**The lognormal age-of-onset distribution in Perthes disease: An analysis of 992 cases.**

**Abstract**

**Aims and Objective:** It has been proposed that the age-of-onset of Perthes disease follows a lognormal distribution, which is important for considering aetiological determinants, and validating the adequacy of epidemiological studies. This study tested conformity to the distribution, and considered the effect of sex, laterality and geography on age-of-onset.

**Methods:** Cases were identified from the Merseyside Perthes Register, between 1976 and 2010. Conformity to a lognormal distribution was tested with chi-squared test of normality and the Shapiro Wilk test. Stratification was conducted by sex, laterality and geographical area in order to identify differences in the age-of-onset distributions.

**Results:** 992 cases of Perthes disease were considered, of which 673 arose from a geographically defined region. The age-of-onset demonstrated a marked positive skew, which fitted well to a lognormal distribution, especially so within the defined geographic region thereby minimizing data truncation. Disease onset was 6-months earlier in females (p=0.01), and 1-year earlier in those who go onto develop bilateral disease (p<0.001). There was no difference in the age-of-onset between geographic regions with differing incidence rates.

**Conclusion:** The age-of-onset in Perthes disease conforms well to a log-normal model, which both allows comparisons to infectious disease epidemiology, and allows validation of epidemiological studies. Premature onset amongst females and those who develop bilateral disease, may offer clues in understanding the aetiological determinant of disease.

**The lognormal age-of-onset distribution in Perthes disease An analysis of 992 cases.**

**Background**

Perthes disease is an idiopathic avascular necrosis of the femoral head in children, and a common precursor to hip osteoarthritis in young adults. The disease aetiology, determinants and mechanism are unknown. Perthes disease typically occurs in boys between 4 and 8 years old. The age-of-onset demonstrates a positive skew, such that the mode of disease onset is around 5 years old, and mean around 6 years old. Transformation of the data suggests that the age-of-onset follows a lognormal distribution [1](http://wizfolio.com/?citation=1&ver=3&ItemID=419&UserID=3268&AccessCode=4BC7EAF00A504F69B29B32627A6BB450&CitationSuffix=)[,2](http://wizfolio.com/?citation=1&ver=3&ItemID=540&UserID=3268&AccessCode=99871EA7140748C8B70CAA1AA9428871&CitationSuffix=).

Lognormal distributions are unusual in chronic disease, and their occurrence is more typically related to incubation periods seen in infectious disease epidemiology [3](http://wizfolio.com/?citation=1&ver=3&ItemID=681&UserID=3268&AccessCode=3C4D9C39619F479B8DCAD6874DEF704A&CitationSuffix=)[,4](http://wizfolio.com/?citation=1&ver=3&ItemID=683&UserID=3268&AccessCode=09527964AF9843DEAAF95CAD619AF7F8&CitationSuffix=). Criticism has been drawn concerning the validity of the lognormal distribution in Perthes disease age-of-onset, and the statistical methods employed [5](http://wizfolio.com/?citation=1&ver=3&ItemID=615&UserID=3268&AccessCode=0&CitationSuffix=). It therefore remains uncertain if a lognormal distribution is the correct approximation in Perthes disease.

Understanding the statistical validity of a lognormal distribution in Perthes disease is useful, both to consider the age-of-onset as an ‘incubation-period’, and in order to act as a tool for externally validating the results of epidemiological studies. Considering the age-of-onset as an ‘incubation-period’ will facilitate greater understanding to the timing of an ‘insult’ and the nature of aetiological determinant(s)

Using the Liverpool Perthes Disease Register we consider the age-of-onset distribution, and it’s fit to a lognormal model, with reference to sex, laterality of disease and region of residence.

**Methods**

**The Liverpool Perthes Register**

The Liverpool Perthes disease register is the largest series of Perthes disease cases, commencing in 1976. The methods underpinning the register have been described previously[6](http://wizfolio.com/?citation=1&ver=3&ItemID=884&UserID=3268&AccessCode=B154BCD47672472091B736F65C96A32A&CitationSuffix=). It is a register of incident cases maintained at Alder Hey Childen’s Hospital, Liverpool, UK. Information collated for each case includes date of birth, date at disease onset (defined by first radiographic diagnosis of disease), laterality of disease, sex and area of residence.

The date of disease onset is recorded as a month and year. All cases were therefore assumed to be incident of the 15th day of the month, to allow the calculation of age of disease onset (in days).

**Analysis**

Data analysis was conducted using Stata 10.0. The age-of-onset distributions were explored and tested for fit to the lognormal model using the chi-squared test of normality and the Shapiro Wilk test. Analysis was conducted for the entire dataset (i.e. tertiary and quaternary referrals) and for the area exclusively supplied by Alder Hey hospital (Liverpool, Sefton and Knowsley) (i.e. tertiary referrals).

Subgroup analyses of laterality, sex and region of diagnosis were undertaken, using parametric tests following data transformation. Comparisons were made using oneway ANOVA, with the Bonferroni correction.

Tests of normality are based on the null hypothesis that the dataset follows a normal distribution. A test of a normality for which p<0.05 therefore indicates that the alternative hypothesis is accepted (i.e. the data does not appear to follow a normal distribution).

**Results**

1082 cases of Perthes disease were recorded within the register to December 2010. 682 of these cases arose from a geographical defined area that Alder Hey Hospital was the sole provider of paediatric orthopaedic services (Liverpool, Sefton and Knowsley). Within the register, 90 cases did not have a clear date of onset – largely as a result of the child been referred for quaternary care of established disease. Such cases were excluded. Of those arising within Liverpool, Sefton and Knowsley, in only nine of the 682 cases was the date of onset absent or ambiguous, and similarly excluded.

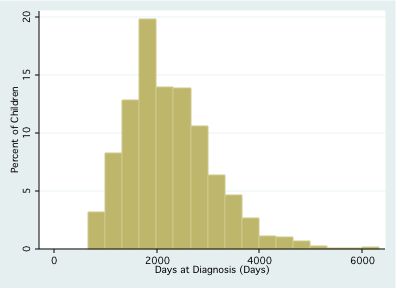
992 cases were available for inclusion. The mean onset of disease was 2260 days (6.2 years), median 2135 days (5.8 years) and mode 1834 years (5.02 years). This therefore corresponds to a distribution with a positive skew (skewness 0.92).

**Testing Lognormality**

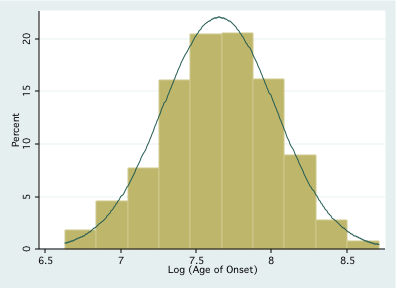
Visually the entire dataset fitted well to a lognormal model (Fig A, B). Statistically this was significant using chi-squared test of normality at the 5% level (p = 0.10), yet not using Shapiro Wilk test (p = 0.01).

Examining those patients from the area exclusively served by Alder Hey hospital (n=673) revealed significant results, using both chi-squared (p = 0.16) and Shapiro Wilk (p = 0.08) tests of normality.

**Fig A –** Histogram of the age-of-onset distribution of cases of Perthes disease (days).

****

**Fig B –** Histogram of Logarithmic transformation of age-of-onset of Perthes disease, with a fitted curve.

****

**Male vs. Female Age of Onset**

Female onset was almost 6 months before males (p=0.01). Table 1.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sex** | **Cases** | **Mean Ln**  **(Age of Onset)** | **Age of Onset**  **(Years)** |
| Female | 167 | 7.59 | 5.40 (5.10 – 5.70) |
| Male | 825 | 7.67 | 5.85 (5.69 – 5.99) |

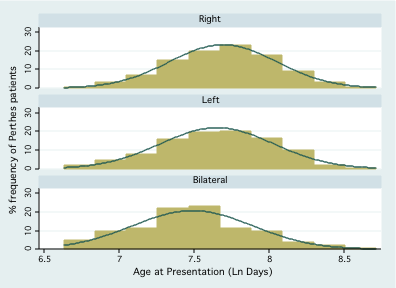
**Side of Disease**

There was no significant difference between the age of onset of right and left disease (p=0.35), yet bilateral disease had a disease onset approximately 1 year earlier than either left or right disease (p<0.001). Illustrated in Fig C.

|  |  |  |  |
| --- | --- | --- | --- |
| **Sex** | **Cases** | **Mean Log Age of Onset**  **(Days)** | **Age of Onset using logarithmic mean (Years)** |
| Right | 457 | 7.69 (7.66 – 7.73) | 6.00 (5.81 – 6.20) |
| Left | 389 | 7.65 (7.61 – 7.69) | 5.77 (5.55 – 5.98) |
| Bilateral | 122 | 7.49 (7.42 – 7.56) | 4.90 (4.56 - 5.26) |

\* In 24 cases the side of disease was not recorded.

**Fig C –** Stacked histograms of Ln(Age of Onset) of Perthes disease by laterality of disease, with fitted normal-density plot.

****

**Region of Residence**

There was no statistically significant change in the age of onset of disease between regions within Merseyside (Liverpool vs. Sefton: p=0.40, Liverpool vs. Knowsley p=1.0, Sefton vs. Knowsley p=1.0).

|  |  |  |  |
| --- | --- | --- | --- |
| **Period** | **Cases** | **Mean Ln**  **(Age of Onset)** | **Age of Onset**  **(Years)** |
| Liverpool | 379 | 7.65 | 5.74 |
| Sefton | 146 | 7.64 | 5.44 |
| Knowsley | 148 | 7.63 | 5.66 |

**Discussion**

The Liverpool Perthes Register demonstrates conformity of the dataset to a lognormal distribution. It also demonstrates a significantly earlier onset in disease amongst girls and amongst children who progress to bilateral disease. No age of onset difference was identified between the three geographical regions studies.

Sartwell first demonstrated that infectious disease incubation periods follow a lognormal distribution [7](http://wizfolio.com/?citation=1&ver=3&ItemID=682&UserID=3268&AccessCode=2544964669CB4B90AE6017B9E11FCFCB&CitationSuffix=) – such graphic representations are now often referred to as ‘Sartwell curves’. In infectious diseases this distribution exists when a disease arises from a point-source outbreak (i.e. episode of food poisoning), yet it is not present in broad ill-defined infective outbreaks (e.g. water borne thypoid). Sartwell suggested that chronic diseases may similarly follow such patterns, which may offer insight into the ‘incubation-period’, and hence the timing of the aetiological insult [8](http://wizfolio.com/?citation=1&ver=3&ItemID=775&UserID=3268&AccessCode=C49D7C1E994C46908D5570DB10500AE7&CitationSuffix=). The suggestion is that the incubation period of a chronic disease begins at the time of exposure to a pathogen (be it environmental or genetic), and ends at the point of disease onset. If the exposure is in-utero then the incubation period is the summation of the gestational period following exposure, and the age of disease onset.

A number of chronic diseases have been investigated using Sartwell’s model. The age of onset of diseases with a recognised genetic biochemical abnormality conform well to the model (e.g. Wilson’s disease, cystinurea and Hartnup disease), whereas those diseases known to have a multifactorial aetiology (e.g systemic lupus erythematosus, idiopathic hypoparathyroidism and sudden infant death syndrome) do not [9](http://wizfolio.com/?citation=1&ver=3&ItemID=684&UserID=3268&AccessCode=58A7805205F845F5A5E13E21820401EE&CitationSuffix=). Diseases with a defined environmental trigger similarly conform; such as occupational lung cancer and pancytopenia following chloramphenicol exposure [10](http://wizfolio.com/?citation=1&ver=3&ItemID=685&UserID=3268&AccessCode=FB929EA8738D409196E54CC8963EFAC7&CitationSuffix=). Interestingly, an investigation comparing the age-of-onset distribution of acute lymphoblastic leukaemia following intrauterine radiation (point source exposure) and idiopathic disease (unknown exposure time), demonstrated that only the former followed a lognormal distribution [10](http://wizfolio.com/?citation=1&ver=3&ItemID=685&UserID=3268&AccessCode=FB929EA8738D409196E54CC8963EFAC7&CitationSuffix=).

Criticism has been leveled against the use of Sartwell curves in chronic disease [5](http://wizfolio.com/?citation=1&ver=3&ItemID=615&UserID=3268&AccessCode=0&CitationSuffix=). The plausibility has been questioned as it is believed that infectious disease conforms to this model through replication of infectious agents (bacteria/ virus), with disease presentation at a threshold. It may be argued that neoplastic disease may be modeled in a similar way through the accrual of malignant cells, yet the mechanism by which other diseases, including Perthes disease, may reflect such a model is unclear. Other criticisms concern the truncation of data, or shifts within populations which, it is argued, may artificially produce the appearance of a lognormal distribution. The statistical test used have also been criticized as the chi-squared tests of normality is usually used in preference to more powerful tests i.e. Shapiro Wilk. This analysis has therefore defined a clear geographic region (Liverpool/ Sefton and Knowsley) in order to limit data truncation, and has used the more powerful Shapiro Wilk statistic. We accept that the biological plausibility of the lognormal assumption in Perthes disease remains questionable, yet we provide clear graphic and statistical evidence that this distribution exists in a large geographically defined case series.

The earliest case of Perthes disease within our dataset was 2.0 years old – a figure similarly identified by Hall and Barker [1](http://wizfolio.com/?citation=1&ver=3&ItemID=419&UserID=3268&AccessCode=4BC7EAF00A504F69B29B32627A6BB450&CitationSuffix=). Applying Sartwell’s ideas to chronic disease epidemiology therefore supports the notion that the aetiological trigger in Perthes disease acts before 2 years old, and suggests that there may be a single vulnerable period during which the aetiological trigger acts – to replicate a ‘point source’ exposure. It is however unclear whether such an ‘exposure’ acts prenatally, or postnatally. It is increasingly apparent that there is a very strong association between Perthes disease and deprivation [11](http://wizfolio.com/?citation=1&ver=3&ItemID=421&UserID=3268&AccessCode=4218FE2BEFAA4602A3AE33A14B46B4CD&CitationSuffix=)[E](http://wizfolio.com/?citation=1&ver=3&ItemID=336&UserID=3268&AccessCode=931B87AAD6FC415B93144BCBED02C38C&CitationSuffix=)[-13](http://wizfolio.com/?citation=1&ver=3&ItemID=533&UserID=3268&AccessCode=5B51967FA99F41A78C35435F29EBD3DC&CitationSuffix=), and therefore it appears unlikely that there is a significant genetic component of disease.

Other observers have speculated that the age distribution of Perthes disease may differ with respect to sex and laterality, but have been unable to confirm observations statistically [14](http://wizfolio.com/?citation=1&ver=3&ItemID=319&UserID=3268&AccessCode=36424C1B61FE4505AD4D9B0AD13DC043&CitationSuffix=)[E](http://wizfolio.com/?citation=1&ver=3&ItemID=633&UserID=3268&AccessCode=0&CitationSuffix=)[-16](http://wizfolio.com/?citation=1&ver=3&ItemID=774&UserID=3268&AccessCode=7893574E8FFA4F7293C0A43B4F5265F4&CitationSuffix=). Our data statistically confirms that the disease onset is younger in females and in cases of bilateral disease. Others have suggested that the age of disease onset is lower in areas of greater disease incidence [17](http://wizfolio.com/?citation=1&ver=3&ItemID=349&UserID=3268&AccessCode=4EA93B51C0594378AE108D53C773010F&CitationSuffix=). It has previously been demonstrated within Merseyside, that the geographical disease distribution varies considerably, with annual incidence rates of around 10 per 100,000 0 -14 years olds within Liverpool and Knowsley and 5 per 100,000 in Sefton [11,](http://wizfolio.com/?citation=1&ver=3&ItemID=421&UserID=3268&AccessCode=4218FE2BEFAA4602A3AE33A14B46B4CD&CitationSuffix=)[13](http://wizfolio.com/?citation=1&ver=3&ItemID=533&UserID=3268&AccessCode=5B51967FA99F41A78C35435F29EBD3DC&CitationSuffix=). No difference in the age-of-onset could be identified between these areas, despite the marked difference in incidence. It is unclear why those with bilateral disease, and girls may be affected earlier. It is predominantly a disease of males, therefore it may postulated that girls have a innate resistance to disease and therefore require a greater ‘dose’ of exposure to precipitate disease, which similarly may be the case in those with bilateral disease, with a consequent reduction in the age-of-onset – though ‘incubation’ is typically independent of dose [5,](http://wizfolio.com/?citation=1&ver=3&ItemID=615&UserID=3268&AccessCode=0&CitationSuffix=)[10](http://wizfolio.com/?citation=1&ver=3&ItemID=685&UserID=3268&AccessCode=FB929EA8738D409196E54CC8963EFAC7&CitationSuffix=). MENTION INDIA HAS A HIGHER AGE OF DISEASE.

This study therefore demonstrates that the lognormal distribution is an appropriate model of Perthes disease age-of-onset, and this may have wider implications for understanding the disease aetiology. This relationship is also be important in validating the appropriateness of Perthes disease datasets. Similar studies concerning the age and gender distribution of childhood malignancy have provided important epidemiological clues regarding their aetiology and pathogenesis, and it is hoped that similar successes may be achieved in understanding of Perthes disease.

[**References**](https://wizfolio.com/?style=1&ver=3&UserID=3268&StyleName=Style%2020101111%20092316)

**1.** Hall AJ, Barker DJ. The age distribution of Legg-Perthes disease. An analysis using Sartwell's incubation period model. Am J Epidemiol. 1984;120(4):531-36.

**2.** Wiig O, Terjesen T, Svenningsen S, Lie SA. The epidemiology and aetiology of Perthes' disease in Norway. A nationwide study of 425 patients. J Bone Joint Surg Br. 2006;88(9):1217-23.

**3.** Sartwell PE. The incubation period and the dynamics of infectious disease. Am J Epidemiol. 1966;83(2):204-06.

**4.** Sartwell PE. The incubation period of poliomyelitis. Am J Public Health Nations Health. 1952;42(11):1403-08.

**5.** Horner RD, Samsa G. Criteria for the use of Sartwell's incubation period model to study chronic diseases with uncertain etiology. J Clin Epidemiol. 1992;45(10):1071-80.

**6.** Perry DC, Bruce CE, Pope D, Dangerfield P, Platt MJ, Hall AJ. Perthes' disease: deprivation and decline. Arch Dis Child. 2011;96(12):1124-28.

**7.** Sartwell PE. The distribution of incubation periods of infectious disease. American journal of hygiene. 1950;51(3):310-18.

**8.** Sartwell P. Some approaches to the epidemiologic study of chronic disease. Am J Public Health Nations Health. 1955;45(5):609-14.

**9.** Armenian HK, Khoury MJ. Age at onset of genetic diseases: an application for Sartwell's model of the distribution of incubation periods. Am J Epidemiol. 1981;113(5):596-605.

**10.** Armenian HK, Lilienfeld AM. The distribution of incubation periods of neoplastic diseases. Am J Epidemiol. 1974;99(2):92-100.

**11.** Hall AJ, Barker DJ, Dangerfield PH, Taylor JF. Perthes' disease of the hip in Liverpool. BMJ. 1983;287(6407):1757-59.

**12.** Kealey WD, Moore AJ, Cook S, Cosgrove AP. Deprivation, urbanisation and Perthes' disease in Northern Ireland. J Bone Joint Surg Br. 2000;82(2):167-71.

**13.** Margetts BM, Perry CA, Taylor JF, Dangerfield PH. The incidence and distribution of Legg-Calvé-Perthes' disease in Liverpool, 1982-95. Arch Dis Child. 2001;84(4):351-54.

**14.** Fisher RL. An epidemiological study of Legg-Perthes disease. J Bone Joint Surg Am. 1972;54(4):769-78.

**15.** Guille JT, Lipton GE, Szöke G, Bowen JR, Harcke HT, Glutting JJ. Legg-Calvé-Perthes disease in girls. A comparison of the results with those seen in boys. J Bone Joint Surg Am. 1998;80(9):1256-63.

**16.** Burch PR, Nevelös AB. Perthes' disease: a new genetic hypothesis. Med Hypotheses. 1979;5(5):513-28.

**17.** Barker DJ, Dixon E, Taylor JF. Perthes' disease of the hip in three regions of England. J Bone Joint Surg Br. 1978;60(4):478-80.