

# SCIENTIFIC REPORTS



OPEN

## Hodgkin Lymphoma has a seasonal pattern of incidence and mortality that depends on latitude

Sven Borchmann, Horst Müller & Andreas Engert

Seasonal variations in incidence and mortality after a Hodgkin lymphoma (HL) diagnosis have been previously described with partly conflicting results. The goal of this analysis is to provide a comprehensive analysis of these seasonal variations. In total, 41,405 HL cases diagnosed between 1973 and 2012 in the 18 Surveillance, Epidemiology, and End Results registries were included. Cosinor analysis and Cox proportional-hazards models were employed to analyze seasonality of incidence and mortality, respectively. HL shows a sinusoid seasonal incidence pattern ( $p < 0.001$ ). Estimated incidence in March is 15.4% [95%-CI: 10.8-20.0] higher than in September. This sinusoid pattern is more pronounced at higher latitudes ( $p = 0.023$ ). The risk of dying within the first three years after a HL diagnosis in winter is significantly increased compared to a HL diagnosis in summer at higher latitudes (HR = 1.082 [95%-CI: 1.009-1.161],  $p = 0.027$ ). Furthermore, increasing northern latitude increases the additional mortality risk conferred by a diagnosis in winter ( $p_{\text{interaction}} 0.033$ ). The seasonality patterns presented here provide epidemiological evidence that Vitamin D might play a protective role in HL. Further evidence on the direct association between Vitamin D levels and the clinical course of HL needs to be collected to advance the understanding of the role of Vitamin D in HL.

Hodgkin lymphoma (HL) is among the hematological malignancies with the best prognosis. Ongoing efforts to improve treatment efficacy have resulted in 5-year overall survival (OS) rates of above 95% being reported in clinical trials in various disease stages<sup>1-4</sup>. Despite this, pathogenesis of HL is poorly understood and there are potential unknown risk factors that might affect HL incidence and mortality. Examining seasonal patterns of incidence and mortality might be helpful in improving the understanding of both, the pathogenesis of the disease and the influence of seasonally variable risk factors on incidence and mortality.

Seasonal variations in incidence and mortality of HL have been previously described in small datasets<sup>5-10</sup> with partly conflicting patterns and results. Interestingly, all published studies examining these patterns were performed in fairly northern countries such as Norway, Sweden and northern parts of England or Scotland.

The goal of the present study is to use Surveillance, Epidemiology, and End Results (SEER) registry data to perform a comprehensive analysis of seasonality patterns in incidence and mortality of HL. By using a large and geographically diverse dataset, it is possible to analyze geographic variations of seasonality patterns, which might allow better insights into possible causes and mechanisms of this malignancy.

### Methods

All data was obtained from the SEER incidence database using the November 2014 submission<sup>11</sup>. The SEER program provides population based cancer registry data from 18 SEER registries covering approximately 28% of the U.S. population. A detailed description of the SEER registries, methods of data collection and follow-up is available through the SEER website<sup>12</sup>.

To calculate the number of incident cases in each month, the SEER\*Stat software<sup>13</sup> was used to select all HL cases diagnosed between 1973 and 2012 in the 18 SEER registries. This selection resulted in 50,179 actively followed cases. All death certificate and autopsy only cases ( $n = 281$ ) were excluded. Following that, all cases without known month of diagnosis ( $n = 257$ ) were excluded. The SEER registries "Alaska natives" ( $n = 17$ ) and "Rural Georgia" ( $n = 57$ ) were excluded due to low case numbers. Finally, some cases had no information available for age of diagnosis ( $n = 5$ ) and Ann-Arbor stage ( $n = 8,157$ ) and were excluded. This selection process resulted in

German Hodgkin Study Group (GHSG), Department I of Internal Medicine, University Hospital Cologne, Cologne, Germany. Correspondence and requests for materials should be addressed to S.B. (email: [sven.borchmann@uk-koeln.de](mailto:sven.borchmann@uk-koeln.de))

		No. of cases
Total		41405 (100%)
Sex	Male	22708 (54.8%)
	Female	18697 (45.2%)
Histological subtype	Lymphocyte rich	1371 (3.3%)
	Mixed cellularity	6172 (14.9%)
	Lymphocyte depleted	684 (1.7%)
	Nodular sclerosis	24631 (59.5%)
	NLPHL	1619 (3.9%)
	Not otherwise specified	6928 (16.7%)
Age group	0–19	5352 (12.9%)
	20–29	9869 (23.8%)
	30–39	7865 (19.0%)
	40–49	5610 (13.6%)
	50–59	4202 (10.1%)
	60–69	3647 (8.8%)
	≥70	4860 (11.8%)
Ann-Arbor stage	I	8824 (21.3%)
	II	16309 (39.4%)
	III	8537 (20.6%)
	IV	7735 (18.7%)
Year of diagnosis	<1990	4608 (11.1%)
	≥1990 & <2000	9024 (21.8%)
	≥2000 & <2010	21278 (51.4%)
	≥2010	6495 (15.7%)
Latitude	<34.19°N	7932 (19.2%)
	≥34.19°N & <38.05°N	12737 (30.8%)
	≥38.05°N & <41.68°N	10152 (24.5%)
	≥41.68°N	10584 (25.5%)

**Table 1.** Patient characteristics and number of cases for each subgroup in the dataset. NLPHL denotes nodular lymphocyte-predominant Hodgkin lymphoma.

41,405 cases. A CONSORT flowchart visualizing the selection process is provided as Supplementary Figure 1. The characteristics of the cases included in the analysis are described in Table 1 and are representative of the general HL population.

All calculations were performed using the software R<sup>14</sup> and R Commander<sup>15,16</sup>. Survival analysis and cosinor analysis were performed using the R-packages survival<sup>17</sup> and cosinor<sup>18</sup>, respectively. All p-values were calculated as two-sided p-values. P-values < 0.05 were considered to be statistically significant in all statistical tests employed. All p-values < 0.001 are presented as such. All confidence intervals presented are 95%-intervals. With regard to the subgroup analysis of the seasonal incidence pattern, correction for multiple testing was performed according to the Bonferroni-Holm method<sup>19</sup> and adjusted p-values are indicated with  $p_{adj}$  where applicable.

**Seasonality of incidence.** The date of diagnosis was defined as the month, day and year the HL was first diagnosed clinically or microscopically by a recognized medical practitioner<sup>12</sup>. Cases were grouped according to month of diagnosis. The number of cases per month were adjusted for differences in the length of the respective month by dividing each month's case count by the number of days in that month, with February having 28.25 days, and multiplied by the average number of days across all months (30.4375 days). Subsequently, adjusted case counts were standardized, so that the average month had an incidence of 1. This was done by dividing each month's adjusted case count by the average case count across all months. Cosinor analysis was performed to examine seasonality in the incidence data using the standardized monthly incidence as input. A detailed description of cosinor analysis is provided by Nelson *et al.* and Tong *et al.*<sup>20,21</sup>. In order to examine seasonality in various subgroups, the cases were stratified by sex, histological subtype, age group, Ann-Arbor stage, year of diagnosis and latitude quartiles. Age at diagnosis was sorted into groups aged 0–19, 20–29, 30–39, 40–49, 50–59, 60–69 and >70, respectively. For each case, the location of diagnosis was available from the SEER database at county level. Latitude quartiles were defined by sorting all counties by their respective latitude and allocating the counties into quartiles by latitude. Data on latitude for the single counties was provided by the U.S. Census Bureau<sup>22</sup>. The peak incidence month is defined as the month closest to the incidence peak estimated by the cosinor model. Amplitude is defined as the amplitude estimated by the cosinor model.

Cases were subdivided into cases diagnosed in the northern half (latitude ≥38.05°N) and southern half (latitude <38.05°N) of counties. Subsequently, a dichotomous variable representing the location of diagnosis (northern half vs. southern half of counties) was included into the cosinor model and its influence on the amplitude tested.

Month	Case count	Adjusted case count	Normalized incidence
January	3698	3631	1.052
February	3454	3721	1.078
March	3846	3776	1.094
April	3701	3755	1.088
May	3412	3350	0.970
June	3466	3517	1.019
July	3283	3223	0.934
August	3318	3258	0.944
September	3154	3200	0.927
October	3355	3294	0.954
November	3298	3346	0.969
December	3420	3358	0.973

**Table 2.** Raw, adjusted and normalized incidence of Hodgkin lymphoma for each month in the dataset. The adjusted case count is rounded to the nearest integer.

**Seasonality of Mortality.** In order to analyze seasonal differences in mortality, a Cox proportional-hazards model was employed<sup>23</sup>. Seasonality was first tested by grouping the month of diagnosis into a dichotomous variable and, thus, dividing the cases into those diagnosed during the winter (September to February) and summer half year (March to August). This dichotomization was chosen on the basis of the meteorological definition of the seasons<sup>24</sup>. In the present study, the summer half year comprises the meteorological spring and summer and the winter half year comprises the meteorological autumn and winter. All known risk factors for a poor outcome of HL available from the SEER database were included into the model to control for their effect: Age at diagnosis, histological subtype, year of diagnosis, sex and Ann-Arbor stage.

Follow-up was censored at three years to focus on short term mortality, taking into consideration the hypothesis that mortality differences by season of diagnosis are less pronounced with increasing follow-up, as patients live through more seasons in total, thus decreasing the relative exposure to a specific season that might be a risk factor. A model with 5 years of follow up was fitted as well in order to increase robustness of results. Analysis was firstly performed for all cases together and then stratified by the above mentioned subdivision into the northern vs. southern half of counties.

In order to evaluate a possible effect of latitude on the difference in mortality after a diagnosis in winter versus summer, an additional Cox proportional-hazards model was fitted. This model included a multiplicative interaction term between the dichotomous variable indicating the season of diagnosis and the continuously modelled latitude of the county the case was diagnosed in.

In order to test for robustness of results with respect to dividing the year into two half years at specific months, a further Cox proportional-hazards model was fitted including a sinusoid term  $x_1$  that was calculated with the following formula:

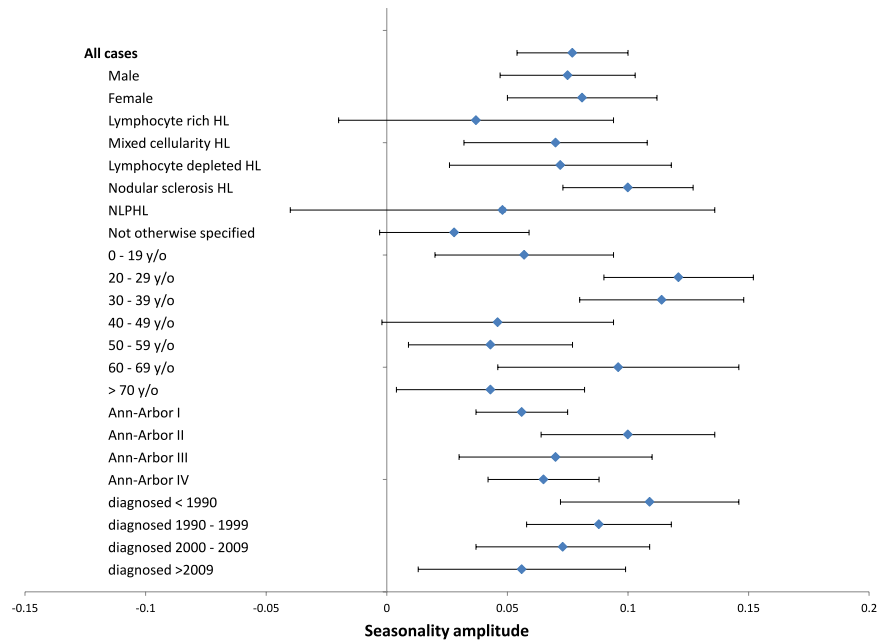
$$x_1 = \cos\left[2\pi\left(\frac{M - M_{\max}}{12}\right)\right] \quad (1)$$

where  $M$  is the month of diagnosis (e.g. January = 1, February = 2...) and  $M_{\max}$  is the month with the maximum risk. For example, for  $M_{\max} = 3$ , this term assigns March the highest possible seasonal risk 1 and September the lowest  $-1$ . For  $M_{\max} = 4$ , the seasonal risk term shifts forward by one month so that April is assigned the highest seasonal risk 1 and October the lowest  $-1$  and so on. Substituting this seasonal risk term for the dichotomized season of diagnosis variable in the Cox proportional-hazards model described above and iteratively determining  $M_{\max}$  so that  $\beta_1$  in the Cox proportional-hazards model is maximized makes it possible to find the highest risk month that describes the true peak and trough of seasonal mortality risk in the data best. A detailed explanation of the concept is given by Efrid *et al.*<sup>25</sup>.

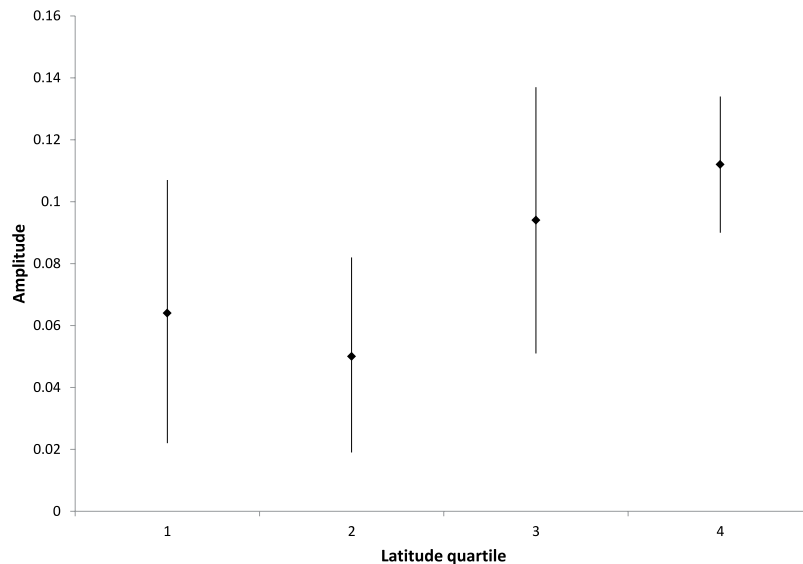
**Data availability.** The data analysed in the current study is available from the SEER registries at <https://seer.cancer.gov/registries/>.

## Results

**Seasonal incidence pattern.** Looking at all incident cases together, a highly significant ( $p < 0.001$ ) seasonal incidence pattern emerged with a peak around March and a trough around September. The estimated amplitude of seasonality was 0.077, indicating that the incidence around September is roughly 15% lower than around March. Table 2 shows the raw, adjusted and normalized incidence for each month for the whole dataset. Figure 1 shows a forest plot with all subgroup analyses. The seasonal pattern was equally present in male and female cases. Stratified by histological subtype, only cases of the mixed cellularity subtype ( $p_{\text{adj}} = 0.027$ ), the nodular sclerosis subtype ( $p_{\text{adj}} = 0.027$ ) and the lymphocyte depleted subtype ( $p_{\text{adj}} = 0.027$ ) showed a significant seasonal incidence pattern when analyzed separately, while the lymphocyte rich subtype ( $p_{\text{adj}} = 0.412$ ), the not otherwise specified (NOS) subtype ( $p_{\text{adj}} = 0.244$ ) and the nodular lymphocyte predominant (NLPHL) subtype ( $p_{\text{adj}} = 0.412$ ) did not. Despite that, there was a clear trend towards seasonality in all histological subgroups. Stratification by age group



**Figure 1.** Forest plot of the subgroup analysis of seasonality of incidence. The seasonality amplitude is shown on the x-axis. Subgroups are shown on the y-axis. The estimated seasonality amplitude with 95% confidence intervals is shown for each subgroup.



**Figure 2.** Relationship between latitude quartile and estimated seasonality amplitude. Latitude quartiles are shown on the x-axis. The first, second, third and fourth latitude quartiles include all counties south of 34.19°N, between 34.19°N and 38.05°N, between 38.05°N and 41.68°N and north of 41.68°N, respectively. The estimated seasonality amplitude of the respective quartile is shown on the y-axis with 95% confidence intervals of the estimated amplitude.

revealed a significant seasonal incidence pattern for all age groups, apart from the age groups 40–49 ( $p_{\text{adj}} = 0.244$ ), 50–59 ( $p_{\text{adj}} = 0.084$ ) and >70 ( $p_{\text{adj}} = 0.160$ ). The amplitude was particularly high for the age groups 20–29 (0.121), 30–39 (0.114) and 60–69 (0.096), coinciding with higher HL incidence age groups<sup>26</sup>. Furthermore, stratification by Ann-Arbor stage and year of diagnosis revealed significant seasonality across these subgroups. Despite some non-significant results after adjusting for multiple testing, there was a clear trend towards seasonality in all subgroups (Fig. 1). The non-significant seasonality in some subgroups might be due to lower case numbers in the respective groups.

Looking at the stratification by latitude quartiles, all quartiles showed significant seasonality patterns when analyzed separately (Fig. 2). The amplitude was dependent on the latitude. Cases diagnosed at a latitude <38.05°N

		Hazard Ratio	Lower 95% CI	Upper 95% CI	p
3-year survival	All cases	1.030	0.981	1.081	0.234
	$\geq 38.05^\circ\text{N}$	<b>1.082</b>	<b>1.009</b>	<b>1.161</b>	<b>0.027</b>
	$< 38.05^\circ\text{N}$	0.990	0.926	1.059	0.772
5-year survival	All cases	1.027	0.983	1.073	0.238
	$\geq 38.05^\circ\text{N}$	<b>1.073</b>	<b>1.007</b>	<b>1.144</b>	<b>0.029</b>
	$< 38.05^\circ\text{N}$	0.991	0.932	1.053	0.764

**Table 3.** Results of the estimated Cox proportional-hazards model for all cases and stratified by latitude. Age at diagnosis, histological subtype, year of diagnosis, sex and Ann-Arbor stage were included into the model as known risk factors to control for their effect. Hazard ratio is the hazard ratio for overall mortality in the first three or five years after the HL diagnosis when being diagnosed in winter (September to February) versus summer (March to August). 95% confidence intervals of the hazard ratio are given. Significant results are shown in bold.

Peak risk month	Hazard Ratio	Lower 95% CI	Upper 95% CI	p
January	1.023	0.973	1.075	0.380
February	0.992	0.944	1.042	0.748
March	0.965	0.919	1.014	0.155
April	0.947	0.902	0.995	0.032
May	0.943	0.897	0.991	0.021
June	0.954	0.908	1.003	0.065
July	0.978	0.931	1.028	0.380
August	1.008	0.956	1.060	0.748
September	1.036	0.987	1.089	0.155
October	1.056	1.005	1.109	0.032
November	1.060	1.009	1.114	0.021
December	1.048	0.997	1.102	0.065

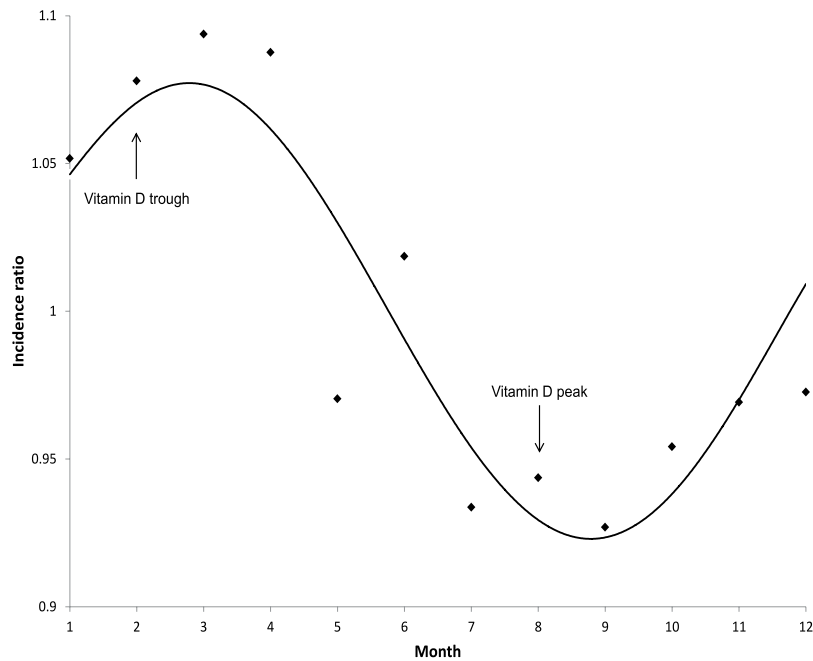
**Table 4.** Hazard ratio for the seasonal risk term for various peak risk months. The Cox proportional-hazards model included only cases diagnosed north of  $38.05^\circ\text{N}$ . The hazard ratios for different peak risk months are shown. A detailed explanation of the calculation of the seasonal risk term is given in the methods section. 95% confidence intervals of the hazard ratios are given.

(southern half of counties) had an amplitude of 0.055 ([0.026; 0.084],  $p < 0.001$ ) compared to 0.102 ([0.073; 0.131],  $p < 0.001$ ) for those diagnosed at a latitude  $\geq 38.05^\circ\text{N}$  (northern half of counties). The difference of 0.05 ([0.01; 0.09],  $p = 0.023$ ) between these groups was statistically significant.

**Impact of seasonality and latitude on outcome.** The results of the Cox proportional-hazards model for overall survival after a HL diagnosis with season of diagnosis as a dichotomous variable are given in Table 3. Looking at all cases together, being diagnosed in winter does not significantly increase the risk of dying compared to a diagnosis in summer within 3 years after diagnosis (HR = 1.030 [0.981, 1.081],  $p = 0.234$ ). However, when analyzed stratified by latitude, cases from the northern half (latitude  $\geq 38.05^\circ\text{N}$ ) of counties exhibited an increased risk of death when the diagnosis of HL was made in winter (HR = 1.082 [1.009, 1.161],  $p = 0.027$ ). In contrast, cases from the southern half of counties (latitude  $< 38.05^\circ\text{N}$ ) showed no increased risk (HR = 0.990 [0.926, 1.059],  $p = 0.772$ ). Confirmatory analysis using 5-year survival data showed similar results (Table 3).

To further examine the relationship between latitude and the increased risk of death after a winter or summer diagnosis, a multiplicative interaction term between the dichotomous season variable and the latitude each case was diagnosed at (in  $10^\circ$  steps) was included into the Cox proportional-hazards model described above. The hazard ratio for this interaction term was 1.119 ([1.009; 1.241],  $p = 0.033$ ) using 3 year survival data and, thus, significantly greater than 1. This interaction term indicates that increasing northern latitude increases the additional mortality risk after a winter versus summer diagnosis. Confirmatory analysis using 5-year survival data showed similar results with a hazard ratio for the interaction term of 1.105 ([1.005; 1.214],  $p = 0.039$ ).

In a final Cox proportional-hazards model with 3-year survival data, robustness of the cutoff months for dividing the year into a summer and a winter half year was assessed. A seasonal risk term, calculated as stated in the methods section, substituted the dichotomized season of diagnosis variable. Table 4 shows the estimated hazard ratios for the seasonal risk term for various  $M_{\max}$ , each reflecting a different peak risk month for cases diagnosed in the northern half (latitude  $\geq 38.05^\circ\text{N}$ ) of counties. The hazard ratio was highest for the seasonal risk term with  $M_{\max} = 11$  indicating the seasonal risk term implying November as the peak mortality risk month describes the seasonal mortality risk best. For  $M_{\max} = 11$ , the seasonal risk term takes positive values for the months September to January, a neutral value of 0 for the months February and August and negative values for



**Figure 3.** Illustration of the relationship between the peak and trough of vitamin D levels in humans in the northern hemisphere and the overall seasonality pattern of HL incidence. The months of the year are shown on the x-axis. The ratio of the respective month's incidence and the average incidence of all months are shown on the y-axis. The line indicates a fitted curve through the data as estimated by the cosinor model.

the months March to July. These intervals coincide with the chosen dichotomization of the year into a summer and a winter half year employed in the previous models. Note how  $\beta_1$  and, thus, the hazard ratio is lowest for  $M_{\min} = 5$ , reflecting the sinusoid property of the seasonal risk term. For  $M_{\min} = 5$ , the seasonal risk term is 1 for cases diagnosed in May, thus implying the lowest seasonal mortality risk in May. Finally, we tested a multiplicative interaction term as described above in this model, this time indicating the interaction between latitude (in  $10^\circ$  steps) and the seasonal risk term with  $M_{\max} = 11$ . The hazard ratio for this interaction term was 1.103 ([1.025; 1.187],  $p = 0.009$ ) and, thus, significantly greater than 1.

## Discussion

Here, we report the most comprehensive analysis of seasonal incidence and mortality patterns in HL performed to date and present several new findings.

The overall seasonal incidence pattern of HL in the northern hemisphere has a peak around March and a trough around September. Douglas and coworkers looked at seasonal incidence patterns for the different histological HL subtypes demonstrating seasonality for the nodular sclerosis subtype and the nodular lymphocyte predominant subtype<sup>10</sup>. We found similar results for the nodular sclerosis subtype and, in addition to that, a significant seasonal incidence pattern for the mixed cellularity and the lymphocyte depleted subtypes. The pattern described by Douglas *et al.*<sup>10</sup> for the nodular lymphocyte predominant subtype was not significant in the present study, although there was a trend towards seasonality in our data as well. Non-significance might be the result of lower case numbers in this rare subgroup. Neilly *et al.*<sup>5</sup> also performed a stratified analysis of seasonal incidence patterns by histological subtype and, very similar to our analysis, also found a significant seasonal incidence pattern for the mixed cellularity and the nodular sclerosis subtypes.

A stratified analysis by age group was undertaken by Chang *et al.*<sup>8</sup>, however, a comparison of the amplitude of the seasonality pattern in the different age groups was not performed. In addition, age groups were defined more broadly than in our analysis, possibly due to lower case numbers. Interestingly, the seasonal incidence pattern was strongly pronounced for the age groups 20–29 and 30–39. An overall incidence peak was observed in these age groups largely associated with Epstein-Barr virus (EBV) negative cases<sup>27</sup>. Data on EBV association are not provided in the SEER database, but one can assume that EBV status is an important factor in determining seasonal incidence patterns and that the seasonal incidence patterns vary between EBV positive and EBV negative cases. We were also able to show that the seasonal incidence pattern is more pronounced at higher latitudes.

We found a significantly increased risk of death in the first three years after a diagnosis of HL made in winter as compared to summer at higher latitudes. Seasonal differences in mortality risk after a diagnosis of HL have previously been described by Porojnicu *et al.*<sup>7</sup> although the authors only found a lower risk of mortality after being diagnosed in autumn versus winter. In our study, the increased risk of mortality after a winter diagnosis was confined to counties at higher latitudes ( $>38.05^\circ\text{N}$ ). All seasonal patterns of mortality or incidence have so far been described in fairly northern countries of high latitude in Europe<sup>5–10</sup>. We were able to show that the increased risk of dying after a winter diagnosis of HL is dependent on the latitude the diagnosis was made at in a way that a more northern latitude increases the risk of dying after being diagnosed with HL in winter. Robustness of our results

is shown in an additional model employing a sinusoid seasonal risk term. This model shows November to be the least favorable month and May as the most favorable month of diagnosis with regard to subsequent mortality risk.

We hypothesize that seasonal differences in vitamin D levels mediated by ultraviolet radiation (UVR) are at least in part responsible for the seasonal patterns described here. The seasonal pattern of vitamin D levels in humans is also more pronounced at higher latitudes<sup>28,29</sup>. Strikingly, the seasonal incidence pattern observed in HL cases almost exactly resembles the seasonal pattern of vitamin D levels in humans<sup>30</sup> (Fig. 3). In addition, the occurrence of a more pronounced seasonal incidence pattern among young patients could be a result of the higher seasonal variation in Vitamin D levels that has been described in this age group<sup>31,32</sup>. Recent epidemiological studies from Australia and the United States showed an increased incidence of HL with longer distances from the equator or lower annual ambient UVR, respectively<sup>33,34</sup>. Furthermore, increased ambient UVR was found to be likely protective, reducing HL incidence in a recent meta analysis<sup>35</sup>. Interestingly, a recent study that examined incidence of HL in Mediterranean countries found a higher HL incidence in countries bordering the northern Mediterranean Sea compared to those bordering the southern Mediterranean Sea<sup>36</sup>. However, these differences could also be explained by differences in sociodemographic characteristics or cancer registry methodology between these countries. A low Vitamin D status has been described as a risk factor for both, developing and suffering from a poor prognosis after a diagnosis of other hematological malignancies. Examples are follicular lymphoma<sup>37</sup> or chronic lymphocytic lymphoma<sup>38</sup>. New treatments for HL, such as immune checkpoint inhibitors<sup>39,40</sup>, might be partially dependent on patients having a functioning immune system. It is known that Vitamin D supports various immune system functions and can act anti-proliferative in various hematological cancers<sup>41,42</sup>.

Other causes of the seasonality described here are certainly possible, for example seasonality of an infectious agent, such as EBV, causing some HL cases. This would, however, not explain the observed seasonal differences in mortality and would contradict the finding that the seasonal incidence pattern of HL is, in fact, strongest in age groups with a fairly high proportion of EBV negative cases. Another possible explanation for the seasonal incidence pattern could be an increase in diagnoses of HL in late winter, due to the coincidence of the common cold season and an increase in consultations resulting in higher health care exposure and consultations for longer lasting reasons like lymph node swellings being postponed until after the holiday season. Lower incidence numbers in late summer could be explained by the summer holiday season. However, these alternative explanations do not explain the observed mortality pattern.

A potential shortcoming of our study is that many HL relevant variables and known risk factors are not available from the SEER program. Therefore it is possible that some variables exist that might at least partly explain the observed incidence pattern or increase mortality after a diagnosis in winter. For example, bulky disease, especially in the mediastinum is a known risk factor in HL<sup>4</sup> and might be more symptomatic in winter. Likewise, data to calculate the International Prognostic Score (IPS), a widely accepted score for risk stratification of HL, is not available. However, by including all available potential risk factors into the Cox proportional-hazards models used to estimate seasonal differences in mortality, we tried to control for potential confounders as good as possible with the available data. Additionally, we compared available disease and patients characteristics between the highest and lowest risk months for mortality (November vs. May) (Supplementary Table 1) and the peak and trough months for incidence (March vs. September) (Supplementary Table 2). Disease and patient characteristics were not different between the compared months.

In conclusion, we found a striking seasonal pattern of incidence and mortality in HL. Further studies are needed in order to better understand the reasons for the seasonal patterns described here.

## References

- Engert, A. *et al.* Reduced-intensity chemotherapy and PET-guided radiotherapy in patients with advanced stage Hodgkin's lymphoma (HD15 trial): a randomised, open-label, phase 3 non-inferiority trial. *Lancet* **379**, 1791–9 (2012).
- Engert, A. *et al.* Reduced treatment intensity in patients with early-stage Hodgkin's lymphoma. *N. Engl. J. Med.* **363**, 640–52 (2010).
- Böll, B. *et al.* ABVD in older patients with early-stage hodgkin lymphoma treated within the german hodgkin study group HD10 and HD11 trials. *J. Clin. Oncol.* **31**, 1522–1529 (2013).
- Borchmann, S., von Tresckow, B. & Engert, A. Current developments in the treatment of early-stage classical Hodgkin lymphoma. *Curr. Opin. Oncol.* **28**, 377–383 (2016).
- Neilly, I. J., Dawson, A. A., Bennett, B. & Douglas, S. Evidence for a seasonal variation in the presentation of Hodgkins disease. *Leuk. Lymphoma* **18**, 325–8 (1995).
- Douglas, S., Cortina-Borja, M. & Cartwright, R. A quest for seasonality in presentation of leukaemia and non-Hodgkin's lymphoma. *Leuk. Lymphoma* **32**, 523–32 (1999).
- Porojnicu, A. C., Robsahm, T. E., Ree, A. H. & Moan, J. Season of diagnosis is a prognostic factor in Hodgkin's lymphoma: a possible role of sun-induced vitamin D. *Br. J. Cancer* **93**, 571–4 (2005).
- Chang, E. T., Blomqvist, P. & Lambe, M. Seasonal variation in the diagnosis of Hodgkin lymphoma in Sweden. *Int. J. Cancer* **115**, 127–30 (2005).
- Westerbeek, R. M. *et al.* Seasonal variations in the onset of childhood leukaemia and lymphoma. *Br. J. Cancer* **78**, 119–24 (1998).
- Douglas, S., Cortina-Borja, M. & Cartwright, R. Seasonal variation in the incidence of Hodgkin's disease. *Br. J. Haematol.* **103**, 653–62 (1998).
- Surveillance, Epidemiology, and End Results (SEER) Program ([www.seer.cancer.gov](http://www.seer.cancer.gov)) SEER\*Stat Database: Incidence - SEER 9 Regs Research Data, Nov 2014 Sub (1973–2012) Katrina/Rita Population Adjustment- Linked To County Attributes - Total U.S., 1969–2013.
- Surveillance, Epidemiology, and End Results (SEER) Program (<http://seer.cancer.gov/data/>).
- Surveillance Research Program, National Cancer Institute SEER\*Stat software ([seer.cancer.gov/seerstat](http://seer.cancer.gov/seerstat)) version 8.2.1.
- The R Foundation for Statistical Computing R version 3.2.2. (<https://www.r-project.org/>).
- R Commander version 2.2.3. (<http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/>).
- Fox, J. The R Commander: A Basic-Statistics Graphical User Interface to R. *J. Stat. Softw.* **14**, 1–42 (2005).
- Therneau, T. M. & Lumley, T. Survival: Survival Analysis. *survival: Survival Analysis* Available at: <https://cran.r-project.org/package=survival> (2016).
- Sachs, M. cosinor: Tools for estimating and predicting the cosinor model. Available at: <https://cran.r-project.org/package=cosinor> (2014).

19. Holm, S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand. J. Stat.* **6**, 65–70 (1979).
20. Nelson, W., Tong, Y. L., Lee, J. K. & Halberg, F. Methods for cosinor-rhythmometry. *Chronobiologia* **6**, 305–23.
21. Tong, Y. L. Parameter estimation in studying circadian rhythms. *Biometrics* **32**, 85–94 (1976).
22. Geography, U. C. B. 2010 Census Gazetteer Files Record Layouts.
23. Cox, D. R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B* **34**, 187–220 (1972).
24. Glossary of the American Meteorological Society. Available at: <http://glossary.ametsoc.org/>.
25. Efrid, J. T. Sinusoidal cox regression—a rare cancer example. *Cancer Inform.* **9**, 265–79 (2010).
26. Correa, P. & O'Connor, G. T. Epidemiologic patterns of Hodgkin's disease. *Int. J. Cancer* **8**, 192–201 (1971).
27. Jarrett, R. F. *et al.* The Scotland and Newcastle epidemiological study of Hodgkin's disease: impact of histopathological review and EBV status on incidence estimates. *J. Clin. Pathol.* **56**, 811–6 (2003).
28. Webb, A. R., Kline, L. & Holick, M. F. Influence of season and latitude on the cutaneous synthesis of vitamin D3: exposure to winter sunlight in Boston and Edmonton will not promote vitamin D3 synthesis in human skin. *J. Clin. Endocrinol. Metab.* **67**, 373–8 (1988).
29. Kimlin, M. G., Olds, W. J. & Moore, M. R. Location and vitamin D synthesis: is the hypothesis validated by geophysical data? *J. Photochem. Photobiol. B.* **86**, 234–9 (2007).
30. Kasahara, A. K., Singh, R. J. & Noymer, A. Vitamin D (25OHD) Serum Seasonality in the United States. *PLoS One* **8**, e65785 (2013).
31. Maeda, S. S. *et al.* Seasonal variation in the serum 25-hydroxyvitamin D levels of young and elderly active and inactive adults in São Paulo, Brazil: The São Paulo Vitamin D Evaluation Study (SPADES). *Dermatoendocrinol.* **5**, 211–7 (2013).
32. Degerud, E. *et al.* Cosinor modelling of seasonal variation in 25-hydroxyvitamin D concentrations in cardiovascular patients in Norway. *Eur. J. Clin. Nutr.* <https://doi.org/10.1038/ejcn.2015.200> (2015).
33. van Leeuwen, M. T. *et al.* Latitude gradients for lymphoid neoplasm subtypes in Australia support an association with ultraviolet radiation exposure. *Int. J. Cancer* **133**, 944–951 (2013).
34. Bowen, E. M. *et al.* Relationship between ambient ultraviolet radiation and Hodgkin lymphoma subtypes in the United States. *Br. J. Cancer* **114**, 826–831 (2016).
35. Monnereau, A. *et al.* Exposure to UV radiation and risk of Hodgkin lymphoma: a pooled analysis. *Blood* **122**, 3492–3499 (2013).
36. Salati, M., Cesaretti, M., Macchia, M., El Mistiri, M. & Federico, M. Epidemiological overview of hodgkin lymphoma across the mediterranean basin. *Mediterr. J. Hematol. Infect. Dis.* **6**, 2014048 (2014).
37. Kelly, J. L. *et al.* Low Serum Vitamin D Levels Are Associated With Inferior Survival in Follicular Lymphoma: A Prospective Evaluation in SWOG and LYSA Studies. *J. Clin. Oncol.* **33**, 1482–90 (2015).
38. Shanafelt, T. D. *et al.* Vitamin D insufficiency and prognosis in chronic lymphocytic leukemia. *Blood* **117**, 1492–8 (2011).
39. Ansell, S. M. *et al.* PD-1 Blockade with Nivolumab in Relapsed or Refractory Hodgkin's Lymphoma. *N. Engl. J. Med.* **372**, 311–319 (2015).
40. Armand, P. *et al.* Programmed Death-1 Blockade With Pembrolizumab in Patients With Classical Hodgkin Lymphoma After Brentuximab Vedotin Failure. *J. Clin. Oncol.* [JCO673467](https://doi.org/10.1200/JCO.2016.67.3467), <https://doi.org/10.1200/JCO.2016.67.3467> (2016).
41. Hall, A. C. & Juckett, M. B. The role of vitamin D in hematologic disease and stem cell transplantation. *Nutrients* **5**, 2206–21 (2013).
42. Baeke, F., Takiishi, T., Korf, H., Gysemans, C. & Mathieu, C. Vitamin D: modulator of the immune system. *Curr. Opin. Pharmacol.* **10**, 482–96 (2010).

## Acknowledgements

S.B. has been supported by a Köln Fortune Grant (181/2016).

## Author Contributions

S.B. designed the study. S.B., H.M. and A.E. analyzed and interpreted the data, contributed to the scientific interpretation of the results, contributed to writing the paper and critically revised it for intellectual content.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-017-14805-y>.

**Competing Interests:** The authors declare that they have no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017