BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([http://bmjopen.bmj.com/site/about/resources/checklist.pdf](http://bmjopen.bmj.com/site/about/resources/checklist.pdf)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

**ARTICLE DETAILS**

<table>
<thead>
<tr>
<th>TITLE (PROVISIONAL)</th>
<th>The incidence of Lyme disease in the UK, a population-based cohort study</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTHORS</td>
<td>Cairns, Victoria; Wallenhorst, Christopher; Rietbrock, Stephan; Martinez, Carlos</td>
</tr>
</tbody>
</table>

**VERSION 1 - REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Kenneth D Mandl&lt;br&gt;Boston Children's Hospital, United States</th>
</tr>
</thead>
<tbody>
<tr>
<td>REVIEW RETURNED</td>
<td>15-Sep-2018</td>
</tr>
</tbody>
</table>

**GENERAL COMMENTS**

Cairns and co-authors have performed an important measurement—an estimated Lyme disease incidence in the UK. They do so using the Clinical Practice Research Datalink (CPRD), a primary care database covering about 8% of the population in the UK.

The main finding is that that the number of cases of LD has increased rapidly over the years 2001 to 2012.

The authors took great care to define the cases using two independent reviewers. Importantly, cases with disagreement were adjudicated between them. Nonetheless, it would be of interest to know the interrater reliability of their assessments.

One method of identifying Lyme disease cases was to use recorded medical codes, notes indicating LD, laboratory tests and use of specific antibiotics. However, the completed codes and antibiotics list for identifying LD cases are not included in the manuscript (it's only in the study protocol). The authors should describe the identification method in detail.

Further, with regards to the case definitions, the term "definite LD" may be misleading as it does not require laboratory confirmation. Further it is unclear if the ECM found in the free text or code was required to be on the day of diagnosis, or could be an ECM-like rash as recalled by the patient. This is an important distinction, as Lyme can be diagnosed clinically with a confirmed ECM rash.

I find it somewhat surprising that given the richness of the data in the CPRD that the specifics of the laboratory testing are not accounted for. How many LD cases received standard 2-tiered LD
lab test? What are the results among all the LD cases which underwent lab testing?

In the Introduction, the authors cited reference 1 to describe that the US Centers for Disease Control and Prevention estimate around 300,000 new cases of LD per year in the US. Authors should consider citing their published papers, rather than the website.

The authors defined a single LD episode using a 365-day gap. Early Lyme disease can be cured within 4 weeks. A 365 days gap is much more than 4 weeks. The authors should explain or discuss the reason they choose 365 days to define an LD episode.

In the Results Section, the authors report 21.8% of LD cases with a record of ECM. The authors should discuss the relatively low percentage of ECM in their cohort.

It is not clear to me how population incidences were calculated by extrapolating from the CPRD. The authors state “The number of LD cases was estimated by region as the observed incidence rate times the population in the region. Some of the healthcare regions were combined to ensure a sufficiently large population for yearly estimates by region.” But (1) what fraction of the population in each region is covered by the primary care practitioners enrolled in the CPRD?; (2) were there changes in GP participation in the CPRD participation over the study period that could affect the measurements?

The authors should consider showing a map to illustrate geographic variability in the incidence of the disease.

REVIEWER
Kiersten Kugeler
US CDC

REVIEW RETURNED
18-Sep-2018

GENERAL COMMENTS
The authors present an analysis of estimated LD incidence in the UK based on a seemingly robust primary care database. The methods seem like they are appropriate yet lack detail on some specific areas. First, the authors do not provide sufficient evidence to support the assertion that this data source is representative of the UK population. The validity of the conclusions hinge upon this, yet characteristics of this cohort are not compared to those of the UK population as a whole. Second, there is missing detail in how LD cases were classified both in terms of LD status, but also laboratory testing, and antibiotics. Some specific comments follow:
1. In Introduction, please spell out World Health Organization prior to first use of WHO.
2. I don’t follow the logic in the introduction that because of an estimate of >200,000 cases in Germany, LD in the UK is underestimated. These are two different places with no doubt different ecology and population size. Why would you assume that risk is similar between these two countries without any additional justification accordingly?
3. Please describe more about laboratory testing for LD in the UK so this manuscript is broadly applicable. For example, what does
“non-negative” actually mean? Are these single tier or two tier tests?
4. How is one determined to be “definite LD” based on free text?
   It’s mentioned several times, but it’s unclear how reading only a chunk of medical notes, this assumption could be made. For example, what if the provider wrote “this person could have Lyme disease, but it could also be X, Y, Z”—how would that be classified? I’d be concerned that “rule out” comments could contribute to overinclusion here.
5. Please add a few more sentences that describe the 4038 persons with LD in your cohort—you have some detail in Table 1, but it would be nice if it was mentioned in text too.
6. Please add more detail to Table 1 to reflect LD “status”/diagnostic methods
7. The sex and age distribution of the persons with LD in this cohort is different than in the US, with which I am most familiar. Is this a result of bias in the data source or is this the known age and sex-related breakdown in LD in the UK? Regardless, it should be mentioned in the discussion.
8. Please consider adding sensitivity analyses results to results section
9. First sentence of summary—please clearly state this is the first in the UK.
10. Why are there no UK-based authors that currently work in academia or public health? It’s a bit odd, as I am not sure how you’d be able to draw any meaningful conclusions or next steps otherwise. Consequently, there are gaps in the “so what” of this manuscript—in the lack of discussion regarding implications for the increasing incidence of LD in the UK and what it means for clinicians, public health officials, and the public.

GENERAL COMMENTS
This is a well written and well described paper that highlights a world-wide trend of increasing diagnoses of Lyme disease.

The authors accurately record that the number of Lyme diagnoses in the UK is increasing and that this increase is spread throughout the UK, which I find to be scientifically very plausible.

There are three main areas that if further explored would lead to a better understanding of the paper:
1. The authors correctly point out that a negative LD serological test, particularly early in the illness, is not definitive evidence against LD and that LD remains to a significant extent a clinical diagnosis. However, if the data were available for patients in the groups of “Definite LD” and “Treated suspected LD” as to how many had serology performed and what the positive rate was, as a general guide to the accuracy of clinical diagnosis of the (in general) non Lyme-expert general practitioner cohort, it would be helpful. The authors do point out that increasing awareness of Lyme among the general population and among health practitioners may have played a role in an increase in diagnosis, but do not discuss that a proportion of this is inevitably over-
diagnosis or misdiagnosis. The proportion which is over-diagnosed or misdiagnosed is unclear but could be informed (though not totally resolved) by comprehensive laboratory data.

2. Related to point 1, above, the description of group 1 as ‘Definite LD’ is perhaps slightly misleading, and I would prefer a name such as ‘Clinically diagnosed LD’, as I would normally reserve an absolute such as ‘definite’ for those patients with laboratory and clinical evidence of LD.

3. The actual number of patients with two reported episodes of LD is unclear (54 or 56) and the description of these cases is insufficient for proper understanding. It is possible that patients may fit definition 1 or 2 without clinical suspicion for a second episode of acute Lyme disease, or in fact may have had relapsed LD, or a diagnosis of ‘chronic Lyme disease’ which remains an uncertain clinical entity. Without epidemiological or laboratory data to support the suggestion, e.g., a high risk profession or hobby, it seems unlikely that of 4083 cases of acute LD, 56 were re-infections. However, it should be noted that even if these cases were excluded from analysis, the authors’ conclusions about rising LD diagnosis, and the importance of prevention and recognition, remain valid.

**VERSION 1 – AUTHOR RESPONSE**

Reviewer 1:

1.1: The authors took great care to define the cases using two independent reviewers. Importantly, cases with disagreement were adjudicated between them. Nonetheless, it would be of interest to know the interrater reliability of their assessments.

The two independent reviewers of the medical notes came to the same conclusion in the large majority of cases, and reached agreement on the remaining 0.8% of cases after reviewing all information available in the database and after some discussion.

1.2: One method of identifying Lyme disease cases was to use recorded medical codes, notes indicating LD, laboratory tests and use of specific antibiotics. However, the completed codes and antibiotics list for identifying LD cases are not included in the manuscript (it’s only in the study protocol). The authors should describe the identification method in detail.

Thank you for your comment. We have clarified this in the “Study cohort and definition of LD” section of the methods. A new table with the full list of codes used to identify diagnoses of LD/ECM and specific laboratory tests has now been added as a supplementary table. It has also been clarified that laboratory LD tests comprised the type of LD laboratory test and/or a qualitative test result.

Furthermore, the list of recommended antibiotics for LD that has been used in the LD algorithm has been added in the section “Study cohort and definition of LD”. This comprises “amoxicillin, azithromycin, cefotaxime, ceftriaxone, cefuroxime, doxycycline and penicillin G”. The identification methods for LD events, including how the provided code and substance lists have been applied, are already available in the section “Study cohort and definition of LD” and in Figure 1.

Supplementary table: Read medical codes used for definition of LD events

<table>
<thead>
<tr>
<th>Read code</th>
<th>Code description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diagnoses</td>
<td></td>
</tr>
</tbody>
</table>
1.3: Further, with regards to the case definitions, the term "definite LD" may be misleading as it does not require laboratory confirmation.

Thank you for your comment. We have replaced the term "definite LD" by "clinically confirmed LD" throughout the manuscript.

1.4: Further it is unclear if the ECM found in the free text or code was required to be on the day of diagnosis, or could be an ECM-like rash as recalled by the patient. This is an important distinction, as Lyme can be diagnosed clinically with a confirmed ECM rash.

Thank you for this point. The cases with an ECM were identified from a Read medical code for an ECM entered in the database, or from an ECM mentioned in the medical notes. The day of the ECM recording was considered the day of the LD diagnosis, also in the absence of an LD diagnosis code. A patient recall of ECM noted by the GP in the medical notes would only form a LD event if it led to an LD or ECM diagnosis by the GP ("clinically confirmed LD") or if it led to the GP starting treatment with recommended antibiotics ("treated suspected LD").

We have added further information to the section "Study cohort and definition of LD" as follows:

"The patients with LD were identified from an algorithm based on medical codes for LD, erythema chronicum migrans (ECM), laboratory tests, and anonymized medical notes. The day of a recording was considered the date of the occurrence of the respective event."

1.5: I find it somewhat surprising that given the richness of the data in the CPRD that the specifics of the laboratory testing are not accounted for. How many LD cases received standard 2-tiered LD lab test? What are the results among all the LD cases which underwent lab testing?

The CPRD does include details on the type of many laboratory tests. However, details on the exact type of laboratory test used for Lyme disease and the specific results beyond positive, equivocal or negative are not available. Tests without information on the qualitative test result were assumed to be tests with an unspecified result. In the section "Study cohort and definition of LD" of the methods, we have clarified it that information on LD laboratory tests was based on Read medical codes and
comprised the type of LD laboratory test and/or a qualitative test result. A table with the available LD laboratory test Read medical codes has been referenced and is included as a supplementary table.

1.6: In the Introduction, the authors cited reference 1 to describe that the US Centers for Disease Control and Prevention estimate around 300,000 new cases of LD per year in the US. Authors should consider citing their published papers, rather than the website.

Thank you for your suggestion. The two published papers on the incidence of Lyme disease in the US(Hinckley, Connally et al. 2014, Nelson, Saha et al. 2015) have now been referenced and the website reference was removed.

1.7: The authors defined a single LD episode using a 365-day gap. Early Lyme disease can be cured within 4 weeks. A 365 days gap is much more than 4 weeks. The authors should explain or discuss the reason they choose 365 days to define an LD episode.

Some patients can have ongoing symptoms long after being infected. We wanted to ensure that a single infection was not counted more than once and so took a conservative approach (see also last paragraph in discussion section “Strengths and limitations”). This has been clarified in the methods as follows:

“To ensure that a single infection was not counted more than once, multiple episodes of LD for any one patient were only counted if there was a gap of more than 365 days between consecutive recordings.”

1.8: In the Results Section, the authors report 21.8% of LD cases with a record of ECM. The authors should discuss the relatively low percentage of ECM in their cohort.

The proportion of 22% of cases with an ECM in our study is actually not relatively low. It is similar to the 25% of cases with an ECM in the original study by Steere et al in 1977 in which Lyme disease was first identified.(Steere, Malawista et al. 1977) That study is the only Lyme disease study in which having an ECM was clearly not one of the inclusion criteria. Therefore, although it was not a large study, that study appears to provide the most reliable estimate of the proportion of Lyme disease cases with an ECM. In the discussion, we have now added that reference and we mention their 25% of cases with an ECM directly after giving the percentage with an ECM in our study.

1.9: It is not clear to me how population incidences were calculated by extrapolating from the CPRD. The authors state "The number of LD cases was estimated by region as the observed incidence rate times the population in the region. Some of the healthcare regions were combined to ensure a sufficiently large population for yearly estimates by region." But (1) what fraction of the population in each region is covered by the primary care practitioners enrolled in the CPRD?; (2) were there changes in GP participation in the CPRD participation over the study period that could affect the measurements?

Overall about 8% of all GP practices in the UK are registered in the CPRD, but this will vary slightly across regions. GPs remain registered in the CPRD over many years, and so it is not expected that there will have been any changes in GP participation that could affect the measurements.

The CPRD contains the total number of people registered in each participating practice and so the incidence rate for each region could be calculated as the total number of Lyme disease cases that were identified divided by the total number of person-years of the patients registered in the participating practices in that region.

Then the observed incidence rate in the registered GP practices in that region was multiplied by the total number of inhabitants (and not by the number of patients registered in the CPRD) in that region to estimate the total number of cases in that region.
The methods section has been modified to clarify this point as follows:

"Annual regional incidence rates with 95% confidence intervals (based on the Poisson distribution) were estimated from the observed number of LD cases in a region divided by the total person-years of the patients registered in the participating practices in that region. The number of LD cases in a region was then estimated as the observed incidence rate times the total number of inhabitants in that region."

1.10: The authors should consider showing a map to illustrate geographic variability in the incidence of the disease.

Thank you for your suggestion. A new figure has now been added with a map showing the incidence of LD in different regions across the UK, as was suggested. It has been observed that the incidence of LD may not be uniform within a region, for example in areas of Scotland, England and Wales. (Smith, O'Connell et al. 2000, Mavin, Watson et al. 2015) The incidence of LD may vary, for example due to variations in the prevalence of ticks infected with genospecies of B. burgdorferi, differences in people’s occupations and leisure activities, and differences in the climate and vegetation. Therefore, we have added a footnote to our new figure saying that "Incidence rates may vary across different parts of a region".

Figure 2: Incidence rates of Lyme disease per 100,000 per year by region, 2010-2012

Incidence rates may vary across different parts of a region

Reviewer 2:

2.1: First, the authors do not provide sufficient evidence to support the assertion that this data source is representative of the UK population. The validity of the conclusions hinge upon this, yet characteristics of this cohort are not compared to those of the UK population as a whole.

Thank you for your comment. When compared with the UK census in 2011, CPRD patients are broadly representative of the UK population in terms of age, sex, ethnicity and body mass index. The
CPRD may not be exactly representative of all practices in the UK based on geography. Therefore, overall UK LD estimates have been standardized for region. A source providing information on the representativeness of the CPRD in the UK is already included in section “Data source”. (Herrett, Gallagher et al. 2015)

2.2 Second, there is missing detail in how LD cases were classified both in terms of LD status, but also laboratory testing, and antibiotics.

Please see our responses to comments 2.5 and 2.6.

2.3: In Introduction, please spell out World Health Organization prior to first use of WHO.

Thank you for the suggestion. The name of the WHO has now been written out in full, rather than abbreviated.

2.4: I don’t follow the logic in the introduction that because of an estimate of >200,000 cases in Germany, LD in the UK is underestimated. These are two different places with no doubt different ecology and population size. Why would you assume that risk is similar between these two countries without any additional justification accordingly?

We agree about the comment on the number of cases seen in Germany. We have removed that statement, and have clarified this point in the introduction as follows:

"The higher incidence rates seen in some neighbouring countries suggest a potential underestimation of the incidence rate and number of cases in the UK. (Sykes and Makiello 2017)"

2.5: Please describe more about laboratory testing for LD in the UK so this manuscript is broadly applicable. For example, what does “non-negative” actually mean? Are these single tier or two tier tests?

Thank you for this point. The CPRD does not include exact details on the type of laboratory test used for Lyme disease, or on the results beyond positive, equivocal or negative. We have clarified this in the methods by inserting a new supplementary table with all Read medical codes used to identify LD tests. We have also replaced the term “non-negative” by “positive, equivocal or unspecified”.

2.6: How is one determined to be “definite LD” based on free text? It’s mentioned several times, but it’s unclear how reading only a chunk of medical notes, this assumption could be made. For example, what if the provider wrote “this person could have Lyme disease, but it could also be X, Y, Z”—how would that be classified? I’d be concerned that “rule out” comments could contribute to over inclusion here.

Thank you for your comment. The information in the medical notes was grouped by the medical reviewers into the categories “definite LD”, “suspected LD”, “no LD” or “insufficient information”. Medical notes with insufficient information or assessed as "no LD" were not counted as Lyme disease cases. Medical notes that listed LD as one of several potential diagnoses, and medical notes indicating that the patient may have had LD but without a confirmation by the GP were assessed as "no LD". Medical notes assessed as “definite LD” had a clear statement of diagnosis by the GP in the free text but were only counted as a “definite LD” event when a medical code for LD or ECM, a secondary care encounter or a prescription for recommended antibiotics was recorded on the same day. Medical notes assessed as “suspected LD” had a statement by the GP that Lyme disease was suspected and were only counted as a "treated suspected LD" event, if these patients also had a prescription for recommended antibiotics on the same day. Any suspected cases that were ruled out by the GP would not have received any antibiotics or seen a secondary care doctor, and so they would not have been counted in our study.
We have replaced the term "definite LD" by "clinically confirmed LD".

2.7: Please add a few more sentences that describe the 4038 persons with LD in your cohort—you have some detail in Table 1, but it would be nice if it was mentioned in text too.

Thank you for your suggestion. The results section has been expanded to summarise some of the details presented in Table 1, as follows:

"Nearly one quarter of all cases were aged under thirty, 53.2% were female, and half of the diagnoses were made in the summer months (Table 1)."

2.8: Please add more detail to Table 1 to reflect LD "status"/diagnostic methods

Although details on the LD classification are available in Figure 1, we have added further details on the number of LD cases identified in each of the diagnostic categories in Table 1, as was suggested.

<table>
<thead>
<tr>
<th>Diagnostic category of Lyme disease</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinically confirmed</td>
<td>1702 (41.7)</td>
</tr>
<tr>
<td>Treated suspected</td>
<td>1913 (46.9)</td>
</tr>
<tr>
<td>Treated possible</td>
<td>468 (11.5)</td>
</tr>
</tbody>
</table>

2.9: The sex and age distribution of the persons with LD in this cohort is different than in the US, with which I am most familiar. Is this a result of bias in the data source or is this the known age and sex-related breakdown in LD in the UK? Regardless, it should be mentioned in the discussion.

The age distribution of the cases in our study is very similar to that seen in the Lyme disease surveillance study in England and Wales. (Smith, O'Connell et al. 2000)

To enable a direct comparison by age group we have categorized age here below using the same age categories as by Smith et al. The age categories do not differ by more than 2 percentage points in each age category.

<table>
<thead>
<tr>
<th>Age (years)</th>
<th>Our study (%)</th>
<th>Smith et al. 2000 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;15</td>
<td>10.2</td>
<td>10.6</td>
</tr>
<tr>
<td>15-24</td>
<td>7.0</td>
<td>5.7</td>
</tr>
<tr>
<td>25-44</td>
<td>29.4</td>
<td>27.6</td>
</tr>
<tr>
<td>45-64</td>
<td>36.9</td>
<td>37.7</td>
</tr>
<tr>
<td>65+</td>
<td>16.7</td>
<td>16.2</td>
</tr>
</tbody>
</table>

We have now addressed this point in the discussion section as follows:

"The age distribution of the cases in this study is very similar to that seen in the LD surveillance study in England and Wales. (Smith, O'Connell et al. 2000) The sex distribution of the cases is also consistent with that seen in other European studies. (Hubalek 2009) In contrast to the USA, where nearly 57% of LD cases are male, (Schwartz, Hinckley et al. 2017) there is a slight preponderance of females in Europe, ranging from 54 to 60%. (Hubalek 2009) which is consistent with the 53.2% seen in our study. Any differences between countries in the observed incidence rate and in the age and sex distribution of LD cases could be due to differences in the distribution of ticks, in the proportion of ticks infected with LD, and also to differences in people’s occupations and leisure activities, and the public awareness of LD. (Nelson, Saha et al. 2015)"

2.10. Please consider adding sensitivity analyses results to results section

As was suggested, information on the results from the sensitivity analyses has been added at the end of the results section as follows:
"Sensitivity analyses indicated that allowing a time window of ±30 days for the antibiotic prescription around the date of coding of a laboratory test would have increased the number of cases by 13.9% in our study cohort and by 14.0% in our overall UK estimate. The age, sex, season and region distribution of LD cases in the sensitivity analyses was consistent with their distribution in the main analysis."

2.11. First sentence of summary—please clearly state this is the first in the UK.

That it is the first such study in the UK has been added in the summary. We have also added at the end of the introduction a statement that the guideline on LD written by the British National Institute for Health and Care Excellence (NICE) stresses the need for an epidemiological study on LD in the UK, and we have added the reference. (National Institute for Clinical Excellence 2018)

2.12: Why are there no UK-based authors that currently work in academia or public health? It’s a bit odd, as I am not sure how you’d be able to draw any meaningful conclusions or next steps otherwise. Consequently, there are gaps in the “so what” of this manuscript—in the lack of discussion regarding implications for the increasing incidence of LD in the UK and what it means for clinicians, public health officials, and the public.

The first author lives in the UK but is retired and so has no affiliation with any organisation. She has already published on Lyme disease in the International Journal of Epidemiology in 2005. Her paper was a meta-analysis of what is now known as post-treatment Lyme disease syndrome. The other authors have published more than 10 peer reviewed articles based on data from the CPRD. We have experience in publishing medical statistics, and our lack of affiliation with any UK institution is not important with regard to this publication.

Reviewer 3:

3.1: The authors correctly point out that a negative LD serological test, particularly early in the illness, is not definitive evidence against LD and that LD remains to a significant extent a clinical diagnosis. However, if the data were available for patients in the groups of 'Definite LD' and 'Treated suspected LD' as to how many had serology performed and what the positive rate was, as a general guide to the accuracy of clinical diagnosis of the (in general) non Lyme-expert general practitioner cohort, it would be helpful. The authors do point out that increasing awareness of Lyme among the general population and among health practitioners may have played a role in an increase in diagnosis, but do not discuss that a proportion of this is inevitably over-diagnosis or misdiagnosis. The proportion which is over-diagnosed or misdiagnosed is unclear but could be informed (though not totally resolved) by comprehensive laboratory data.

The reasons for a clinical diagnosis of Lyme disease vary. Patients in this study with a record of definite LD or suspected LD in the medical notes did not have a record of a laboratory test result on the same day, and they were only counted in this study if they were given antibiotics on the same day, or if they had a record of a secondary care encounter. Patients who had a laboratory test and met the study criteria for a diagnosis were counted separately from those identified from the medical notes. Some other patients with medical notes indicative of Lyme disease may have had a laboratory test on a different day, but this was not evaluated, as it was not part of the study criteria.

The CPRD does not include details on the type of laboratory test used for Lyme disease, or on the specific results beyond positive, equivocal or negative. A new supplementary table has been added with details on the Read medical codes that were used to identify LD tests. We have added information in the manuscript that information on the tests was qualitative and based on Read medical codes and medical notes only. Based on the data available in CPRD, it is not possible to say what proportion of Lyme disease cases are over or under diagnosed.
We have clarified this point by adding information in the Study cohort and definition of LD section of the methods tests were qualitative and based on Read medical codes and medical notes only, as follows:

“The algorithm identified patients with a medical code for a laboratory LD test (Supplementary table), comprising the type of LD laboratory test and/or a qualitative test result, separated into those recorded with a positive result, an equivocal result, or an unspecified result, together with a prescription for a recommended LD antibiotic recorded by the GP on the same day.”

3.2: Related to point 1, above, the description of group 1 as ‘Definite LD’ is perhaps slightly misleading, and I would prefer a name such as ‘Clinically diagnosed LD’, as I would normally reserve an absolute such as ‘definite’ for those patients with laboratory and clinical evidence of LD.

Thank you for this point. The term "definite LD" has been re-phrased as "clinically diagnosed LD", as was suggested.

3.3: The actual number of patients with two reported episodes of LD is unclear (54 or 56) and the description of these cases is insufficient for proper understanding. It is possible that patients may fit definition 1 or 2 without clinical suspicion for a second episode of acute Lyme disease, or in fact may have had relapsed LD, or a diagnosis of ‘chronic Lyme disease’ which remains an uncertain clinical entity. Without epidemiological or laboratory data to support the suggestion, e.g., a high risk profession or hobby, it seems unlikely that of 4083 cases of acute LD, 56 were re-infections. However, it should be noted that even if these cases were excluded from analysis, the authors’ conclusions about rising LD diagnosis, and the importance of prevention and recognition, remain valid.

A conservative approach was taken in order not to count patients with long term problems as multiple cases. Therefore, a patient had to be free of any previous LD recording for >365 days to be counted as having a re-infection according to our algorithm. Patients with very long term problems following a single infection are not expected to have had multiple separate diagnoses that met the study criteria, and so it is unlikely that such a case was counted more than once.

We have modified the text in the results section to clarify how many re-infections were observed as follows:

"Only 56 of the 4025 patients (1.4%) appeared to have had more than one LD infection. Of those, 54 patients had exactly one LD re-infection and two had exactly two re-infections, i.e. a total of 58 re-infections, based on our 365 day re-infection blocking-time-window."

References


**VERSION 2 – REVIEW**

<table>
<thead>
<tr>
<th>REVIEWER</th>
<th>Kenneth Mandl</th>
</tr>
</thead>
<tbody>
<tr>
<td>Boston Children's Hospital, and Harvard Medical School USA</td>
<td></td>
</tr>
</tbody>
</table>

| REVIEW RETURNED | 05-Dec-2018 |

<table>
<thead>
<tr>
<th>GENERAL COMMENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dr. Yi Ju Tseng of Chang-Gung University assisted in the review.</td>
</tr>
<tr>
<td>The authors have responded thoroughly to the comments. Thank you. We have a few remaining comments.</td>
</tr>
</tbody>
</table>

1. The authors clearly describe the case identification with a figure and the disagreement rate of medical notes review. The interrater reliability of their assessments is high. It should be reported in the text of the manuscript.

2. It is reasonable that the authors chose a 365-day period to define an LD episode. This ensures that a single infection was not counted more than once, and is a conservative approach. However, most LD cases can be treated within 4 weeks. How many LD diagnoses were recorded, per patient, within 365 days in the CPRD? It would be useful to understand the temporal patterns to know whether some episodes several months apart could represent distinct infections.

3. Reference 16 from Public Health England provides annual numbers and rates of positive laboratory diagnoses of LD in England and Wales. The authors only discussed the case numbers but the rates should be also compared and discussed.
## GENERAL COMMENTS

Overall, I was disappointed in this revision. I thought all three reviewers provided some very insightful questions of the authors, most of which generated no ensuing change to the manuscript to address those questions. Simply replying to reviewers with an answer without acknowledging the source of their question doesn’t answer the underlying improvements needed in the manuscript. A few examples of such insufficiencies include: lack of interrater reliability data in the manuscript and lack of information of the generalizability of the data source. Regarding the former, a reviewer asked for this information and a loose statement that the reviewers “came to the same conclusion in the large majority of cases” is not only vague, but no detail was added to the manuscript itself. Secondly, stating that the dataset is “broadly generalizable” is not supported by any evidence, despite a reviewer request to provide some data to support the statement. Other concerns include: the authors changed “definite LD” to “clinically confirmed LD” which is a term that is not routinely used in Lyme disease research. Moreover, the authors informed the third reviewer that the text change was to “clinically diagnosed LD”. Although this latter term makes more intuitive sense, this is not at all how the text was modified, exemplifying a lack of attention to detail.

### VERSION 2 – AUTHOR RESPONSE

**Reviewer 1 (Kenneth Mandl):**

The authors have responded thoroughly to the comments. Thank you. We have a few remaining comments.

1.1: The authors clearly describe the case identification with a figure and the disagreement rate of medical notes review. The interrater reliability of their assessments is high. It should be reported in the text of the manuscript.

We have provided information on the interrater variability and on how it was derived by adding the following in the methods section, based on an amended result from a more detailed evaluation:

The free text of all potential cases from the criteria (ii) “patients with clinically confirmed LD or ECM in the medical notes together with a medical code indicating a secondary care visit or referral to secondary care, and (iii) patients with mentioning of LD or ECM in the medical notes (separated into clinically confirmed LD or LD suspected by the GP) together with a prescription for a recommended LD antibiotic (amoxicillin, azithromycin, cefotaxime, ceftriaxone, cefuroxime, doxycycline or penicillin G) recorded by the GP on the same day” was reviewed manually by two reviewers (both MDs) independently using a reiterative approach consisting of three steps. First, the reviewers allocated the medical notes into (i) clinically diagnosed LD or erythema chronicum migrans (ECM), (ii) suspected LD/ECM or an LD-specific laboratory test, and (iii) no evidence of LD/ECM or insufficient information.
In a second step, all medical notes with a non-concordant assessment were reviewed again by the same reviewers independently and without knowledge of the previous assessment. This procedure was repeated a third time. This resulted in a total concordance of 99.4%. The discrepancies in the categories to which events were allocated had been evaluated until a final assessment was agreed. Potential cases were allocated to ‘clinically diagnosed LD’, ‘suspected LD’, or ‘no LD diagnosis or insufficient information’.

The methods section of the manuscript was expanded as follows:

The free text of all potential cases from the criteria (ii) and (iii) above was reviewed manually by two reviewers (both MDs) independently using a reiterative approach consisting of three steps. First, the reviewers allocated the medical notes into (i) clinically diagnosed LD or erythema chronicum migrans (ECM), (ii) suspected LD/ECM or an LD-specific laboratory test, and (iii) no evidence of LD/ECM or insufficient information. In a second step, all medical notes with a non-concordant assessment were reviewed again by the same reviewers independently and without knowledge of the previous assessment. This procedure was repeated a third time. This resulted in a total concordance of 99.4%. The discrepancies in the categories to which events were allocated had been evaluated until a final assessment was agreed. Potential cases were allocated to ‘clinically diagnosed LD’, ‘suspected LD’, or ‘no LD diagnosis or insufficient information’.

1.2: It is reasonable that the authors chose a 365-day period to define an LD episode. This ensures that a single infection was not counted more than once, and is a conservative approach. However, most LD cases can be treated within 4 weeks. How many LD diagnoses were recorded, per patient, within 365 days in the CPRD? It would be useful to understand the temporal patterns to know whether some episodes several months apart could represent distinct infections.

We agree to include more information about the multiple recordings of LD seen in our study, and the pattern of those recordings.

We have therefore added the following in the new version of our manuscript at the end of the results section:

"There were 437 patients (10.9%) who had at least one additional recording of LD between 29 and 365 days after the initial LD diagnosis. The median time between the initial LD recording and the first subsequent LD recording was 50 days. 25% of the patients had a subsequent recording between 29 and up to 38 days after the initial LD recording. Another 25% of the patients had a subsequent recording at least 84 days after the initial LD recording.

Of the 437 patients who had a subsequent recording of LD within 365 days, only 2 (0.5%) received intravenous ceftriaxone treatment at the time of their subsequent recording. Overall, 14 (0.3%) of the patients in this study had a record of intravenous antibiotic treatment, which is consistent with treatment for late stage LD."

The following text has been added to the study limitations:

"If it is caught early, most cases of LD can be treated successfully within four weeks. It is therefore possible that some of the 437 patients with a second LD recording between 29 and 365 days later may have had successful treatment, and then a second LD infection. Sensitivity analyses showed that using a time window of 28 instead of 365 days for concatenation of LD episodes, would have resulted in an additional 230 calculated cases in the UK in the year 2012, on top of the calculated 7738, which reflects a potential underestimation of the incidence of LD of up to 3%."
In the study that we cite by Public Health England in 2011, they write that over 1000 serologically confirmed infections are reported annually in the UK and that it has been estimated that there may also be between 1000 and 2000 unconfirmed cases per year. Sources were provided with estimates based on laboratory data in England and Wales, and centralized reporting in Scotland, but these are different diagnostic categories, and we do not have any comparable data from other studies with the number of cases by region. There is a lack of epidemiological data on Lyme disease in the UK, and so we have not been able to make any changes to our manuscript based on this suggestion.

Reviewer 2 (Kiersten Kugeler):

Overall, I was disappointed in this revision. I thought all three reviewers provided some very insightful questions of the authors, most of which generated no ensuing change to the manuscript to address those questions. Simply replying to reviewers with an answer without acknowledging the source of their question doesn’t answer the underlying improvements needed in the manuscript.

We are sorry that the reviewer felt that we had not added enough clarification in the manuscript to answer the reviewers’ questions. We have now added some statements to cover the remaining issues.

2.1: A few examples of such insufficiencies include: lack of interrater reliability data in the manuscript and lack of information of the generalizability of the data source. Regarding the former, a reviewer asked for this information and a loose statement that the reviewers "came to the same conclusion in the large majority of cases" is not only vague, but no detail was added to the manuscript itself.

We have provided information on the interrater variability and on how it was derived by adding the following in the methods section, based on an amended result from a more detailed evaluation:

The free text of all potential cases from the criteria (ii) "patients with clinically confirmed LD or ECM in the medical notes together with a medical code indicating a secondary care visit or referral to secondary care, and (iii) patients with mentioning of LD or ECM in the medical notes (separated into clinically confirmed LD or LD suspected by the GP) together with a prescription for a recommended LD antibiotic (amoxicillin, azithromycin, cefotaxime, ceftriaxone, cefuroxime, doxycycline or penicillin G) recorded by the GP on the same day" was reviewed manually by two reviewers (both MDs) independently using a reiterative approach consisting of three steps. First, the reviewers allocated the medical notes into (i) clinically diagnosed LD or erythema chronicum migrans (ECM), (ii) suspected LD/ECM or an LD-specific laboratory test, and (iii) no evidence of LD/ECM or insufficient information. In a second step, all medical notes with a non-concordant assessment were reviewed again by the same reviewers independently and without knowledge of the previous assessment. This procedure was repeated a third time. This resulted in a total concordance of 99.4%. The discrepancies in the categories to which events were allocated had been evaluated until a final assessment was agreed. Potential cases were allocated to ‘clinically diagnosed LD’, ‘suspected LD’, or ‘no LD diagnosis or insufficient information’.

The methods section of the manuscript was expanded as follows:

The free text of all potential cases from the criteria (ii) and (iii) above was reviewed manually by two reviewers (both MDs) independently using a reiterative approach consisting of three steps. First, the reviewers allocated the medical notes into (i) clinically diagnosed LD or erythema chronicum migrans (ECM), (ii) suspected LD/ECM or an LD-specific laboratory test, and (iii) no evidence of LD/ECM or insufficient information. In a second step, all medical notes with a non-concordant assessment were reviewed again by the same reviewers independently and without knowledge of the previous assessment. This procedure was repeated a third time. This resulted in a total concordance of 99.4%. The discrepancies in the categories to which events were allocated had been evaluated until a final
assessment was agreed. Potential cases were allocated to ‘clinically diagnosed LD’, ‘suspected LD’, or ‘no LD diagnosis or insufficient information’.

2.2: Secondly, stating that the dataset is "broadly generalizable" is not supported by any evidence, despite a reviewer request to provide some data to support the statement.

With regard to the generalizability of the database, we have expanded our statement in the data source section as follows, based on information given in the reference that we cite:

"Over 98% of the UK population are registered with a primary care general practitioner, and a subset of the GP practices participate in the CPRD linkage scheme and provide patient-level information. Approximately 8% of the UK population is currently included in the CPRD.

When compared with the UK census in 2011, CPRD patients were found to be broadly representative of the UK population in terms of age and sex, and also comparable in terms of ethnicity.(Herrett, Gallagher et al. 2015)"

The following text has been added to the study limitations:

"The CPRD may not be representative of all practices in the UK based on geography and size,(Herrett, Gallagher et al. 2015) but this should not have affected our estimates, which are based on rates calculated by region from the participating GP practices."

2.3: Other concerns include: the authors changed "definite LD" to "clinically confirmed LD" which is a term that is not routinely used in Lyme disease research. Moreover, the authors informed the third reviewer that the text change was to "clinically diagnosed LD". Although this latter term makes more intuitive sense, this is not at all how the text was modified, exemplifying a lack of attention to detail.

We had an error in our response to reviewer number 3, in which we said that "definite LD" was re-phrased as "clinically diagnosed LD". Unfortunately, the word diagnosed was not later corrected in that statement, although, in the version of our manuscript that we submitted, the text had been changed throughout from "definite LD" to "clinically confirmed LD", as we intended. However, we now agree with reviewer number 2, who says that she thinks "clinically diagnosed LD" makes more intuitive sense, and so we have changed the text in the manuscript again, this time to "clinically diagnosed LD" throughout.

**GENERAL COMMENTS**

1. The authors clearly described the method of clinical note review and reported the interrater reliability in the manuscript.
2. The authors have included an analysis of LD visits patterns. About 50% of LD patients had a subsequent recording at least 12 weeks after the initial LD recording. The sensitivity analysis shows potential underestimation of the incidence of LD.