Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Cancer Epidemiology 51 (2017) 118–124

Contents lists available at ScienceDirect

Cancer Epidemiology

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

Original Article

Seasonal trends of diagnosis of childhood malignant diseases and viral prevalence in South Korea

Kyu Seok Shim[a], Min Hyung Kim[a], Choong Nam Shim[b], Minkyu Han[c], In Seok Lim[a,d], Soo Ahn Chae[a,d], Sin Weon Yun[a,d], Na Mi Lee[a], Dae Yong Yi[a,d], Hyery Kim[d,e,*]

[a] Department of Pediatrics, Chung-Ang University Hospital, Seoul, Republic of Korea
[b] Department of Internal Medicine, Nasaret International Hospital, Incheon, Republic of Korea
[c] Department of Clinical Epidemiology and Biostatistics, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea
[d] College of Medicine, Chung-Ang University, Seoul, Republic of Korea
[e] Department of Pediatrics, Asan Medical Center, Children's Hospital, University of Ulsan College of Medicine, Seoul, Republic of Korea

ARTICLE INFO

Keywords:
Childhood neoplasm
Seasonality
Time trend
Viral infections

ABSTRACT

Background: Several studies have reported a seasonal trend in the diagnosis of childhood cancer suggesting seasonal factors such as infection. The present study aimed to analyze the diagnosis pattern of childhood malignant diseases using public health data, and to compare this pattern with seasonal viral infection trends.

Method: Using the open data source of the Health Insurance Review and Assessment Service, we extracted data regarding all patients under 21 years of age and who had any cancer, aplastic anemia or myelodysplastic syndrome between September 2009 and December 2013. The positive detection rates of 11 viruses was collected from the surveillance data of Korea Centers for Disease Control and Prevention, and seasonality analysis were conducted with both data.

Results: In total, 9085 patients were diagnosed with malignant disease during the study period; there were about 175 new cases per month on average. Monthly stacked time series by year showed an apparent seasonal variation with the highest monthly average in January as 236, and the lowest in September as 120. In winter, significantly more patients were diagnosed with acute lymphoblastic leukemia, acute myeloid leukemia, neuroblastoma, and Hodgkin's lymphoma than in other seasons. There was a temporal correlation with the diagnostic trends of several diseases and the prevalence of recent human parainfluenza virus.

Conclusion: This study tentatively suggests that the diagnosis of childhood malignancy follows a seasonal trend in Korea, and has a possible correlation with viral prevalence in several diseases. Further long-term analysis of epidemiological data is needed to explore possible causality.

1. Introduction

Malignant tumors of childhood are the most common disease-related cause of death among patients younger than 20 years old in Korea [1]. Annually, 16.6 per 100,000 children are diagnosed with malignancies; this accounts for about 1% of all cancer patients.

Despite decades of research, the etiology of childhood cancer remains largely unknown. However, one hypothesis has suggested that pediatric cancer is related to infection. In particular, precursor B cell acute lymphoblastic leukemia (ALL), several subtypes of lymphoma, and nasopharyngeal cancer are reportedly associated with viral infection [2–5].

Several studies have investigated temporal variations in the diagnosis of childhood cancer. Summer peaks of ALL diagnosis have been observed amongst both children and adults in United Kingdom [6]. In the United States, summer peaks of ALL, hepatoblastoma, and rhabdomyosarcoma diagnosis have been observed amongst children, and a winter peak of central nervous system (CNS) tumor diagnosis has been observed amongst children in Southern USA [7]. In Denmark, seasonal variation in diagnosis—with a peak in March—was found amongst children and adolescents with Hodgkin's lymphoma (HL) [8]. According to the recent study of Northern England, there was statistically significant sinusoidal variation in girls for all lymphomas (peak in March) and HL (peak in January), and in boys for osteosarcoma (peak in October) [9].

Such variations in the rate of diagnosis may arise from seasonal

* Corresponding author at: Division of Pediatric Hematology/Oncology, Department of Pediatrics, Asan Medical Center Children's Hospital, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Republic of Korea.
E-mail address: taban@hanmail.net (H. Kim).

http://dx.doi.org/10.1016/j.canep.2017.11.003
Received 11 July 2017; Received in revised form 17 October 2017; Accepted 1 November 2017
Available online 08 November 2017
1877-7821/ © 2017 Elsevier Ltd. All rights reserved.

differences in environmental risk factors, such as infections. That is, childhood cancer might develop in a two-step process: exposure to infection may be the "second hit" that promotes overt malignancy in children with underlying genetic vulnerability [4,5]. If a triggering infection of this kind exhibited a seasonal pattern, as influenza or gastro-enteritis does, some seasonal variation would be expected in the date of childhood cancer diagnosis.

In this regard, we searched public health data for evidence of (1) any seasonal variation in the diagnosis of childhood malignant diseases, and (2) any correlation between these trends and the prevalence of a specific virus. No previous studies have reported seasonal trends in pediatric malignant disease diagnosis in Korea. This study aimed to analyze the diagnosis pattern of pediatric malignant disease, and to compare the pattern with seasonal trends in viral infection.

2. Materials and methods

2.1. Data of childhood malignant disease diagnosis

We extracted data regarding all pediatric malignant disease from the open data source of the Health Insurance Review and Assessment Service (HIRA). South Korea has a universal health coverage system, and the National Health Insurance covers approximately 98% of the South Korean population. Claims data are collected by the HIRA when healthcare service providers in South Korea seek reimbursements for services that are covered by the National Health Insurance Corporation. The claims data of the HIRA are open access to approved researchers who submit a profile of their research. They contain the information of 46 million patients per year, which accounts for 90% of the total population of South Korea; the claims data include information regarding patients' diagnoses, hospitalization, and prescription drugs; such information constitutes a valuable resource for healthcare research.

In Korea, the disease code registered in the HIRA database is entered directly by the patient's primary care physician. Since 2005, the "National cancer registration and reimbursement program" has been implemented in Korea, which reimburse for 90% of the medical expenses of patients with malignant diseases. Therefore, all the patients with malignant diseases should be registered in the database after a definite diagnosis, so that the diagnosis of the patient registered in the HIRA database would be relatively correct. Using this open claims data, we used selected diagnostic codes to extract the data of patients with pediatric malignant diseases. The source population comprised all patients under 21 years of age who had been diagnosed with cancer, aplastic anemia (AA) or myelodysplastic syndrome (MDS), as defined by the International Classification of Diseases (ICD, 10th revision), between September 2009 and December 2013 (Table 1). Data selection and mining were conducted using SAS (version 9.2) and R (3.2.2; http://www.r-project.org/).

2.2. The surveillance data of virus

The Korea Centers for Disease Control and Prevention collect data regarding pathogens that cause acute respiratory infection and enteric diseases. During each year of the study period, more than 1000 respiratory specimens were collected from about 100 participating hospitals across the country, and the causative pathogens were identified using standardized diagnostic procedures in a central laboratory. In addition, the prevalence of enteric pathogens was under weekly surveillance, with cohorts comprising 17 local environmental and health institutes, as well as 105 participating hospitals. Each week, the percentage of positive cases among all specimens was reported for each pathogen. We calculated the average monthly positive detection rates (PDRs) of 7 respiratory viruses (adenovirus, parainfluenza virus, respiratory syncytial virus, influenza virus, corona virus, rhinovirus, and bocavirus) and 4 acute diarrhea viruses (adenovirus, rotavirus, norovirus, and astrovirus). The PDR data was collected from 2009 to 2013.

Table 1
International Classification of Diseases 10 for childhood cancer, aplastic anemia and myelodysplastic syndrome.

Disease classification	ICD 10 codes
Aplastic anemia	D61.0–D61.9
Acute lymphoblastic leukemia	C91.0, C91.3, C91.6, C91.7, C91.9
Acute myeloid leukemia	C92.0, C92.3–C93.0, C94.2
Bone tumor	C40.0, C41.0–C41.4, C41.8, C41.9
Chronic myeloid leukemia	C92.1–C92.2
Chronic myelomonocytic leukemia	C93.1
Central nervous system tumor	C70.0, C70.1, C70.9, C71.0–C71.9, C72.0–C72.9
Germ cell tumor	C56.0, C56.1, C56.9, C57.0, C57.3, C57.4, C57.7–C58.0, C60.0 ~
Hepatoblastoma	C22.2
Hepatocellular carcinoma	C22.0
Hodgkin's lymphoma	C81.0, C81.1, C81.2, C81.3, C81.4, C81.7, C91.9
Juvenile myelomonocytic leukemia	C93.3
Myelodysplastic syndrome	C94.6, D46.0–D46.9
Myelofibrosis/Myeloproliferative neoplasm	C94.5
Neuroblastoma	C74
Non-Hodgkin lymphoma	C82, C83, C84, C85, C86, C88.4
Nasopharyngeal carcinoma	C11
Retinoblastoma	C69.1, C69.2, C69.3, C69.4, C69.6, C69.9
Renal tumor	C64
Sarcoma	C49

ICD, International Classification of Diseases.

2.3. Statistical analysis

Because seasonality is a general component of monthly time series, we used the autoregressive integrated moving average (ARIMA) modeling approach to build a model of seasonal variation in the diagnosis of childhood cancer. The seasonal ARIMA model assumes that the current observation is related to past observations through time. The general multiplicative form of the ARIMA model is written as follows: ARIMA (p, d, q) × (P, D, Q). This model has non-seasonal, autoregressive, and moving average parameters of orders p and q, seasonal, autoregressive, and moving average parameters of order P and Q, and two non-seasonal and seasonal differencing parameters of orders d and D, respectively. To determine the general form of the model to be fit, residual ACF (Autocorrelation Function) was examined. Considering the ACF graphs, different ARIMA models were identified for model selection (Fig. 1, Supplemental Fig. 1). The model of the minimum Akaike's information criterion (AIC) was selectd as the best-fit model. The Granger approach [10] was used to investigate how many of the current values in the time series y could be explained by other values in the time series y. The commonly used EViews software was originally developed for use in economics, but it can be useful in other statistical applications; EViews version 5.0 (http://www.eviews.com/) was used to analyze the data. The null hypotheses were that series x does not Granger-cause series y in the first regression, and that y does not Granger-cause x in the second regression. Seasonal differences in newly diagnosed childhood cancer were analyzed using the Chi-Square test, and P-values < 0.05 were considered statistically significant.

3. Results

3.1. Trend analysis of diagnosis of childhood malignant disease and time series modeling

Between September 2009 and December 2013, 9085 patients had pediatric cancer, AA and MDS; there were 175 new cases per month on average. The total numbers of newly diagnosed patients were 933 in 2009 (September–December), 2322 in 2010, 2091 in 2011, 1933 in

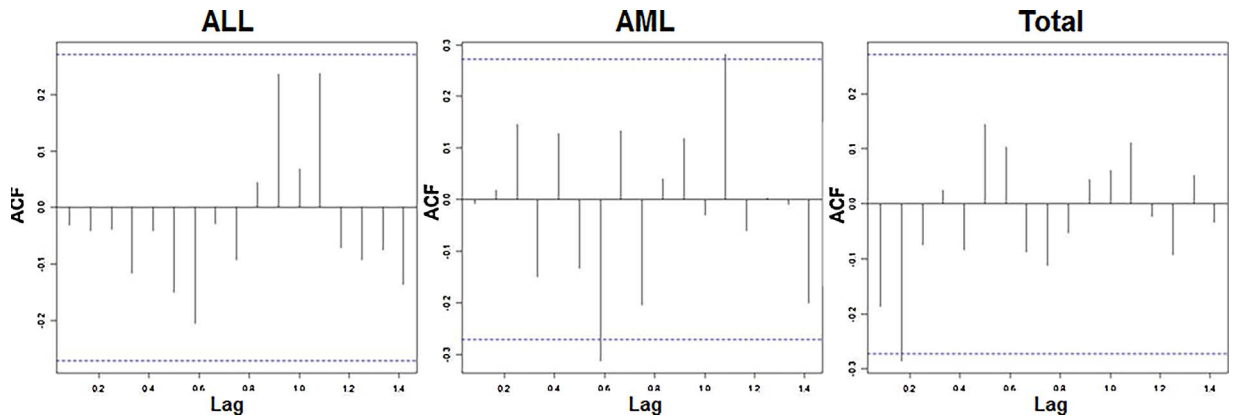


Fig. 1. Redidual ACF correlogram and 95% confidence limits for newly diagnosed acute lymphoblastic leukemia, acute myeloid leukemia and total childhood malignant disease.

2012 and 1806 in 2013, respectively.

Fig. 2 illustrates the variation in the total number of newly diagnosed childhood cancers including AA and MDS between September 2009 and December 2013. More specifically, Fig. 2A, which illustrates the entire study period longitudinally, shows a remarkable variability in the diagnosis trend (Fig. 2A). In Fig. 2B, these longitudinal data were stacked in terms of month, and a seasonal variation became apparent.

Table 2 shows the average numbers of newly diagnosed pediatric patients with childhood malignant diseases per month. The average numbers of patients with newly diagnosed pediatric cancer were 233, 194, 174, 161, and 151 in 2009 (September–December), 2010, 2011, 2012, and 2013, respectively. An average of 175 patients were diagnosed each month; the highest monthly average was 236 patients (in January), and the lowest was 120 patients (in September). In 2010 and 2011, October had the fewest diagnoses, whereas in 2012 and 2013 September had the fewest. Conversely, in 2010 and 2011, December had the most diagnoses; in 2012, November had the most, and in 2013, January had the most.

To carry out time series modeling, we fitted an ARIMA model to the residuals of an order 3, polynomial, time-trend model with a significant trend. Table 3 shows the parameters of the ARIMA models in newly diagnosed childhood malignant disease patients, and the AIC parameters of the model.

3.2. Trend analysis by disease category

Table 1 shows the classification of disease groups, and Supplemental Fig. 2 shows a monthly time series of each disease group. Seasonal trends were apparent in acute myeloid leukemia (AML), Non-Hodgkin's lymphoma (NHL), and especially ALL and germ cell tumors (GCTs). In the case of ALL, the number of diagnoses was lowest in September; this increased towards a peak in December and January.

To ascertain seasonal differences in newly diagnosed childhood cancer by disease category, we carried out Chi-Square test, which showed that AML, HL, and neuroblastoma (NBL) had statistically significant seasonal differences (Table 4). That is, significantly more patients were diagnosed with AML or NBL in winter, and HL was diagnosed more in winter than in autumn.

3.3. Positive detection rates of virus

Supplemental Table 1 outlines the PDRs of viruses. The PDRs of most viruses showed apparent seasonal variation. Specifically, the PDR of the parainfluenza virus was highest from May to August. That of adenovirus was highest in September, and those of respiratory syncytial virus and norovirus were highest from November to December. Influenza and norovirus had their highest PDR from January to April, while the PDR of coronavirus was highest from December to February. Finally, the bocavirus PDR was highest from April to June. To assess approximate viral trends, the monthly PDRs of all viruses were summed up; this total PDR was lowest in September and highest in December.

3.4. Seasonal correlation between childhood malignancy diagnosis and virus prevalence

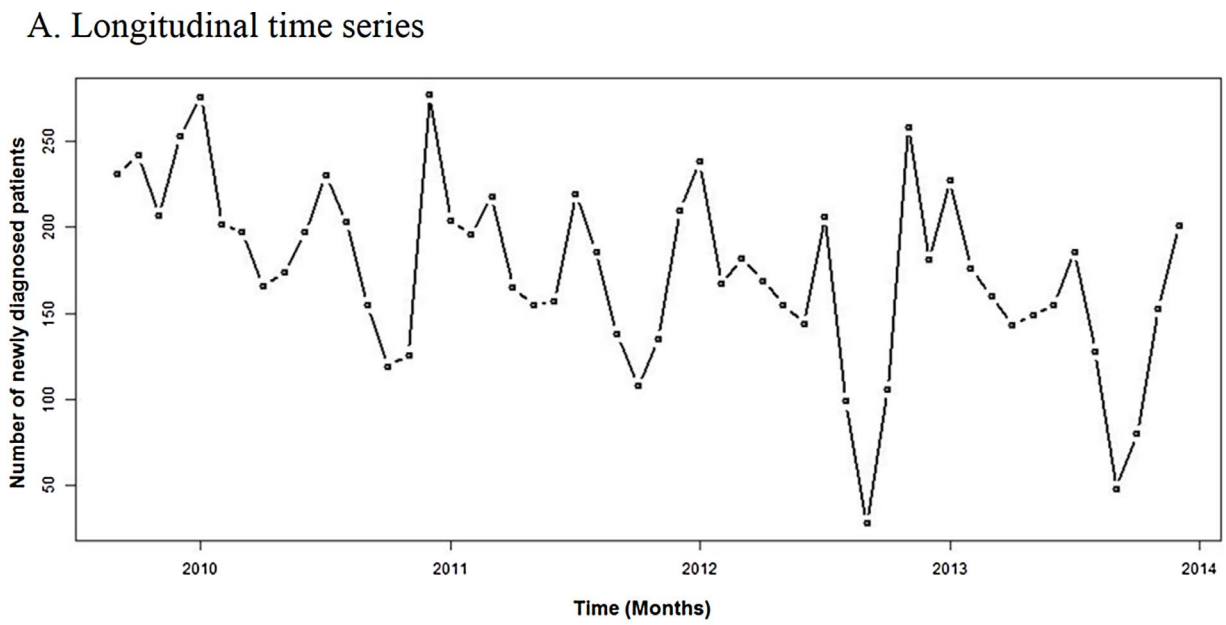
If any virus prevalent time affected cancer diagnosis, the prevalence of that virus might increase before a peak in cancer diagnosis. Thus, a Granger causality test was conducted between the virus PDR data and the cancer diagnostic data from 1 to 2 months later. The results of this Granger causality test are shown in Table 5. Among the seven respiratory viruses analyzed, the prevalence time of human parainfluenza virus was related to the diagnostic data of 9 disease categories after 1 month, as well as to the summed data (ALL, AML, bone tumors, CNS tumors, GCT, NHL, nasopharyngeal cancer, retinoblastoma, sarcoma, and all childhood cancers). After 2 months, it was related to the data for ALL, AML, bone tumors, CNS tumors, NHL, retinoblastoma, and all childhood cancers. In addition, respiratory syncytial virus was related to the HL and NBL data after 1 month and to the aplastic anemia and AML data after 2 months. Among the enteric viruses, astrovirus was related to newly diagnosed sarcoma after 1–2 months.

4. Discussion

This study was carried out to investigate whether the diagnosis of childhood malignant diseases follows any seasonal trend, and there is any correlation between this trend and the prevalence timing of specific viruses. Seasonal trends were apparent in AML, NHL, and especially in ALL and GCTs. These trends suggest that viruses or allergic causes, which have a seasonal prevalence pattern throughout the year, might affect the occurrence of childhood cancer.

Previous studies into the seasonal patterns of childhood cancer have mainly focused on ALL, and several studies have reported a summer peak in the diagnosis of childhood ALL [6,7,11,12]. Other investigation has found a winter peak in childhood ALL diagnosis [13]. Harris et al. compared the seasonal risk of childhood ALL between cases diagnosed in the northern USA (greater than 40°N latitude) and those diagnosed in the southern USA (less than 40°N latitude) [14]. They found complex, trimodal patterns, with seasonal peaks in April, August, and December at northern latitudes, and in February, July, and October in the southern locations. They suggested that these peaks coincided with seasonal elevations in allergic and infectious processes: the indexes of tree and grass pollen are higher in the spring, and that of ragweed pollen is higher in the summer; conversely, influenza is more prevalent in the winter.

In addition to ALL, a cancer seasonality study involving adolescents in United Kingdom revealed significant evidence of seasonality in the



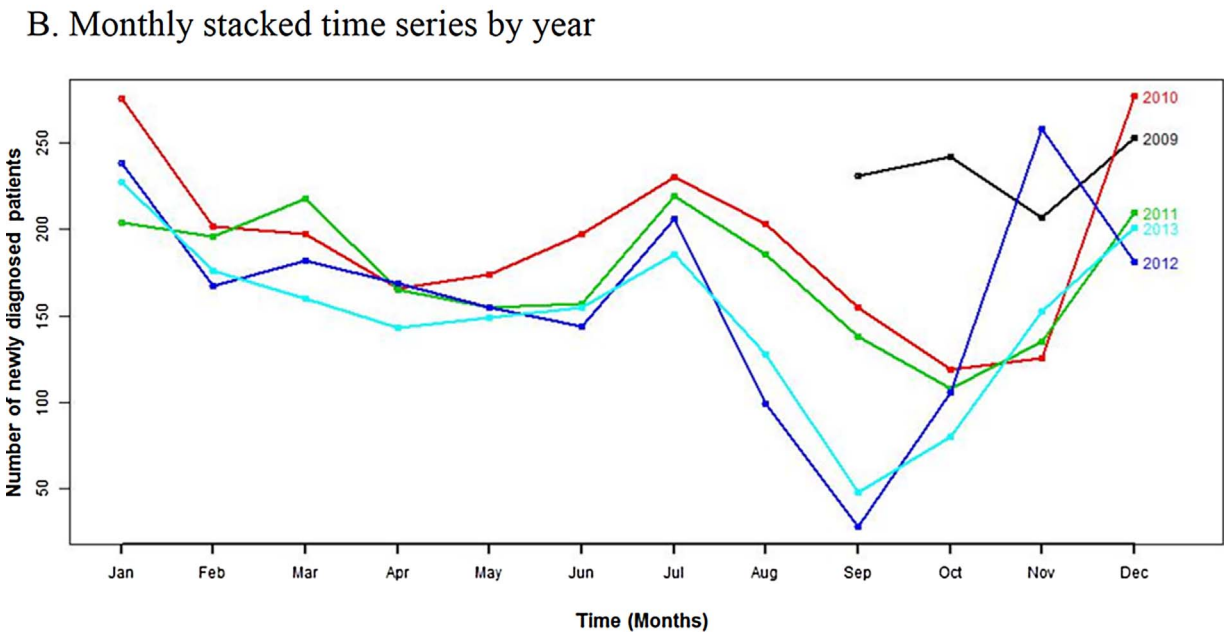


Fig. 2. Monthly time series for numbers of newly diagnosed childhood cancer: (A) Longitudinal time series shows the variation in the total numbers of patients with newly diagnosed childhood malignant disease between September 2009 and December 2013; (B) Monthly stacked time series by year shows an apparent seasonal variation so that the annual lowest numbers were between September and November.

Table 2
Monthly numbers of newly diagnosed patients with childhood malignant diseases in Korea.

	Mon												
Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Ave.
2009									231	242	207	253	233
2010	276	202	197	166	174	197	230	203	155	119	126	277	194
2011	204	196	218	165	155	157	219	186	138	108	135	210	174
2012	238	167	182	169	155	144	206	99	28	106	258	181	161
2013	227	176	160	143	149	155	186	128	48	80	153	201	151
Ave.	236	185	189	161	158	163	210	154	120	131	176	224	175

Mon, month; Ave., average.

Table 3
Parameters of ARIMA models for newly diagnosed childhood malignant disease patients.

Parameters		AA	ALL	AML	BT	CML	CNS	GCT	HBL	HL	MDS	NBL	NHL	NPca	RBL	RT	Sar	Total
p	Nonseasonal AR	0	3	1	0	0	0	3	0	0	0	0	0	1	0	1	0	0
q	Nonseasonal MA	1	0	0	0	0	1	0	0	1	0	0	1	0	1	0	0	0
P	Seasonal AR	0	0	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1
Q	Seasonal MA	1	0	0	1	0	1	0	0	1	0	0	1	0	1	0	0	0
d	Nonseasonal Diff	0	1	0	0	0	1	1	0	0	0	0	0	1	1	1	0	1
D	Seasonal Diff	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
Final model AIC		347.14	367.99	282.83	180.18	202.73	14.03	379.62	341.56	212.84	190.37	216.00	88.43	233.77	−55.90	253.03	334.05	173.42

AA, aplastic anemia; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; AR, autoregressive; BT, bone tumor; CML, chronic myeloid leukemia; CNS, central nervous system tumors; Diff, difference; GCT, germ cell tumor; HBL, hepatoblastoma; HL, Hodgkin's lymphoma; MA, moving average; MDS, myelodysplastic syndrome; NBL, neuroblastoma; NHL, Non-Hodgkin's lymphoma; NPca, nasopharyngeal cancer; RBL, retinoblastoma; RT, renal tumor; Sar, sarcoma; AIC, Akaike's Information Criterion

diagnosis of HL, with a peak in November and December [11]. These findings are more consistent with an infectious exposure than with any other environmental exposures. Variation in the peak seasons would be caused by climate changes, which clearly influence the timing of particular infectious outbreaks [5].

In our study, childhood ALL showed a seasonal trend: diagnosis increased in the winter, and there was decreasing trend subsequently until September, when the lowest number of cases were diagnosed. Childhood ALL is a relatively well-known disease associated with infection, with a peak age of development of between 2 and 5 years. The disease is associated with industrialization and modern societies, and increased prevalence has led to the formulation of two infection-based hypotheses: Kinlen's "population-mixing" hypothesis, and Greaves' "delayed-infection" hypothesis [15,16]. Kinlen's hypothesis predicts that clusters of childhood ALL cases result from the exposure of susceptible individuals to common, but fairly innocuous infections after population-mixing with carriers [15]. The delayed-infection hypothesis of Greaves is based on a minimal "two-hit" model; it suggests that some susceptible individuals have a prenatally acquired, pre-leukemic allele, and that these individuals had little or no exposure to common infections early in life because they live in an affluent hygienic environment [16]. Such insulation from infection predisposes the immune system of these individuals to aberrant or pathological responses after subsequent or delayed exposure to common infections at an age when there is increased lymphoid-cell proliferation.

It might be noticeable that the diagnosis of childhood AML, HL, NHL, and NBL increased during the winter in our study. In particular, there have been limited epidemiologic data on NBL. Recently, an Italian neuroblastoma cohort study reported that stage 4S NBL had a significant peak among July births; however no trend in the date of diagnosis was found [17]. In the present study, we could not classify the patients' stages; however, stage 4S NBL accounts for less than 5% of all NBL cases, so our seasonal NBL pattern is surely different from that reported in the Italian study [18]. With regards to AML, while ALL and AML are biologically different diseases with distinct etiologies [19], it is possible that different infectious agents play a role in the development of both. Our preliminary evidence for seasonal peaks in several diseases warrants further investigation. We also found significant July and winter peaks in the diagnosis of GCTs. In Western countries, No GCT time series have been published; however, our results suggest that seasonal factors could influence the development of GCT.

Infection is known to affect cancer development, but only a few viruses have a confirmed causal relationship. Approximately 18% of the global cancer burden has been attributed to infectious agents, with estimates ranging from 7% in developed countries to about 22% in developing countries [20]. Chronic infections by the hepatitis B and C viruses, human papilloma virus, and Helicobacter pylori are reportedly responsible for approximately 15% of all human cancers, especially in adults. Malignancies known to be associated with infection have been considered as resulting from the prolonged latency that occurs in chronic infections. In general, pathogenic infections could be necessary, but not sufficient for the initiation and progression of many cancers. That is, cancer initiation may require additional co-factors, including secondary infections [5]. Therefore, in patients with chronic infection by one agent, a secondary co-infection with another agent may be an important co-factor in cancer initiation and progression. Co-infections are relatively common in areas with a high prevalence of infectious agents, especially in developing countries. These co-infections can

Table 4
Seasonal difference analysis in newly diagnosed childhood malignancy by disease category.

Season (months)	Spring (3, 4, 5)	Summer (6, 7, 8)	Autumn (9, 10, 11)	Winter (12, 1, 2)	*P*
AA	16.58 ± 4.25	19.00 ± 7.11	12.87 ± 8.31	19.69 ± 6.06	0.759
ALL	22.33 ± 5.37	20.42 ± 6.72	21.87 ± 11.72	28.54 ± 7.83	0.071
AML	9.58 ± 2.50	8.50 ± 3.26	9.67 ± 4.22	13.62 ± 2.99	0.004
BT	14.83 ± 3.66	14.83 ± 4.53	13.40 ± 8.10	18.38 ± 4.68	0.221
CML	2.00 [0.50; 2.50]	3.00 [2.00; 3.50]	2.00 [1.00; 2.00]	3.00 [2.00; 3.00]	0.108
CNS	35.42 ± 5.76	34.75 ± 10.06	29.73 ± 16.39	42.92 ± 13.76	0.305
GCT	17.08 ± 3.00	20.25 ± 6.31	11.67 ± 6.63	22.92 ± 6.16	0.359
HBL	2.50 [1.50; 4.00]	2.50 [1.00; 3.50]	3.00 [1.00; 4.00]	1.00 [1.00; 3.00]	0.77
HL	3.00 [1.50; 5.00]	3.50 [2.00; 5.00]	2.00 [1.00; 2.50]	4.00 [3.00; 5.00]	0.013
MDS	3.50 [1.50; 5.00]	4.00 [3.00; 4.00]	2.00 [1.00; 3.00]	3.00 [2.00; 5.00]	0.321
NBL	4.00 [2.50; 5.50]	4.50 [3.00; 6.50]	5.00 [3.50; 5.00]	7.00 [6.00; 8.00]	0.012
NHL	19.83 ± 4.78	21.25 ± 5.40	13.07 ± 6.97	22.54 ± 3.93	0.945
NPca	1.00 [1.00; 2.00]	1.00 [0.00; 1.50]	1.00 [0.00; 1.00]	2.00 [1.00; 2.00]	0.198
RBL	4.67 ± 2.06	4.58 ± 2.15	4.67 ± 3.83	7.00 ± 3.27	0.069
RT	4.58 ± 2.43	4.42 ± 2.78	4.07 ± 2.63	6.69 ± 3.66	0.12
Sar	6.58 ± 2.39	7.50 ± 3.97	6.20 ± 4.02	9.31 ± 3.61	0.149
SUM	169.42 ± 21.22	175.83 ± 39.42	142.27 ± 68.07	216.00 ± 35.97	0.119

AA, aplastic anemia; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BT, bone tumor; CML, chronic myeloid leukemia; CNS, central nervous system tumors; GCT, germ cell tumor; HBL, hepatoblastoma; HL, Hodgkin's lymphoma; MDS, myelodysplastic syndrome; NBL, neuroblastoma; NHL, Non-Hodgkin's lymphoma; NPca, nasopharyngeal cancer; RBL, retinoblastoma; RT, renal tumor; Sar, sarcoma; SUM, summation of all cases.

Table 5
Values of the Granger causality test between the time series of childhood malignancy diagnosis and the time points of positive detection rates of virus, with < 0.05 indicating significance (written in bold).

A. Virus data & Diagnostic data after 1 month

	Ds.																
Virus	AA	ALL	AML	BT	CML	CNS	GCT	HBL	HL	MDS	NBL	NHL	NPca	RBL	RT	Sar	SUM
HAdV	0.735	0.203	0.791	0.052	0.847	0.423	0.251	0.201	0.125	0.912	0.602	**0.029**	**0.020**	0.316	0.112	0.907	0.243
HPIV	0.066	**0.001**	**0.001**	**0.002**	0.548	**0.017**	**0.041**	0.246	0.078	0.065	0.081	**0.021**	**0.012**	**0.014**	0.061	**0.012**	**0.001**
HRSV	0.570	0.083	**0.011**	0.327	0.486	0.286	0.355	0.706	**0.025**	0.812	**0.017**	0.102	0.208	0.528	0.928	0.515	0.070
IFV	0.244	0.617	0.449	0.785	0.440	0.546	0.810	0.696	0.439	0.745	0.494	0.585	0.649	0.532	0.651	0.672	0.633
HCoV	0.095	0.049	0.267	0.124	0.858	0.541	0.079	0.875	0.075	0.395	0.373	0.054	**0.041**	0.062	0.629	0.327	0.115
HRV	0.627	0.976	0.962	0.389	0.591	0.896	0.665	0.856	0.449	0.279	0.749	0.916	0.190	0.508	0.062	0.528	0.933
HBoV	0.102	0.073	0.098	0.613	0.366	0.890	0.428	0.905	0.945	0.557	0.480	0.790	0.280	0.280	0.475	0.298	0.809
Rota	0.425	0.832	0.419	0.885	0.947	0.905	0.655	0.572	0.595	0.939	0.136	0.442	0.422	0.869	0.858	0.284	0.931
Norovirus	**0.006**	0.949	0.078	0.184	0.118	0.904	0.244	0.995	0.371	0.101	0.061	0.197	0.183	0.041	0.703	0.446	0.223
EntericAdenovirus	0.996	0.172	0.620	0.970	0.992	0.889	0.492	0.127	0.828	0.501	0.831	0.468	0.762	0.256	0.766	0.521	0.803
Astrovirus	0.070	0.496	0.028	0.141	0.630	0.551	0.241	0.613	0.055	0.398	0.042	0.104	0.071	0.489	0.191	**0.009**	0.129

B. Virus data & Diagnostic data after 2 months

	Ds.																
Virus	AA	ALL	AML	BT	CML	CNS	GCT	HBL	HL	MDS	NBL	NHL	NPca	RBL	RT	Sar	SUM
HAdV	0.486	0.350	0.667	0.114	0.405	0.788	0.434	0.392	0.210	0.801	0.327	0.065	0.064	0.414	0.637	0.758	0.353
HPIV	0.186	**0.002**	**0.003**	**0.006**	0.774	**0.041**	0.118	0.634	0.216	0.135	0.104	**0.045**	0.059	**0.034**	0.268	0.050	**0.002**
HRSV	**0.040**	0.230	**0.029**	0.526	0.830	0.403	0.258	0.807	0.114	0.619	0.073	0.235	0.301	0.787	0.428	0.079	0.357
IFV	0.577	0.847	0.540	0.843	0.502	0.609	0.701	0.679	0.754	0.602	0.571	0.304	0.964	0.643	0.428	0.868	0.752
HCoV	0.334	0.070	0.411	0.261	0.632	0.643	0.185	0.893	0.169	0.335	**0.050**	0.121	0.238	**0.013**	0.480	0.338	0.120
HRV	0.633	0.297	0.810	0.618	0.830	0.467	0.252	0.998	0.558	0.209	0.699	0.213	0.506	0.171	0.592	0.511	0.111
HBoV	0.057	**0.020**	0.095	0.436	0.517	0.184	0.186	0.955	0.102	0.383	0.464	0.194	0.516	0.177	0.448	0.197	0.087
Rota	0.465	0.308	0.462	0.938	0.868	0.905	0.366	0.813	0.752	0.822	0.102	0.229	0.560	0.467	0.834	0.592	0.760
Norovirus	0.081	0.187	0.135	0.322	0.323	0.691	0.768	0.712	0.514	0.264	0.097	0.196	0.267	0.085	0.908	0.628	0.450
Enteric Adenovirus	0.945	0.370	**0.032**	0.329	0.936	0.820	0.868	0.105	0.745	0.571	**0.040**	0.930	0.919	0.527	0.920	0.412	0.836
Astrovirus	0.079	0.727	0.097	0.334	0.318	0.301	0.613	0.661	**0.020**	0.734	0.090	0.253	0.186	0.741	0.379	**0.035**	0.242

AA, aplastic anemia; ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BT, bone tumor; CML, chronic myeloid leukemia; CNS, central nervous system tumors; Ds., disease; GCT, germ cell tumor; HBL, hepatoblastoma; HL, Hodgkin's lymphoma; MDS, myelodysplastic syndrome; NBL, neuroblastoma; NHL, Non-Hodgkin's lymphoma; NPca, nasopharyngeal cancer; RBL, retinoblastoma; RT, renal tumor; Sar, sarcoma; Sum, summation of all cases.
P-values < 0.05 were considered as indicating causality.

cause an imbalance in the host immune system by affecting the persistence of, and susceptibility to, malignant infections.

In this study, we used the surveillance data for viruses from about 100 participating hospitals across the country. The participating hospitals are located in all the 16 provinces throughout the country; not only university hospitals but also primary medical centers are participating. Therefore, this data could reflect the nationwide viral prevalence and infection trends in Korea.

No respiratory or enteric viruses have been postulated as risk factors in cancer yet. However, in the present study, a period of high parainfluenza virus prevalence was temporaly associated with the diagnosis of several diseases. This temporal association does not confirm a causal relationship between viral infection and cancer development, and it has not been reported that parainfluenza virus affects tumorigenesis.

Such correlations might occur because frequent hospital visiting during viral infections could lead to the diagnosis of malignant diseases. However, there were several disease groups without seasonality, and GCT showing an apparent seasonality had a summer peak not corresponding to any viral prevalence trend in this study. Thus, hospital visiting factors could not explain all the data in this study.

In addition, cellular immune response such as T cell dysfunction occurs during parainfluenza virus infection [21], which might affect host immune dysregulation and cancer susceptibility. If infection affects tumorigenesis, the seasonal prevalence pattern of childhood cancer diagnosis would vary by region and country according to differing infection patterns. Further longitudinal multifactorial study is needed to elucidate infection related etiological factors.

The present study was the first to investigate seasonal trends in childhood malignancy diagnosis in Korea. However, unavoidable limitations need to be considered. First, the present study was an ecological analysis. We had no information at the individual level regarding potential confounding factors. Second, we had no information regarding the month of onset of clinical symptoms, which may be a more important indicator because there may be considerable variability in the time between the first appearance of symptoms and diagnosis. Thirdly, there may have been selection bias: although duplicated patients were excluded, it is possible that some of the patients had been diagnosed before the study period. This selection bias is inevitable in cross-sectional cohort studies; however, more long-term study could minimize such bias. The fourth, this type of study cannot evaluate exposure to specific infectious agents, and individual exposures to certain virus could not be confirmed in the open data source. Finally, because the etiology of most cancers is likely multifactorial, not all cases would necessarily have an infectious cause. However, the combinational effects would probably result in an underestimate, rather than an overestimate, of any putative seasonal influence.

In conclusion, this study tentatively suggests that the diagnosis of childhood malignancy follows a seasonal trend in Korea, and has a possible correlation with viral prevalence in several diseases. Further long-term analysis of epidemiological data is needed to explore possible causality.

Authorship contribution statement

H Kim designed the study, collected, reviewed, and analyzed data; KS Shim, MH Kim, and CN Shim wrote the paper; M Han performed

statistical and bioinformatic analysis; IS Lim, SA Chae, SW Yun, and NM Lee contributed data collection; DY Yi organized writing process. All authors made substantial contributions to the concept and design of the study, and gave final approval for the manuscript to be submitted.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of interest

Conflict of interest relevant to this article was not reported.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.canep.2017.11.003.

References

[1] H.J. Park, E.K. Moon, J.Y. Yoon, et al., Incidence and Survival of Childhood Cancer in Korea, Cancer Res. Treat. 48 (3) (2016) 869–882.
[2] S.C. Peh, V.S. Nadarajah, Y.C. Tai, et al., Pattern of Epstein-Barr virus association in childhood non-Hodgkin's lymphoma: experience of university of malaya medical center, Pathol. Int. 54 (3) (2004) 151–157.
[3] G. Niedobitek, Epstein-Barr virus infection in the pathogenesis of nasopharyngeal carcinoma, Mol. Pathol. 53 (5) (2000) 248–254.
[4] C.H. Pui, L.L. Robison, A.T. Look, Acute lymphoblastic leukaemia, Lancet 371 (9617) (2008) 1030–1043.
[5] R.J. McNally, T.O. Eden, An infectious aetiology for childhood acute leukaemia: a review of the evidence, Br. J. Haematol. 127 (3) (2004) 243–263.
[6] P. Badrinath, N.E. Day, D. Stockton, Seasonality in the diagnosis of acute lymphocytic leukaemia, Br. J. Cancer 75 (11) (1997) 1711–1713.
[7] J.A. Ross, R.K. Severson, A.R. Swensen, et al., Seasonal variations in the diagnosis of childhood cancer in the United States, Br. J. Cancer 81 (3) (1999) 549–553.
[8] V. Langagergaard, B. Norgard, L. Mellemkjaer, et al., Seasonal variation in month of birth and diagnosis in children and adolescents with Hodgkin disease and non-Hodgkin lymphoma, J. Pediatr. Hematol. Oncol. 25 (7) (2003) 534–538.
[9] N.O. Basta, P.W. James, A.W. Craft, et al., Season of birth and diagnosis for childhood cancer in Northern England, 1968–2005, Paediatr. Perinat. Epidemiol. 24 (3) (2010) 309–318.
[10] C.W. Granger, Investigating causal relations by econometric models and cross-spectral methods, Econometrica 37 (1969) 424–438.
[11] R.M. Westerbeek, V. Blair, O.B. Eden, et al., Seasonal variations in the onset of childhood leukaemia and lymphoma, Br. J. Cancer 78 (1) (1998) 119–124.
[12] C.D. Higgins, I. dos-Santos-Silva, C.A. Stiller, et al., Season of birth and diagnosis of children with leukaemia: an analysis of over 15 000 UK cases occurring from 1953-95, Br. J. Cancer 84 (3) (2001) 406–412.
[13] M. Karimi, H. Yarmohammadi, Seasonal variations in the onset of childhood leukemia/lymphoma: April 1996 to March 2000, Shiraz, Iran, Hematol. Oncol. 21 (2) (2003) 51–55.
[14] R.E. Harris, F.E. Harrell Jr., K.D. Patil, et al., The seasonal risk of pediatric/juvenile acute lymphocytic leukemia in the United States, J. Chronic Dis. 40 (10) (1987) 915–923.
[15] L. Kinlen, Infections and immune factors in cancer: the role of epidemiology, Oncogene 23 (38) (2004) 6341–6348.
[16] M. Greaves, Infection, immune responses and the aetiology of childhood leukaemia, Nature reviews, Cancer 6 (3) (2006) 193–203.
[17] S. Parodi, V. Fontana, R. Haupt, et al., Seasonal variations of date of diagnosis and birth for neuroblastoma patients in Italy, Cancer Epidemiol. 37 (5) (2013) 575–578.
[18] Y. Al-Tonbary, M. Badr, A. Mansour, et al., Clinico-epidemiology of neuroblastoma in north east Egypt: a 5-year multicenter study, Oncol. Lett. 10 (2) (2015) 1054–1062.
[19] J.A. Ross, S.M. Davies, J.D. Potter, et al., Epidemiology of childhood leukemia, with a focus on infants, Epidemiol. Rev. 16 (2) (1994) 243–272.
[20] V. Vedham, R.L. Divi, V.L. Starks, et al., Multiple infections and cancer: implications in epidemiology, Technol. Cancer Res. Treat. 13 (2) (2014) 177–194.
[21] P.M. Gray, S. Arimilli, E.M. Palmer, et al., Altered function in CD8+ T cells following paramyxovirus infection of the respiratory tract, J. Virol. 79 (6) (2005) 3339–3349.