Study protocol 13_210R - approved by the Independent Scientific Advisory Committee for Clinical Practice Research Datalink research

The incidence of Lyme disease derived from primary care data in the UK

Lay Summary

The estimated number of new cases of Lyme disease (LD) in Europe per year varies widely between and within countries. Public Health England estimated the number of new cases in England and Wales in 2011 to be 959, but this is based on laboratory statistics and excludes clinically diagnosed cases. In Scotland, where reporting of LD is mandatory, the number of LD cases reported in 2011 was 229. In contrast, an estimate of over 200,000 new cases per year has been reported for Germany which suggests a potential underestimation of LD in the UK. To address the risk that LD is currently being underestimated, it is proposed to investigate the number of new cases of LD recorded by general practitioners (GPs) in the UK in the years 2001 to 2013. Accurate estimates of the incidence of LD may increase awareness of the disease and thus contribute to improving diagnosis and early treatment, so data from the CPRD will provide a valuable addition to the literature on LD.

Objectives

The primary aim of the study will be to estimate the incidence rate of LD as recorded by GPs in the UK, by calendar year since 2001 and by region.

Additional aims will be to estimate the frequency of use of GP-initiated laboratory tests and antibiotic treatment among patients with incident LD.

Background

LD (also known as Lyme borreliosis) is a bacterial infection spread by ticks and is the most common tick-borne infection in many parts of Europe and the United States. The reported numbers of cases of LD in Europe vary widely between and within countries. In a WHO report published in 2006 [1], the authors wrote that about 85,000 cases are reported annually in Europe but they also write that this is a considerable underestimate as many LD infections go undiagnosed and the methods used to count cases of LD are inconsistent and incomplete. For 2005, the European Commission [2] published a total of 20,129 LD cases in Europe (although with missing data for many countries).

Some areas have established mandatory reporting of LD, including Scotland and six states of northeastern Germany. Health Protection Scotland reported a total 229 LD cases in 2011 [3]. In contrast, in the six states of northeastern Germany, 5568 new cases of LD were reported in 2008 which translates to roughly 28,000 incident cases for all of Germany [4]. This estimate

is well below the estimate of 60,000–100,000 incident cases per year in Germany based on an older seroprevalence study in a single region of Germany [5], suggesting that a large proportion of cases are not registered in the six northeastern states despite the reporting being mandatory.

Presentation with the characteristic rash erythema migrans is considered sufficient for a clinical diagnosis of LD without any laboratory confirmation [6,7], and in continental Europe roughly 90% of cases present with erythema migrans [7]. However, there are geographic variations in the distribution of the different Borrelia genospecies, and the proportion of LD cases with erythema migrans may be lower in the UK. The British Infection Association noted that a significant proportion of infected UK ticks carry Borrelia valaisiana which only rarely causes erythema migrans [7].

Another approach to estimate the incidence of LD is based on laboratory testing. Public Health England estimated the number of new cases in England and Wales to be 959 in 2011 [8], but this is based only on laboratory statistics and excludes clinically diagnosed cases. However, the test validity with both false negative and false positive serologic tests limits its use. It has therefore been recommended that "serology should only be used to confirm but not to primarily establish the diagnosis of LD" [9].

The most reliable set of data for Germany, the country with the largest number of cases in Europe, appears to come from Mueller et al [9] who evaluated all patients with a diagnosis of LD recorded by the German health insurance company Deutsche Angestellten Krankenkasse (DAK) between 2006 and 2008. The DAK insures patients across all German states and covers the health costs of 7.3% of the German population. Mueller et al looked at all patients with a diagnosis of LD with erythema migrans (ICD-10 code A69.2). They found a yearly incidence rate of 261 per 100,000 in the population. Their extrapolation to the whole German population (ca. 82 million people) resulted in an incidence estimate of 213,913 new cases of LD with erythema migrans in Germany per year. That is well beyond all other estimates for Germany.

Similarly, in the U.S. where LD is notifiable, the Centers for Disease Control (CDC) has recently increased their estimate of the number of new cases per year in the U.S. from 30,000 based on the number reported to the CDC per year to 300,000 based on survey data and medical claims data [10].

Early diagnosis and rapid antibiotic treatment of LD is important to help prevent the development of more serious illness and, as seen in a small proportion of LD patients, long-term morbidity persisting after the recommended antibiotic treatment [11]. Accurate estimates of the incidence of LD may increase awareness of the disease and thus improve diagnosis and

early treatment, so data from the CPRD will provide a valuable addition to the literature on LD.

It is proposed to take an approach for the UK similar to that taken by the CDC and Mueller et al by analyzing data from the CPRD in order to obtain more reliable estimates of the incidence of LD recorded by GPs in the UK.

Study type

This will be a purely descriptive study.

Study design

Cohort study of patients with an incident diagnosis of LD.

Comparison group

Not applicable.

Primary study population – patients with an incident diagnosis of LD

Given the healthcare system in the UK, it is possible that LD is diagnosed by the GP or by a specialist. Therefore, the LD study population will be formed from:

> (1) patients with a Read medical code for LD (Table 1),
>
> (2) patients with a mentioning of LD in the medical notes <u>and</u> with a Read medical code indicating a secondary care visit or a referral to secondary care identified from entity type,
>
> (3) patients with a mentioning of LD in the medical notes <u>and</u> with a prescription of a defined antibacterial medication (see Table 2) issued by the GP on the day of the LD diagnosis, and
>
> (4) patients with a recording or mentioning of a non-negative Borrelia test (Table 3) <u>and</u> with a prescription of a defined antibacterial medication issued by the GP on the day of the Borrelia test.

To obtain free text on LD from the medical notes, a search for the following key words will be made: borreli, borrelli, boreli, borelli, lyme dis, erythema chron migrans. The surrounding text (up to 20 words before and after) will be extracted and anonymized, so that diagnoses of LD can be detected in a subsequent manual review.

Table 1: Read medical codes for the diagnosis of LD

Read medical code	Read term
A871000	Lyme disease
N010A00	Arthritis in Lyme disease
AA41.00	Erythema chronicum migrans

Table 2: Recommended antibiotic treatment for LD

Antibacterial medication
Doxycycline
Azithomycin
Amoxicillin
Cefuroxime axetil
Ceftriaxone
Cefotaxime
Penicillin G

Table 3: Read medical codes for LD laboratory tests (including negative test results)

Read medical code	Read term
43e1.00	Borrelia burgdorferi antibody level
43e0.00	Borrelia burgdorferi IgG level
43e3.00	Borrelia burgdorferi blot test
43ek.00	Borrelia burgdorferi IgM level
43j2.00	Borrelia burgdorferi nucleic acid detection
43T..00	Lyme disease test
43T0200	Lyme ELISA equivocal
43T0000	Lyme ELISA negative
43T0100	Lyme ELISA positive
43T0300	Lyme ELISA reactive
43T0.00	Lyme ELISA test
43T1200	Lyme immunoblot equivocal
43T1000	Lyme immunoblot negative
43T1100	Lyme immunoblot positive
43T1.00	Lyme immunoblot test

All patients meeting any one of the above four criteria for LD will be counted in the incidence assessments. Patients meeting more than one criterion will have the earliest used for the incidence calculation. To identify incident cases ~~each year~~, only those patients who did not have an LD diagnosis in the previous 12 months before the diagnosis will be included in the cohort, and only patients with at least 12 months prior registration in the database will be evaluated. As patients can be re-infected with LD, patients with a new diagnosis of LD more than 12 months apart from an LD diagnosis meeting all case definition criteria will be counted as having more than one incident diagnosis of LD over the full study period.

Validity of LD diagnosis

All diagnoses of LD entered in HES will be evaluated to test the sensitivity of our case definition based on GP records and notes.

Study period

January 2001 to June 2013.

Sample size

The sample size is expected to generate an incidence estimate for recent years higher than the total of 1188 cases reported by Public Health England and Health Protection Scotland together for 2011, and so the number of cases in CPRD over the years 2001-2013 is expected to be several thousand.

Baseline characteristics of patients with an incident diagnosis of LD

Baseline characteristics of interest will include age, gender, calendar year, region, season of incident diagnosis, socioeconomic status and source of diagnosis.

Data analysis

Incident LD cohort

The descriptive evaluation of the incident LD cohort will include

      1) the patients' baseline characteristics and

      2) the crude incidence of LD per calendar year by age, gender, season, region, socioeconomic status and history of LD disease.

Overall, annual and stratified incidence rates and incidence rate differences between the different strata will be estimated based on the Poisson distribution. Only the first recording of LD will be counted.

Patient group involvement

Kate Bloor from the Lyme Research Group in the UK has provided valuable suggestions for this study.

Study limitations

Identification of LD it appears that many GPs do not enter LD diagnoses as Read codes in particular when LD was diagnosed in secondary care. In such case the LD diagnosed would only be contained in the unstructured medical notes. Distinguishing the positive LD diagnoses from comments on suspected LD or dismissal of LD is not without error. A conservative approach regarding the free text will be taken by only including potential LD cases from the medical notes which were associated with a prescription of a predefined antibacterial medication on the same day or associated with a Read code indicating an encounter with a secondary care physician. In addition, patients with a non-negative Borrelia test associated with a prescription of a predefined antibacterial medication on the same day will be included.

Some patients receive a diagnosis of LD from health professionals other than their GP. These may be entered as Read codes or simply in the medical notes, or not entered. As noted above, such diagnoses mentioned in the medical notes will only be included when associated with a Read code indicating the encounter with a third party. Diagnoses of LD entered in HES will be evaluated to assess the sensitivity of the case finding approach, i.e. by calculating the proportion of HES cases also recorded in GPRD, whether with a Read code or in the medical notes.

Dissemination of results

It is planned to publish the results, if possible in a peer-reviewed journal read by GPs.

References

(1) WHO Report 'Lyme borreliosis in Europe: influences of climate and climate change, epidemiology, ecology and adaptation measures' by: Elisabet Lindgren and Thomas G.T. Jaenson (2006).
http://www.euro.who.int/__data/assets/pdf_file/0006/96819/E89522.pdf
(2) European Commission Health and Consumers Directorate-General (2006):
http://ec.europa.eu/health/ph_information/dissemination/echi/docs/lyme_en.pdf
(3) Health Protection Scotland: http://www.documents.hps.scot.nhs.uk/giz/10-year-tables/lyme.pdf
(4) Robert Koch-Institut (2010). "Lyme-borreliose: Analyse der gemeldeten Erkrankungsfaelle der Jahre 2007 bis 2009 aus den sechs oestlichen Bundeslaendern"[Lyme borreliosis: Analysis of reported cases in the years 2007-2009 in six eastern German counties], Epidemiol Bulletin 12:101–110

(5) Poggensee G, Fingerle V, Hunfeld KP et al (2008). "Lyme-Borreliose: Forschungsbedarf und Forschungsansaetze. Ergebnisse eines interdisziplinaren Expertentreffens am Robert Koch-Institut. [Lyme borreliosis: research gaps and research approaches. Results from an interdisciplinary expert meeting at the Robert Koch institute]", Bundesgesundheitsblatt Gesundheitsforschung - Gesundheitsschutz 51(11):1329–1339

(6) Stanek G, Fingerle V, Hunfeld KP, et al. (2011). Lyme borreliosis: Clinical case definitions for diagnosis and management in Europe. Clin Microbiol Infect 17: 69–79

(7) British Infection Association (2011). The epidemiology, prevention, investigation and treatment of Lyme borreliosis in United Kingdom patients: A position statement by the British Infection Association. Journal of Infection (2011) 62, 329-338

(8) Public Health England: www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/LymeDisease/EpidemiologicalData/lymLymeepidemiology/

(9) Mueller I, Freitag MH, Poggensee G et al (2012). 'Evaluating Frequency, Diagnostic Quality, and Cost of Lyme Borreliosis Testing in Germany: A Retrospective Model Analysis.' Clinical and Developmental Immunology.

(10) Kuehn BM (2013). CDC estimates 300000 US cases of Lyme disease annually. JAMA 310(11):1110

(11) Cairns V and Goodwin J (2005). Post-Lyme borreliosis syndrome: a meta-analysis of reported symptoms. International Journal of Epidemiology 34:1340–1345