cell precursor or T-lineage acute lymphoblastic leukaemia; CPR number, civil personal registration number; HSCT in CR1, haematopoietic stem cell transplantation in first complete remission; SMN, second malignant neoplasm.

209x297mm (300 x 300 DPI)



General practice health care utilisation

General practice utilisation by months before event for cases* (n=60) compared with references† (n=295). (A) Daytime. (B) Out-of-hours. (C) Diagnostic procedures.

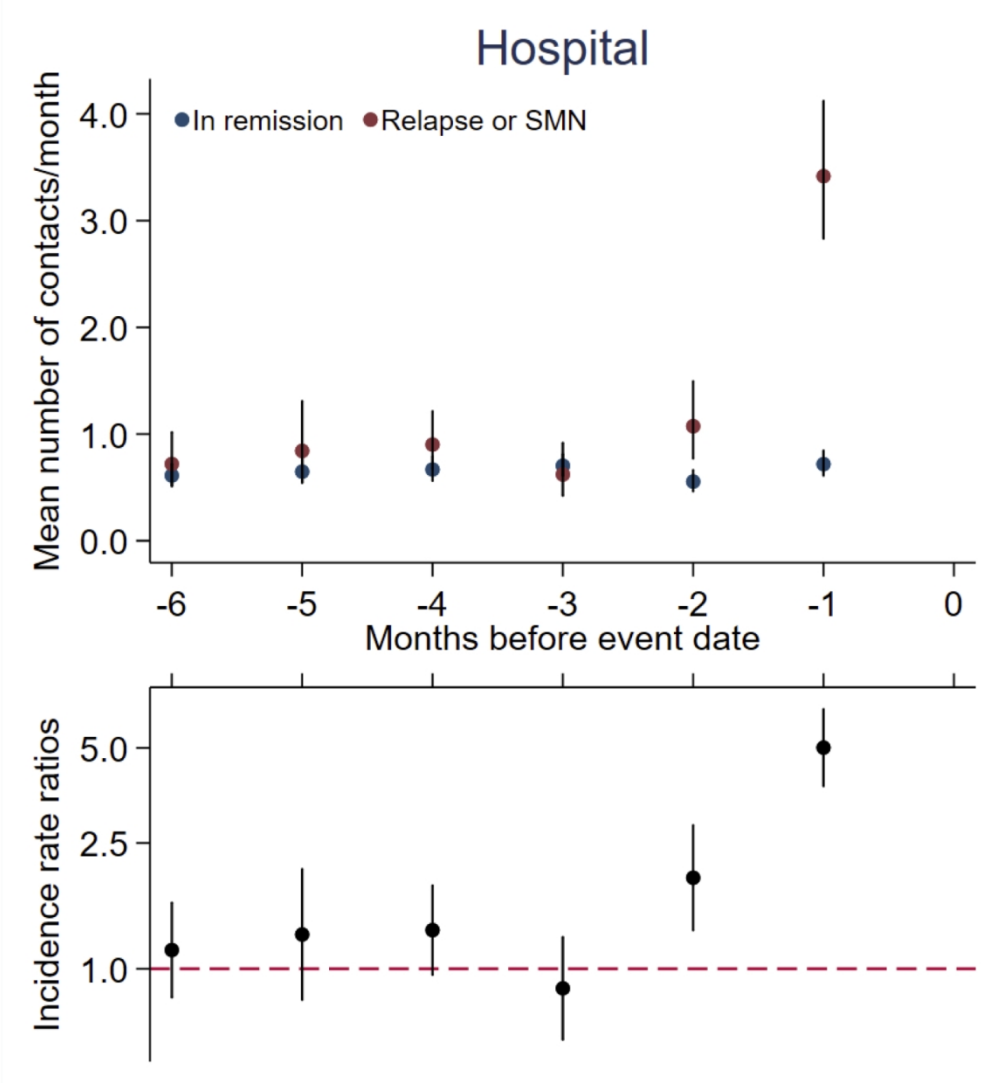
Top panel: Contacts/diagnostic procedure mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.

†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

ALL, acute lymphoblastic leukaemia; SMN, Second malignant neoplasm.

282x105mm (300 x 300 DPI)



Hospital health care utilization

Hospital health care utilisation by months before event for cases* (n=60) compared with references† (n=295).

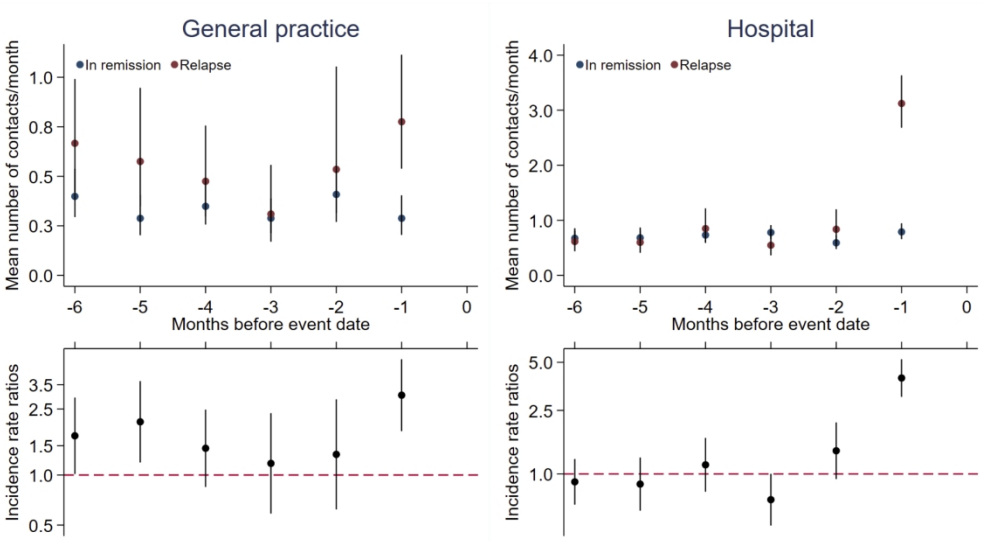
Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*Cases, survivors of childhood ALL developing a relapse or an SMN as the first event.

†References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

133x144mm (300 x 300 DPI)



Health care utilisation

Health care utilisation by months before event for cases (n=49) compared with references* (n=243). Cases are survivors of childhood ALL developing a relapse as the first event (cases developing an SMN are excluded in this analysis).

Top panel: Contacts mean rates per month presented as crude rates. Bottom panel: Incidence rate ratios adjusted for age, sex and time since diagnosis. Vertical lines represent 95% confidence intervals.

*References, survivors of childhood ALL still in first remission matched on age, sex, treatment protocol and risk group.

ALL, acute lymphoblastic leukaemia; SMN, second malignant neoplasm.

265x146mm (300 x 300 DPI)

Table S1

The RECORD statement – checklist of items extended from the STROBE statement to be reported in observational studies using routinely collected health data.

	Item No.	STROBE items	Location in manuscript where items are reported	RECORD items	Location in manuscript where items are reported
Title and abstract					
	1	(a) Indicate the study's design with a commonly used term in the title or the abstract (b) Provide in the abstract an informative and balanced summary of what was done and what was found	Pages 1 and 2	RECORD 1.1: The type of data used should be specified in the title or abstract. When possible, the name of the databases used should be included. RECORD 1.2: If applicable, the geographic region and timeframe within which the study took place should be reported in the title or abstract. RECORD 1.3: If linkage between databases was conducted for the study, this should be clearly stated in the title or abstract.	Page 2 Page 2 Page 2
Introduction					
Background rationale	2	Explain the scientific background and rationale for the investigation being reported	Pages 3-4		
Objectives	3	State specific objectives, including any prespecified hypotheses	Page 4		
Methods					
Study Design	4	Present key elements of study design early in the paper	Page 4		
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	Pages 4-5		
Participants	6	*(a) Cohort study* - Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study* - Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls	Page 5	RECORD 6.1: The methods of study population selection (such as codes or algorithms used to identify subjects) should be listed in detail. If this is not possible, an explanation should be provided. RECORD 6.2: Any validation studies of the codes or algorithms used to select the population should be referenced. If validation was conducted for this study and not	Page 5 N/A

		Cross-sectional study - Give the eligibility criteria, and the sources and methods of selection of participants *(b) Cohort study* - For matched studies, give matching criteria and number of exposed and unexposed *Case-control study* - For matched studies, give matching criteria and the number of controls per case	Page 5	published elsewhere, detailed methods and results should be provided. RECORD 6.3: If the study involved linkage of databases, consider use of a flow diagram or other graphical display to demonstrate the data linkage process, including the number of individuals with linked data at each stage.	Figure 1
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable.	Pages 5-7	RECORD 7.1: A complete list of codes and algorithms used to classify exposures, outcomes, confounders, and effect modifiers should be provided. If these cannot be reported, an explanation should be provided.	Supplementary Table S2
Data sources/ measurement	8	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	Text box		
Bias	9	Describe any efforts to address potential sources of bias	Pages 10-11		
Study size	10	Explain how the study size was arrived at	Page 5		
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	N/A		
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding (b) Describe any methods used to examine subgroups and interactions (c) Explain how missing data were addressed (d) *Cohort study* - If applicable, explain how loss to follow-up was addressed *Case-control study* - If applicable, explain how matching of cases and controls was addressed	Pages 7-8 Page 8 N/A N/A		

		Cross-sectional study - If applicable, describe analytical methods taking account of sampling strategy (e) Describe any sensitivity analyses	Page 8		
Data access and cleaning methods		..		RECORD 12.1: Authors should describe the extent to which the investigators had access to the database population used to create the study population. RECORD 12.2: Authors should provide information on the data cleaning methods used in the study.	Page 5 N/A
Linkage		..		RECORD 12.3: State whether the study included person-level, institutional-level, or other data linkage across two or more databases. The methods of linkage and methods of linkage quality evaluation should be provided.	Page 5
Results					
Participants	13	(a) Report the numbers of individuals at each stage of the study (*e.g.*, numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed) (b) Give reasons for non-participation at each stage. (c) Consider use of a flow diagram	Figure 1 Figure 1 Figure 1	RECORD 13.1: Describe in detail the selection of the persons included in the study (*i.e.,* study population selection) including filtering based on data quality, data availability and linkage. The selection of included persons can be described in the text and/or by means of the study flow diagram.	Page 5 and Figure 1
Descriptive data	14	(a) Give characteristics of study participants (*e.g.*, demographic, clinical, social) and information on exposures and potential confounders (b) Indicate the number of participants with missing data for each variable of interest (c) *Cohort study* - summarise follow-up time (*e.g.*, average and total amount)	Table 1 Table 1 Table 1		
Outcome data	15	*Cohort study* - Report numbers of outcome events or summary measures over time *Case-control study* - Report numbers in each exposure category, or summary measures of exposure	Page 8-10		

		Cross-sectional study - Report numbers of outcome events or summary measures			
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (e.g., 95% confidence interval). Make clear which confounders were adjusted for and why they were included (b) Report category boundaries when continuous variables were categorized (c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	Pages 8-10 and Figures 2-3 Table 1 N/A		
Other analyses	17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	Page 10, Figure 4		
Discussion					
Key results	18	Summarise key results with reference to study objectives	Page 10		
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	Pages 10-11	RECORD 19.1: Discuss the implications of using data that were not created or collected to answer the specific research question(s). Include discussion of misclassification bias, unmeasured confounding, missing data, and changing eligibility over time, as they pertain to the study being reported.	Pages 10-11
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	Page 12		
Generalisability	21	Discuss the generalisability (external validity) of the study results	Page 11		
Other Information					
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	Page 14		
Accessibility of protocol, raw data, and programming code		..		RECORD 22.1: Authors should provide information on how to access any supplemental information such as the study protocol, raw data, or programming code.	Page 15

*Reference: Benchimol EI, Smeeth L, Guttmann A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM, the RECORD Working Committee. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) Statement. *PLoS Medicine* 2015; in press.

Table S2.
Information about contacts to general practice obtained from the Danish National Health Insurance Service Register

Contact type	Speciality code	Time code	Services
Daytime contacts	80	1	
			0101 0102 0105 0201 0411 0421 0431 0441 0451 0461 0491
Out-of-hours contacts	80	8 9	
			0101 0102 0471 0501
Diagnostic procedures	80	1 8 9	
Blood test			2101 2601 4309 4311 4312 4544 4611 7108 7110 7115 7120 7125 7126 7136 7150 7159 7168

			7177 7184 7186 7256 7263 7301 7302 7305 7309 7330 7403
Urine test			2132 4308 7101 7102 7122 7189
Strep throat test			4310 7109
Pulmonary functions test			4543 7113 7121 7183
Electrocardiogram			4313 7155 7156