Highlights

- Thirteen percent of patients met criteria for post-treatment Lyme disease
- Persistent fatigue and pain occurred despite early diagnosis and prompt treatment
- Differences were significant after controlling for co-morbidities and other factors
Risk of post-treatment Lyme disease in patients with ideally-treated early Lyme disease: A prospective cohort study.

Author List
John N. Aucott MD1*, Ting Yang PhD1, Isaac Yoon MD2, Debra Powell MD3, Steven A. Geller MD4, Alison W. Rebman MPH1

Affiliations
1Lyme Disease Research Center, Division of Rheumatology, Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, MD, USA
2Patient First, Lutherville, MD USA
3Reading Hospital Tower Health, West Reading, PA, USA
4Centennial Medical Group, Elkridge, MD, USA

*Corresponding author
2360 W Joppa Road, Suite 320
Lutherville, MD 21093 (USA)
jaucott2@jhmi.edu
410-616-7596
ORCID: 0000-0002-2648-0896
Abstract

Purpose

Post-treatment Lyme disease (PTLD) is characterized by patient-reported symptoms following treatment for *B. burgdorferi* infection. The primary aim of this study was to assess whether participants with a history of Lyme disease (LD) would be more likely to meet criteria for PTLD than those without.

Methods

We conducted a longitudinal, prospective study among 234 participants with and 49 participants without prior LD. All completed survey metrics for fatigue, pain, sleep, depression, and quality of life. An operationalized PTLD definition was applied to both cohorts, and the distributions of clinical outcomes and symptoms were examined.

Results

13·7% with a history of prior LD met criteria for PTLD compared to 4·1% of those without. Participants with prior LD were approximately 5·28 times as likely to meet PTLD criteria compared to those without (*p*=0·042) and had 2-3 times as high odds of reporting moderate or severe fatigue and muscle pain (*p*=0·002, 0·047, respectively). Risk of meeting PTLD criteria was also independently increased among females and those with higher exposure to prior traumatic life events.

Conclusion
Participants ideally diagnosed and treated for prior LD reported more symptoms on standardized surveys and were more likely to meet criteria for PTLD than those without.

**Keywords**
Lyme Disease, Post-Treatment Lyme Disease, Symptom Measurement

**Introduction**
Lyme disease (LD) is a vector-borne disease caused by various genospecies of the spirochetal bacteria *Borrelia burgdorferi* sensu lato (Steere et al., 2016). When untreated, the infection can cause specific early and/or late dermatologic, cardiac, neurologic, and joint manifestations (Steere et al., 2016). Following appropriate antibiotic treatment, LD has also been associated with persistent symptoms in a subset of patients, particularly fatigue, musculoskeletal pain, and cognitive difficulties, which can impact quality of life (Klempner et al., 2001; Rebman et al., 2017b). The association of infectious disease with a proximate chronic condition is not unique to LD (Carapetis et al., 2005; Hickie et al., 2006; Nalbandian et al., 2021).

In the United States and Europe, a small number of longitudinal studies have examined the prevalence and severity of persistent symptoms following treatment for LD, (Aucott et al., 2013b; Smith et al., 2002; Weitzner et al., 2015; Wills et al., 2016) with much of this data coming from early clinical trials (Dattwyler et al., 1997; Luft et al., 1996; Luger et al., 1995; Robert et al., 1992; Ursinus et al., 2021; Wormser et al., 2003). Variability in treatment paradigms, study design, enrollment criteria, and especially outcome classification have led to a
wide range in estimated prevalence of persistent symptoms in this population (0-50%, however 10-20% is frequently cited) (Marques, 2008). Even fewer prospective cohort studies have included controls unexposed to LD so as to compare symptom prevalence to previously uninfected populations. In both our prior work (Bechtold et al., 2017) and in a recent publication by Wormser et al., (Wormser et al., 2020) overall symptom reporting among patients as a whole months to years following treatment for early LD did not differ from unexposed controls. However, also in both studies, a patient subset could be identified as having post-treatment Lyme disease (PTLD), either through standardized symptom surveys (Bechtold et al., 2017) or medical record review by a clinician (Wormser et al., 2020). A recent European study of PTLD showed that in well-defined treated LD patients, persistent symptoms were significantly more prevalent and severe than in patients without LD (Ursinus et al., 2021).

The Infectious Diseases Society of America (IDSA) previously proposed a research case definition for studying patients with PTLD, a distinct subgroup of patients with persistent symptoms who may present with a clinical diagnosis of chronic LD (Feder Jr. et al., 2007; Rebman et al., 2017b; Wormser et al., 2006). Patients who meet criteria for PTLD have prior documented LD, do not have other specific co-morbidities, and have both specific symptoms and functional impact of symptoms at least 6 months from their initial LD diagnosis and treatment. In previous publications, we have operationalized this IDSA research case definition through use of survey instruments and applied it to patients with a history of LD (Aucott et al., 2013a; Bechtold et al., 2017).
In the current study, we applied these PTLD-defining symptom and functional impact criteria to two cohorts of participants followed prospectively; those exposed and those unexposed to early LD. The primary aim of this study was to assess whether those with a history of LD would be more likely to meet these criteria than those without. This represents the first time that such differences have been examined using standardized criteria for PTLD in a large US sample.

Materials and Methods

Study Participants

For this longitudinal, prospective cohort study, adult patients (≥ 18 years) with early LD self-referred or were recruited from primary or urgent care settings in 2008-2020. Eligible participants were primarily enrolled at study sites in Maryland with a small number enrolled at a satellite site in southeastern Pennsylvania. At enrollment, participants were required to have a visible (erythema migrans) EM ≥ 5 cm in diameter diagnosed by a health care provider and to have ≤ 72 hours of appropriate antibiotic treatment for early LD. Additional details and baseline clinical characteristics of this sample have been previously published (Rebman et al., 2021).

Participants without a clinical or serologic history of LD were recruited from similar primary care settings, through community recruitment using flyers, and online advertising. This cohort was required to be two-tier seronegative for antibodies to *B. burgdorferi* at the time of enrollment and at all subsequent visits, as well as be free of any history of prior clinical LD. All participants in both groups were excluded for the same range of self-reported prior medical conditions, (Wormser et al., 2006) specifically: chronic fatigue syndrome, fibromyalgia, unexplained chronic pain, sleep apnea or narcolepsy, autoimmune disease, chronic neurologic
disease, liver disease, HIV, cancer or malignancy in the past 2 years, major psychiatric illness, or drug or alcohol abuse. We did not exclude participants from either group based on medical conditions not referenced in the IDSA’s PTLD proposed case definition.

The Institutional Review Board of the Johns Hopkins University School of Medicine approved this study, and all participants signed written consent prior to initiation of study activities.

Data Collection

At the baseline study visit, all participants self-reported demographic characteristics and prior physician-diagnosed co-morbidities. Participants also completed the Life Events Checklist (LEC), (Gray et al., 2004) a measure of exposure to potentially traumatic events. At each study visit, all participants self-administered a series of symptom-based surveys, including the Post-Lyme Questionnaire of Symptoms (PLQS), a 36-item survey developed based on prior clinical and research experience among patients with PTLD. For each symptom, participants indicated presence and severity (“absent” [0], “mild” [1], “moderate” [2], or “severe” [3]) over the past 2 weeks.

Additionally, all participants completed the following: the Fatigue Severity Scale (FSS), a 9-item fatigue metric with summary scores ranging from 9 to 63; (Krupp et al., 1989) the Short-Form McGill Pain Questionnaire (SF-MPQ), a 15-item measure of pain with summary scores ranging from 0 to 45; (Melzack, 1975) the Pittsburgh Sleep Quality Index (PSQI) sleep metric with
summary scores ranging from 0 to 21; (Buysse et al., 1989) and the Beck Depression Inventory-II (BDI-II), a 21-item depression metric which can be separated into somatic and affective subscales (Beck et al., 1996; Storch et al., 2004). For these measures, higher scores indicate a worse experience of symptoms. Finally, functional impact was measured by the Short-Form Health Survey, Version 2 (SF-36). This 36-item quality of life measure has 8 norm-based subscale scores as well as Physical and Mental Component summary scores (PCS and MCS, respectively) which can also be compared with the US population mean (50 ± 10) (Ware et al., 2000). For this measure, a higher score indicates better life functioning.

**Outcome Definition**
Outcome data for participants with prior LD were drawn from two follow-up study visits conducted approximately 6 months and 1 year following the end of appropriate antibiotic treatment for their early LD. Similarly, data from participants without prior LD were extracted from two study visits occurring 6 months apart.

At each visit, participants were considered to meet criteria for ‘symptoms’ if fatigue, pain, and/or cognitive complaints were present at the ‘moderate’ or ‘severe’ level on the PLQs. Similarly, participants were considered to meet criteria for ‘functional impact’ if an average composite score of 4 specific norm-based subscales on the SF-36 was 0.5 SD below the population mean, as previously identified (Aucott et al., 2013a).

At each visit, we categorized all participants into 3 outcome groups based on whether they met these ‘symptoms’/‘functional impact’ criteria. Those meeting criteria for both symptoms and
functional impact met our operationalized case definition for PTLD (‘PTLD’ group). Those meeting criteria for symptoms but not functional impact were considered a ‘Symptoms Only’ group. The remaining participants were considered ‘Return to Health.’ Final status determination across the two visits was hierarchically determined. Participants meeting PTLD criteria at either visit were determined to be in this group. Among remaining participants, those meeting Symptoms Only criteria at either visit were determined to be in this group. All remaining participants were considered Return to Health, including the small (4·2%) proportion meeting criteria at either visit for functional impact but not symptoms.

To assess symptom differences by outcome status, PLOS, FSS, SF-MPQ, PSQI, and BDI-II scores were averaged across the two time points.

Statistical Analyses
To test the hypothesis that participants exposed to prior, treated LD would be more likely to meet criteria for “PTLD” than unexposed control participants, we first tested for statistical difference between groups using Fisher’s exact test. We then fit a multinomial logistic regression model with symptoms/functional impact status as the outcome variable and participant cohort as the primary exposure. Relevant demographic, psychosocial, and co-morbidity variables were compared between participant groups, and both $p<0.15$ and clinical relevance were used to select potential confounders for inclusion in the adjusted model.

Further analyses then explored specific symptom differences by participant cohort. First, we presented the distribution of scores for the PLQS, FSS, SF-MPQ, PSQI, and BDI-II. Next, these
scores, as well as a subset of core PTLD symptoms chosen based on review of the literature by a clinical expert (JA) from the PLQS, were compared between participant groups in adjusted analyses. We fit individual linear regression models with each score as the outcome and participant group as the primary exposure, adjusting for potential confounders. Finally, we conducted analyses using individual logistic regression models to compare the odds of ‘moderate’ or ‘severe’ symptom reporting between participant groups for specific PLQS symptoms.

A \( p < 0.05 \) was considered statistically significant. All analyses were performed using R (version 4.0.2).

**Results**

A total of 293 participants with and 54 without prior LD were enrolled in the study. Of those enrolled at the time of early LD diagnosis, 87% and 83% returned for their 6-month and 1-year follow-up visit, respectively. Ninety-three percent of those without prior LD completed both time points. A final sample of 234 participants with and 49 without prior LD were available for analysis, as described in Table 1.

Although not statistically significant, participants with LD had a higher median age, were a lower percentage female, and had a lower number of years of education (Table 1). A significantly higher percentage of participants with prior LD were white, non-Hispanic. The two groups were similar in LEC score. Among self-reported co-morbidities at the baseline visit, those present in \( \geq 5\% \) of the sample were: thyroid disease, heart disease/hypertension, migraine headaches, and carpal tunnel syndrome. Although none were statistically significant by group,
participants with prior LD reported a higher percentage of heart disease/hypertension. There was an overall statistically significant difference in the distribution of symptoms/functional impact status by participant group ($p = 0.038$), with a significantly higher percentage of participants with prior LD that met criteria for PTLD (13.7% vs. 4.1%). A total of 43.1% of participants with prior LD met criteria for either PTLD or Symptoms Only compared to 24.5% without a history of prior LD ($p=0.016$).

We then fit a multinomial logistic regression model with outcome status as the response and patient group as the primary predictor, adjusting for sex, race, years of education, LEC, and heart disease/hypertension (Table 2). Although no significant differences in LEC score were found between groups, it was included as we considered it potentially relevant to outcome status. We found that participants with prior LD were approximately 5.28 times as likely to meet PTLD than Return to Health criteria compared to those without ($p = 0.042$). We also found that females were 4.17 times as likely as males to meet PTLD than Return to Health criteria ($p=0.001$), and that for each 1-unit increase in LEC score, reflecting an additional potentially traumatic life event, the risk of meeting PTLD criteria increased 30% ($p=0.003$).

Figure 1 shows the distribution of average PLQS scores, as well as the percentage reporting each symptom at an average moderate or severe level by group. Descriptively, participants with prior LD had a higher severity of many symptoms. Among the 10 symptoms we selected, on average participants with prior LD had scores approximately 0.2 to 0.3 points higher than those without in joint pain, memory changes and depression ($p=0.006, 0.027, \text{and } 0.046$, respectively).
adjusting for sex and LEC (Table 3). They also had 2 to 3 times as high odds of developing moderate or severe fatigue and muscle pain ($p=0.002$, and $0.047$, respectively).

The distributions of the scores for the standardized self-administered questionnaires are presented in Figure 2. Individual linear regression models adjusting for sex and LEC showed that participants with prior LD had significantly worse scores on the SF-MPQ and the SF-36 PCS compared to those without (Table 4, $p=0.006$ for both).

**Discussion**

In our prospective study, participants with history of early LD were more likely to meet symptom and functional impact criteria for PTLD than those without. These differences remained significant in multivariate analyses controlling for relevant co-morbidity, demographic, and psychosocial factors. This is important because it represents the first time that such differences have been examined and found using the standardized IDSA criteria for PTLD in a large US sample.

Studying PTLD is challenging, primarily due to its symptom-based phenotype. Without specific physical findings or practical laboratory biomarkers linked to a defined pathophysiology, PTLD diagnosis, progression, and response to treatment remain vague. Even *B. burgdorferi* antibody levels are not specific or sensitive in the post-treatment period, and are not included in the IDSA proposed case definition (Wormser et al., 2006). Investigators are therefore left to interpret and operationalize ‘symptoms’ and ‘functional impact’ in study-specific ways. These are often the circumstances for other infection-associated conditions as well, including recently described
“post-acute” or “long” COVID following SARS-CoV-2 infection, (Nalbandian et al., 2021) which anecdotally shares a similar clinical phenotype to PTLD. While others have relied on physician evaluation, (Wormser et al., 2020) in the current study and in our prior publication, (Bechtold et al., 2017) operationalization of PTLD criteria was performed using surveys. Using the instruments detailed above among a Mid-Atlantic cohort of participants, we found that PTLD criteria can be fulfilled in approximately 14% with history of prior LD compared to approximately 4% of those without. The 14% is within the 10-20% range of estimated persistent symptom prevalence frequently cited in the literature (Marques, 2008). It is also similar to the 20% recently reported by Wormser and colleagues based on physician assessment and review of records (Wormser et al., 2020). While these estimates indicate that the majority of patients promptly treated for early LD do generally return to their pre-morbid health, it also suggests that PTLD is not necessarily rare, nor an inconsequential public health concern. Especially considering ongoing tick habitat expansion, (Sonenshine, 2018) LD now affects an estimated 476,000 patients annually in the United States (Kugeler et al., 2021). Furthermore, although the majority of patients may not progress to PTLD, this should not invalidate the experience of the subset of those patients who do. The focus of longitudinal studies, including the current one, has historically been to examine whether samples of patients with LD as a whole report symptoms over time, regardless of clinical outcome. This is a different research aim from understanding the clinical phenotype and life impact among the subgroup of LD patients with PTLD, which can be significant (Klempner et al., 2001; Rebman et al., 2017b, 2017a). Indeed, the distribution of SF-36 PCS and MCS scores was generally higher than the US population mean among our sample
of participants with prior LD, despite the statistically significant difference in SF-36 PCS scores between groups.

The most frequently identified PTLD symptoms are fatigue, pain, and cognitive dysfunction (Rebman et al., 2017b; Wormser et al., 2006). These symptoms are often non-specific and are often reported in general clinical populations as well as in patients with prior LD. In the past, this has led to the conclusion that symptoms in PTLD are likely representative of an expected background “noise” of daily life (Wormser et al., 2006). These statements have historically relied upon extrapolating symptom frequencies across sample populations and symptom measures never intended nor designed to be comparable, potentially masking severity differences and risking false conclusions. In the current study, 24% of participants without a history of prior LD reported these symptoms at a moderate or severe level (our estimate of symptom background “noise”), which is statistically significant compared to 43% with prior LD ($p=0.016$). On both our PLQS symptom survey as well as more broadly-utilized standardized instruments, specific cohort differences appeared most prominently in these known PTLD symptoms.

The current study is also the first to examine and control for the impact of potential confounders on the relationship between cohort group and PTLD. Sex, race, and education were included in the initial model, given statistically significant or borderline significant demographic differences between the cohort groups. Furthermore, as the presence of co-morbid conditions has been found to be associated with increased symptom reporting in LD, (Wills et al., 2016) we also controlled for heart disease/hypertension, the only condition present in $\geq 5\%$ of the sample which appeared to differ by group. Lastly, mental health conditions and prior exposure to traumatic life events
have previously been found and/or theorized to be associated with persistent symptoms following LD (Sigal and Hassett, 2005; Solomon et al., 1998; Wills et al., 2016; Wormser et al., 2006). Therefore, we included LEC score, a measure of potential prior traumatic life events, in multivariate models, and we specifically examined depression both in our PLQS survey and through the use of the BDI-II.

Notably, we found that participants with prior LD remained over 5 times as likely to meet PTLD than Return to Health criteria compared to those without after controlling for these other factors, and there was a statistical trend in both the descriptive and multivariate analyses for these groups to differ in the risk of Symptoms Only as well. We also found that female sex and a higher LEC score remained significant predictors in multivariate models, indicating that these factors are independently associated with meeting PTLD criteria. The reasons for these findings are unclear. It is possible that female sex and/or increased exposure to traumatic life events impact the initial biologic response to *B. burgdorferi* infection and increase risk for persistent symptoms (Klein and Flanagan, 2016; Seiler et al., 2020). However, as these factors remained significant after controlling for cohort group, this supports an association with increased symptoms and functional impact unrelated to prior LD exposure (Barsky et al., 2001; McFarlane, 2010). Finally, although cohort group did predict PLQS depression score, it did not predict moderate/severe level for this variable, nor did it predict BDI-II or SF-36 MCS score. This is consistent with our previous publications that mild or moderate depression is likely not a primary driver of increased symptom reporting in PTLD, at least among patients who are promptly diagnosed and treated (Bechtold et al., 2017).
This study does have limitations, in particular the smaller number of control participants. We accounted for this as best we could by: a) generally recruiting both cohorts from similar underlying geographic and clinical populations, and b) controlling statistically for demographic and clinical factors observed to vary by group. Additionally, in order to increase specificity of our outcomes, and to mirror the IDSA proposed case definition, we excluded participants for a range of conditions with significant symptom overlap with PTLD. Nonetheless, patients with these conditions can be exposed to *B. burgdorferi* infection as well, therefore future studies will be needed to assess generalizability of our findings to more clinically and immunologically diverse patient populations. Furthermore, our study of ideally-diagnosed and treated participants with early LD limits generalizability to those with risk factors for PTLD, namely delayed diagnosis and treatment. Lastly, a small number of participants (9 cases and 3 controls) met criteria for functional impact but not symptoms, and were characterized as Return to Health. To be certain that this did not influence interpretation of our results however, we conducted a sensitivity analysis with these 12 participants removed, and it did not considerably change our reported results.

In this US Mid-Atlantic prospective cohort study, participants ideally diagnosed and treated for prior Lyme disease reported more symptoms on standardized surveys and were more likely to meet an operationalized definition of PTLD than those without, even after controlling for relevant confounding factors. These findings contribute to a growing literature with the common goal of better understanding and characterizing PTLD.
Contributions

All authors contributed to the study conception and design. Funding acquisition was performed by JNA. Data collection was performed and supervised by JNA, IY, DP, SAG, and AWR. Data curation and statistical analyses were performed by TY and AWR. The first draft of the manuscript was written by AWR and TY, and all authors commented on previous versions of the manuscript. All authors verify the underlying data, and have read and approved the final manuscript.

We are grateful to the research participants who contributed their time and effort towards this study. We would also like to thank Erica Mihm and Cheryl Novak for their efforts in patient recruitment and conducting study visits.

Funding Source

This work was supported by the Steven and Alexandra Cohen Foundation; the Global Lyme Alliance (GLA); and the Bay Area Lyme Foundation (BALF). This publication was also made possible by the Johns Hopkins Institute for Clinical and Translational Research (ICTR), which is funded in part by Grant Number UL1 TR003098 from the National Center for Advancing Translational Sciences (NCATS) a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research, and the Johns Hopkins Clinical Research Network (JHCRN). Its contents are solely the responsibility of the authors and do not necessarily represent the official view of the Johns Hopkins ICTR, JHCRN, NCATS, NIH, the Steven and Alexandra Cohen Foundation, GLA, or BALF. The funders did not have any role in study design; collection, analysis, and interpretation of data; writing of this manuscript; or the decision to submit for publication.
**Ethical Approval**

We have read the policy of the journal on ethical consent and the standards of animal care. This work has been approved by the Johns Hopkins Institutional Review Board and complies with its human subject research informed consent regulations. No animals were a part to this research.

**Conflicts of Interest/Competing Interests**

The following financial relationships have been disclosed by the authors, all are not directly related to the submitted work. JNA received grants for work on Lyme disease from the Steven and Alexandra Cohen Foundation, the Global Lyme Alliance, and the Bay Area Lyme Foundation; support for the present manuscript from the Steven and Alexandra Cohen Foundation; received consulting fees from Pfizer North America Lyme Disease Vaccine Ad Board and Tarsus Pharmaceuticals Inc, Lyme Disease Advisory Board; received payment for expert testimony in a 2021 plaintiff malpractice case for Lyme disease; serves on the Bay Area Lyme Disease Scientific Advisory Board; is a Past Chair, 2018, HHS Tick-borne Disease Working Group, Office of HIV/AIDS and Infectious Disease Policy, Office of the Assistant Secretary of Health, Department of Health and Human Services; and has the following patent – John Aucott, MD, Inventor, Title: Elevated CCL19 after Completion of Therapy for Acute Lyme disease Identifies Patients at Risk for Development of Post-treatment Lyme disease syndrome who will benefit from further antibiotic therapy. Patent approved US patent application (16/058,316). JNA, TY, and AWR currently receive salary support from the Steven and Alexandra Cohen Foundation. IY, DP, and SAG have nothing to declare.
The authors attest that they have not been paid to write this article by a pharmaceutical company or other agency. All authors had full access to the data in the study and accept responsibility to submit for publication.

*Availability of data and material*

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

*Code availability*

The code generated for analyses during the current study is available from the corresponding author on reasonable request.

*Consent for publication*

Not Applicable
References

Aucott J, Crowder L, Kortte K. Development of a foundation for a case definition of post-treatment Lyme disease syndrome. Int J Infect Dis 2013a;17:e443-9. https://doi.org/10.1016/j.ijid.2013.01.008.

Aucott J, Rebman A, Crowder L, Kortte K. Post-treatment Lyme disease syndrome symptomatology and the impact on life functioning: Is there something here? Qual Life Res 2013b;22. https://doi.org/10.1007/s11136-012-0126-6.

Barsky AJ, Peekna HM, Borus JF. Somatic symptom reporting in women and men. J Gen Intern Med 2001;16:266–26675. https://doi.org/10.1046/j.1525-1497.2001.00229.x.

Bechtold KT, Rebman AW, Crowder LA, Johnson-Greene D, Aucott JN. Standardized symptom measurement of individuals with early Lyme disease over time. Arch Clin Neuropsychol 2017;32. https://doi.org/10.1093/arclin/acw098.

Beck AT, Steer RA, Brown GK. Beck Depression Inventory - Second Edition Manual. San Antonio: The Psychological Corporation; 1996.

Buysse DJ, Reynolds 3rd CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213. https://doi.org/10.1016/0165-1781(89)90047-4.

Carapetis JR, Steer AC, Mulholland EK, Weber M. The global burden of group A streptococcal diseases. Lancet Infect Dis 2005;5:685–94. https://doi.org/10.1016/S1473-3099(05)70267-X.

Dattwyler RJ, Luft BJ, Kunkel MJ, Finkel MF, Wormser GP, Rush TJ, et al. Ceftriaxone compared with doxycycline for the treatment of acute disseminated Lyme disease. N Engl J Med
1997;337:289–94. https://doi.org/10.1056/NEJM199707313370501.

Feder Jr. HM, Johnson BJ, O’Connell S, Shapiro ED, Steere AC, Wormser GP, et al. A critical appraisal of “chronic Lyme disease.” N Engl J Med 2007;357:1422–30. https://doi.org/10.1056/NEJMra072023.

Gray MJ, Litz BT, Hsu JL, Lombardo TW. Psychometric properties of the life events checklist. Assessment 2004;11:330–41. https://doi.org/10.1177/1073191104269954.

Hickie I, Davenport T, Wakefield D, Vollmer-Conna U, Cameron B, Vernon SD, et al. Post-infective and chronic fatigue syndromes precipitated by viral and non-viral pathogens: prospective cohort study. BMJ 2006;333:575. https://doi.org/10.1136/bmj.38933.585764.AE.

Klein SL, Flanagan KL. Sex differences in immune responses. Nat Rev Immunol 2016;16:626–38. https://doi.org/10.1038/nri.2016.90.

Klempner MS, Hu LT, Evans J, Schmid CH, Johnson GM, Trevino RP, et al. Two controlled trials of antibiotic treatment in patients with persistent symptoms and a history of Lyme disease. N Engl J Med 2001;345:85–92. https://doi.org/10.1056/NEJM200107123450202.

Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol 1989;46:1121–3. https://doi.org/10.1001/archneur.1989.00520460115022.

Kugeler KJ, Schwartz AM, Delorey MJ, Mead PS, Hinckley AF. Estimating the frequency of Lyme disease diagnoses, United States, 2010-2018. Emerg Infect Dis 2021;27:616–9. https://doi.org/10.3201/eid2702.202731.

Luft BJ, Dattwyler RJ, Johnson RC, Luger SW, Bosler EM, Rahn DW, et al. Azithromycin
compared with amoxicillin in the treatment of erythema migrans. A double-blind, randomized, controlled trial. Ann Intern Med 1996;124:785–91.

Luger SW, Paparone P, Wormser GP, Nadelman RB, Grunwaldt E, Gomez G, et al. Comparison of cefuroxime axetil and doxycycline in treatment of patients with early Lyme disease associated with erythema migrans. Antimicrob Agents Chemother 1995;39:661–7.

Marques A. Chronic Lyme disease: a review. Infect Dis Clin North Am 2008;22:341–60, vii–viii. https://doi.org/10.1016/j.idc.2007.12.011.

McFarlane AC. The long-term costs of traumatic stress: Intertwined physical and psychological consequences. World Psychiatry 2010;9:3–10. https://doi.org/10.1002/j.2051-5545.2010.tb00254.x.

Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain 1975;1:277–99. https://doi.org/10.1016/0304-3959(75)90044-5.

Nalbandian A, Sehgal K, Gupta A, Madhavan M V., McGroder C, Stevens JS, et al. Post-acute COVID-19 syndrome. Nat Med 2021;27. https://doi.org/10.1038/s41591-021-01283-z.

Rebman A, Aucott J, Weinstein E, Bechtold K, Smith K, Leonard L. Living in limbo: contested narratives of patients with chronic symptoms following Lyme disease. Qual Health Res 2017a;27:534–46. https://doi.org/10.1177/1049732315619380.

Rebman A, Bechtold K, Yang T, Mihm E, Soloski M, Novak C, et al. The clinical, symptom, and quality-of-life characterization of a well-defined group of patients with posttreatment Lyme disease syndrome. Front Med 2017b;4:224. https://doi.org/10.3389/fmed.2017.00224.

Rebman AW, Yang T, Mihm EA, Novak CB, Yoon I, Powell D, et al. The presenting
characteristics of erythema migrans vary by age, sex, duration, and body location. Infection 2021;49:685–92. https://doi.org/10.1007/S15010-021-01590-0.

Robert B, Lugcr SW, Fratik E, Wisniewski M, Collins JJ, Wormser GP, et al. Comparison of cefuroxime axetil and doxycycline in the treatment of early Lyme disease. Ann Intern Med 1992;117:273–80.

Seiler A, Fagundes CP, Christian LM. The Impact of Everyday Stressors on the Immune System and Health. In: Choukèr A, editor. Stress Challenges Immun. Sp. From Mech. to Monit. Prev. Strateg., Cham: Springer International Publishing; 2020, p. 71–92. https://doi.org/10.1007/978-3-030-16996-1_6.

Sigal LH, Hassett AL. Commentary: “What’s in a name? That which we call a rose by any other name would smell as sweet.” Shakespeare W. Romeo and Juliet, II, ii(47-48). Int J Epidemiol 2005;34:1345–7. https://doi.org/10.1093/ije/dyi180.

Smith RP, Schoen RT, Rahn DW, Sikand VK, Nowakowski J, Parenti DL, et al. Clinical characteristics and treatment outcome of early Lyme disease in patients with microbiologically confirmed erythema migrans. Ann Intern Med 2002;136:421–8.

Solomon SP, Hilton E, Weinschel BS, Pollack S, Grolnick E. Psychological factors in the prediction of Lyme disease course. Arthritis Care Res 1998;11:419–26.

Sonenshine DE. Range expansion of tick disease vectors in North America: Implications for spread of tick-borne disease. Int J Environ Res Public Health 2018;15. https://doi.org/10.3390/ijerph15030478.

Steere AC, Strle F, Wormser GP, Hu LT, Branda JA, Hovius JW, et al. Lyme borreliosis. Nat
Rev Dis Prim 2016;2:16090. https://doi.org/10.1038/nrdp.2016.90.

Storch EA, Roberti JW, Roth DA. Factor structure, concurrent validity, and internal consistency of the Beck Depression Inventory-Second Edition in a sample of college students. Depress Anxiety 2004;19:187–9. https://doi.org/10.1002/da.20002.

Ursinus J, Vrijmoeth HD, Harms MG, Tulen AD, Knoop H, Gauw SA, et al. Prevalence of persistent symptoms after treatment for lyme borreliosis: A prospective observational cohort study. Lancet Reg Heal - Eur 2021;6:100142. https://doi.org/10.1016/j.lanepe.2021.100142.

Ware JE, Kosinski M, Dewey JE. How to score version 2 of the SF-36 health survey. Lincoln, RI: Quality Metric Inc.; 2000.

Weitzner E, McKenna D, Nowakowski J, Scavarda C, Dornbush R, Bittker S, et al. Long-term assessment of post-treatment symptoms in patients with culture-confirmed early Lyme disease. Clin Infect Dis 2015;61:1800–6. https://doi.org/10.1093/cid/civ735.

Wills AB, Spaulding AB, Adjemian J, Prevots DR, Turk SP, Williams C, et al. Long-term follow-up of patients with Lyme disease: longitudinal analysis of clinical and quality-of-life measures. Clin Infect Dis 2016;62:1546–51. https://doi.org/10.1093/cid/ciw189.

Wormser GP, Dattwyler RJ, Shapiro ED, Halperin JJ, Steere AC, Klemperner MS, et al. The clinical assessment, treatment, and prevention of Lyme disease, human granulocytic anaplasmosis, and babesiosis: clinical practice guidelines by the Infectious Diseases Society of America. Clin Infect Dis 2006;43:1089–134. https://doi.org/10.1086/508667.

Wormser GP, McKenna D, Karmen CL, Shaffer KD, Silverman JH, Nowakowski J, et al. Prospective evaluation of the frequency and severity of symptoms in Lyme disease patients with
erythema migrans compared with matched controls at baseline, 6 months, and 12 months. Clin Infect Dis 2020;71:3118–24. https://doi.org/10.1093/cid/ciz1215.

Wormser GP, Ramanathan R, Nowakowski J, McKenna D, Holmgren D, Visintainer P, et al. Duration of antibiotic therapy for early Lyme disease. A randomized, double-blind, placebo-controlled trial. Ann Intern Med 2003;138:697–704.
Legends for Figures

Figure 1. Distributions by participant group of average scores (Panel a) and the percentage reporting ‘moderate’ or ‘severe’ (Panel b) for individual symptoms included on the PLQS. Five symptoms (poor coordination, heart palpitations, urination changes, facial numbness/tingling, and double vision), were removed from Panel b, as no participants reported moderate or severe symptoms. Core post-treatment Lyme disease symptoms selected based on review of the literature by a clinical expert on the study team (JA) for further analysis are marked in red.
Figure 2. Distributions of standardized questionnaires average scores. For all questionnaires except SF-36, a higher score indicates more severe symptom reporting. For SF-36, a higher score indicates a higher quality of life. Dots are displayed at locations in proportion to the number of observations within each group. For all symptom surveys, the dotted lines represent clinically relevant cut-offs. For the SF-36, the dotted lines represent population means.
## Tables

### Table 1. Demographic, psychosocial characteristics, comorbidities, and symptoms/functional impact status by participant group.

|                               | Participants with prior early Lyme disease | Participants without prior Lyme disease | p-value<sup>a</sup> |
|-------------------------------|-------------------------------------------|----------------------------------------|---------------------|
| **DEMOGRAPHIC**              |                                           |                                        |                     |
| Age                           | 53·00 [38·00, 64·00] (20·00, 82·00)       | 50·00 [35·00, 63·00] (22·00, 73·00)    | 0·330               |
| Female Sex                    | 111 (47·4%)                               | 30 (61·2%)                             | 0·110               |
| White, Non-Hispanic           | 227 (97·0%)                               | 36 (73·5%)                             | < 0·001             |
| Years of Education            | 16·00 [16·00, 19·00] (9·00, 26·00)        | 18·00 [16·00, 19·00] (12·00, 27·00)    | 0·156               |
| Annual Household Income<sup>b</sup> |                                    |                                        | 0·510               |
| < $50,000                     | 55/219 (25·1%)                            | 9/45 (20·0%)                           |                     |
| $50,000 - $149,000            | 101/219 (46·1%)                           | 25/45 (55·6%)                          |                     |
| > $150,000                    | 63/219 (28·8%)                            | 11/45 (24·4%)                          |                     |
| **PSYCHOSOCIAL**              |                                           |                                        |                     |
| Life Events Checklist Score (LEC)<sup>c</sup> | 2·00 [1·00, 3·00] (0·00, 10·00)        | 2·00 [1·00, 3·00] (0·00, 10·00)          | 0·748               |
| **CO-MORBIDITIES**           |                                           |                                        |                     |
| Thyroid disease               | 24 (10·3%)                                | 3 (6·1%)                               | 0·592               |
| Heart disease or hypertension | 37 (15·8%)                                | 3 (6·1%)                               | 0·112               |
| Migraine headaches            | 31 (13·2%)                                | 5 (10·2%)                              | 0·646               |
| Carpal tunnel syndrome        | 15 (6·4%)                                 | 3 (6·1%)                               | 1·000               |
| **OUTCOME**                  |                                           |                                        |                     |
| Symptoms/Functional Impact Status |                                    |                                        | 0·038               |
| Return to Health              | 133 (56·8%)                               | 37 (75·5%)                             |                     |
| Symptoms Only                 | 69 (29·5%)                                | 10 (20·4%)                             | 0·117<sup>d</sup>   |
| Post-Treatment Lyme Disease   | 32 (13·7%)                                | 2 (4·1%)                               | 0·032<sup>e</sup>   |

N (%) are reported for categorical and median (25<sup>th</sup> percentile, 75<sup>th</sup> percentile) [minimum, maximum] are reported for continuous variables.

<sup>a</sup>Group comparisons were conducted using Fisher’s exact tests for categorical variables, and t tests or Wilcoxon rank sum tests for continuous variables, as appropriate.

<sup>b</sup>Nineteen participants (6.7%) missing Income data.

<sup>c</sup>Eleven participants (3.9%) missing Life Events Checklist data.

<sup>d</sup>P-value derived from comparing Symptoms Only with Return to Health.
P-value derived from comparing Post-Treatment Lyme Disease with Return to Health.

Table 2. Estimated increase in risk of meeting symptom/functional impact criteria for participants with prior early Lyme disease relative to those without prior Lyme disease, based on a multinomial logistic regression model controlling for sex, race, years of education, LEC score, and heart disease or hypertension.

| Primary Predictor                  | Symptom (+) / Functional Impact (-) |                | Symptom (+) / Functional Impact (+) |                |
|-----------------------------------|--------------------------------------|----------------|--------------------------------------|----------------|
|                                   | RRR* (95% CI)                        | p-value        | RRR* (95% CI)                        | p-value        |
| Participants with LD              | 2.11 (0.88, 5.04)                    | 0.094          | 5.28 (1.06, 26.20)                   | 0.042          |
| Potential Confounders Included in the Model |                                      |                |                                      |                |
| Female Sex                        | 1.35 (0.77, 2.38)                    | 0.295          | 4.17 (1.75, 9.92)                    | 0.001          |
| White, Non-Hispanic               | 0.80 (0.24, 2.62)                    | 0.711          | 0.64 (0.11, 3.63)                    | 0.614          |
| Years of education                | 1.02 (0.91, 1.12)                    | 0.649          | 0.94 (0.82, 1.08)                    | 0.388          |
| Life Events Checklist Score       | 1.15 (1.01, 1.30)                    | 0.039          | 1.30 (1.09, 1.53)                    | 0.003          |
| Heart disease or hypertension     | 0.36 (0.02, 3.00)                    | 0.443          | 0.80 (0.24, 2.70)                    | 0.722          |

*RRR: ratio of relative risks.
Table 3. Individual regression models with PLQS symptom score or PLQS binary score as the outcome. Participant group was included as the primary exposure, with sex and Life Events Checklist score included as potential confounders in each model.

| Symptom               | PLQS Score | PLQS Moderate or Severe (Binary) |
|-----------------------|------------|----------------------------------|
|                       | Estimated Difference (95% CI) | p-value | OR (95% CI) | p-value |
| Fatigue               | 0.09 (-0.13, 0.31) | 0.431 | 2.74 (-0.07, 5.56) | 0.002 |
| Muscle pain           | 0.08 (-0.11, 0.28) | 0.393 | 2.10 (-0.75, 4.95) | 0.047 |
| Joint pain            | 0.31 (0.09, 0.52) | 0.006 | 0.79 (-0.57, 2.15) | 0.217 |
| Difficulty finding words | 0.08 (-0.09, 0.25) | 0.337 | 1.19 (-1.59, 3.96) | 0.339 |
| Difficulty sleeping   | 0.11 (-0.07, 0.28) | 0.221 | 1.18 (-1.60, 3.95) | 0.343 |
| Difficulty focusing   | 0.09 (-0.08, 0.26) | 0.285 | 1.27 (-1.50, 4.04) | 0.292 |
| Memory changes        | 0.17 (0.02, 0.32) | 0.027 | 0.92 (-1.86, 3.70) | 0.486 |
| Neck pain             | -0.06 (-0.22, 0.10) | 0.484 | 0.24 (-1.51, 1.99) | 0.785 |
| Numbness hands/feet   | 0.07 (-0.06, 0.21) | 0.300 | 0.59 (-2.18, 3.35) | 0.672 |
| Depression            | 0.15 (0.00, 0.29) | 0.046 | 1.14 (-1.63, 3.91) | 0.360 |
Table 4. Individual regression models with standardized symptom questionnaire score as the outcome. Participant group was included as the primary exposure, with sex and Life Events Checklist score included as potential confounders in each model.

| Standardized Symptom Questionnaire | Estimated Difference (95% CI) | p-value |
|------------------------------------|-------------------------------|---------|
| Fatigue Severity Scale Total Score (FSS) | 1·08 (–2·28, 4·44) | 0·527 |
| Short-Form McGill Pain Questionnaire Total Score (SF-MPQ) | 1·51 (0·43, 2·59) | 0·006 |
| Pittsburgh Sleep Quality Index Total Score (PSQI) | 0·07 (–0·82, 0·96) | 0·881 |
| Beck Depression Inventory-II Affective Subscale Score (BDI-II Affective) | 0·57 (–0·63, 1·78) | 0·350 |
| Beck Depression Inventory-II Somatic Subscale Score (BSI-II Somatic) | 0·25 (–0·31, 0·80) | 0·388 |
| SF-36 Physical Component Score (SF-36 PCS) | –2·73 (–4·68, –0·78) | 0·006 |
| SF-36 Mental Component Score (SF-36 MCS) | –0·17 (–2·86, 2·51) | 0·900 |