Longitudinal Trajectories of Neurocognitive Functioning in Childhood Acute Lymphoblastic Leukemia

Marita Partanen,¹ PhD, Sean Phipps,² PhD, Kathryn Russell,² PhD, Doralina L. Anghelescu,² MD, Joshua Wolf,² PhD, MBBS, Heather M. Conklin,² PhD, Kevin R. Krull,² PhD, Hiroto Inaba,² MD, PhD, Ching-Hon Pui,² MD, and Lisa M. Jacola,² PhD, ABPP-CN

¹Princess Máxima Center for Pediatric Oncology and ²St. Jude Children's Research Hospital

All correspondence concerning this article should be addressed to Lisa M. Jacola, PhD, ABPP-CN, Department of Psychology, St. Jude Children's Research Hospital, 262 Danny Thomas Place, MS 740, Memphis, TN 38105, USA. E-mail: lisa.jacola@stjude.org

Received 24 May 2020; revisions received 28 August 2020; accepted 1 September 2020

Abstract

Objective Children with acute lymphoblastic leukemia (ALL) are at risk for neurocognitive deficits, and examining individual variability is essential to understand these risks. This study evaluated latent longitudinal trajectories and risk factors of neurocognitive outcomes in childhood ALL.

Methods There were 233 participants with ALL who were enrolled on a phase 3, risk-stratified chemotherapy-only clinical trial (NCT00137111) and who completed protocol-directed neurocognitive assessments [47.6% female, mean (SD) = 6.6 (3.7) years]. Measures of sustained attention, learning/memory, and parent ratings of attention were completed during and after treatment. Longitudinal latent class analyses were used to classify participants into distinct trajectories. Logistic regression was used to identify predictors of class membership. **Results** Within the overall group, attention performance was below age expectations across time (Conners Continuous Performance Test detectability/variability, $p < 0.01$); memory performance and parent ratings were below expectations at later phases (California Verbal Learning Test learning slope, $p < 0.05$; Conners Parent Rating Scale, Revised attention/learning, $p < 0.05$). Most participants (80–89%) had stable neurocognitive profiles; smaller groups showed declining (3–6%) or improving (3–11%) trajectories. Older age ($p = 0.020$), female sex ($p = 0.018$), and experiencing sepsis ($p = 0.047$) were associated with greater attention problems over time. Lower baseline IQ was associated with improved memory ($p = 0.035$) and fewer ratings of attention problems ($p = 0.013$) over time.

Conclusions Most patients with ALL have stable neurocognitive profiles. Smaller groups have significant impairments shortly after diagnosis or have worsening performance over time. A tiered assessment approach, which includes consideration of individual and clinical risk factors, may be useful for monitoring neurocognitive functioning during treatment and survivorship.

Key words: cancer and oncology; longitudinal research; neuropsychology.

Introduction

Acute lymphoblastic leukemia (ALL) is the most common childhood cancer with 5-year overall survival

rates of >90% (Pui et al., 2019). Current treatment protocols use systemic and intrathecal chemotherapy for central nervous system (CNS)-directed

prophylaxis, sparing most patients from neurocognitive changes associated with cranial irradiation (Harila et al., 2009). However, patients continue to be at risk for neurocognitive deficits, particularly in areas of attention, executive functioning, and processing speed (Iyer et al., 2015). These difficulties persist during survivorship and impact functional outcomes and quality of life (Jacola et al., 2016a; Kunin-Batson et al., 2014). Established risk factors include younger age at diagnosis and greater treatment intensity (Espy et al., 2001; Halsey et al., 2011). Female sex and lower socioeconomic status may also adversely impact neurocognitive outcomes (Espy et al., 2001; Hardy et al., 2017a; Jacola et al., 2016b). Common adverse events related to treatment, including neuropathic pain (Anghelescu et al., 2011) and severe infection (Inaba et al., 2017), may also impact neurocognitive outcomes (Cheung et al., 2018; Partanen et al., 2020); however, their contribution to neurocognitive trajectories is unknown.

There are limited longitudinal studies of neurocognitive outcomes in survivors starting during therapy and continuing into survivorship (Halsey et al., 2011; Jacola et al., 2016b; Jansen et al., 2008; Nassar et al., 2017; Sleurs et al., 2017). Results are inconsistent, likely due to differences in design (retrospective vs. prospective), cohort (clinically indicated vs. protocol-directed), study interval, and sample size. In addition, population-level statistics that are used to compare differences between groups do not capture variability among individual participants. In this regard, some patients may have stable cognitive performance after treatment and thus intervention may not be required; whereas, others show declines for whom additional services should be implemented. The ability to determine these different trajectories and the associated risk factors would allow providers to implement services in a timely manner to prevent further difficulties. An examination of these “latent” neurocognitive trajectories has not been completed in childhood ALL.

The primary purpose of this study was to examine the latent longitudinal trajectories of neurocognitive performance in childhood ALL, using prospective, protocol-driven testing completed at defined time points during and after treatment. Previous reports with this cohort showed greater attention deficits at the end of therapy (Conklin et al., 2012) and 2 years post-therapy (Jacola et al., 2016b) when compared to population norms. Furthermore, younger age at diagnosis, greater treatment intensity, and sex were associated with impaired neurocognitive outcomes. The strength of the current study is the ability to examine longitudinal patterns of performance with latent modeling. This allowed us to evaluate *individual*—rather than *group*—level changes in performance over time; the primary goal was to examine individual,

data-driven variability and risk factors associated with this variability. We hypothesized that there would be unique trajectories of neurocognitive performance and that demographic, treatment, or clinical factors would be associated with these trajectories. Ultimately, this knowledge can inform providers about the timing and nature of neurocognitive monitoring and intervention in childhood ALL.

Methods

This research was approved by the St. Jude Children’s Research Hospital’s (SJCRH) Institutional Review Board. Adult participants or legally authorized representatives provided written informed consent, and children provided assent to participation, as appropriate.

Participants

Participants diagnosed with ALL were enrolled on a frontline SJCRH Total Therapy XV protocol from 2000 to 2007 (ClinicalTrials.gov: NCT00137111). The treatment and outcomes have been described previously (Pui et al., 2009). Briefly, patients were assigned to risk groups (low, standard/high) based on clinical presentation and early treatment response. Treatment intensity was risk stratified. All patients received triple intrathecal chemotherapy with methotrexate, cytarabine, and hydrocortisone as CNS-directed therapy. The total duration of treatment was 120 weeks for girls and 146 weeks for boys. No patients received prophylactic cranial irradiation.

Participants completed protocol-directed neurocognitive testing at four time points: induction, reinduction II, end of therapy, and 2 years post-therapy (Supplementary Figure S1). Some participants were not eligible for neurocognitive assessments ($n=10$ had Down syndrome or pre-existing neurodevelopmental condition; $n=16$ did not speak English as primary language). There were 345 participants who had at least one performance measure or parent rating completed at any time point, and we included participants in the current analyses if they contributed data for at least two time points ($n=169$ performance measures; $n=233$ parent ratings). Missing data were due to various reasons (i.e., refractory or progressive disease, refusals, scheduling difficulties) and thus sample size varied by time point and measure (Supplementary Table S1).

Demographic, Medical, and Treatment Variables

Demographic, medical, and treatment variables were abstracted from patients’ medical records at SJCRH. This included sex, race, age at diagnosis, and treatment risk arm (low-risk, standard/high-risk). Episodes of bacteremia and sepsis were identified retrospectively via medical chart review, and sepsis was defined

according to consensus guidelines or requirement of urgent intervention for organ dysfunction (Cheung et al., 2018; Inaba et al., 2017; Wolf et al., 2017). Patients were classified as having: no episodes of bacteremia, at least one bacteremia episode but no sepsis, or at least one bacteremia episode with sepsis during treatment. Neuropathic pain was defined and graded using National Cancer Institute (NCI) Common Toxicity Criteria V2.0 according to whether it interfered with function or activities of daily living (Common Terminology Criteria v2.0, 1999). Neuropathic pain events were identified as adverse events prospectively as part of the Total Therapy XV protocol and were associated with specific treatment (i.e., vincristine administration) retrospectively (Anghelescu et al., 2011). Patients were classified as having zero or at least one neuropathic pain event during treatment.

Neurocognitive Measures

For the current study, we included measures that were completed at protocol-directed time points during and after completion of therapy. We focused on measures that were completed across the four time points (Supplementary Figure S1) and are commonly used in clinical and research practice. Patients ≥ 6 years old completed measures of sustained attention [Conners Continuous Performance Test (CPT)] (Conners, 1995) and verbal learning/memory [California Verbal Learning Test (CVLT)] (Delis et al., 1987, 1994). Caregivers of patients ≥ 3 years old completed rating scales that assessed concerns with attention, learning, or impulsivity-hyperactivity [Conners Parent Rating Scale, Revised (CPRS-R)] (Conners, 1997). Estimated IQ was obtained using age-appropriate Wechsler scales for patients ≥ 3 years old (WPPSI-III, WISC-III, or WAIS-III) (Wechsler, 1991, 1997, 2002). Age-standardized norms included Z-scores (attention and memory, $M = 0$, $SD = 1$; lower scores are worse); T-scores (parent ratings, $M = 50$, $SD = 10$; higher scores indicate more problems), and standard scores (baseline IQ, $M = 100$, $SD = 15$; lower scores are worse). All measures had estimates of adequate reliability and validity.

Analysis

Likelihood ratio tests and independent sample *t*-tests were used to examine differences in demographic and clinical characteristics between participants with and without longitudinal neurocognitive data (≥ 2 time points). One sample *t*-tests were used to compare performance on the neurocognitive measures to age-standardized norms for each time point.

Longitudinal latent class modeling was conducted in the *multlcm* function of the *lcmm* program in R (Proust-Lima et al., 2013). The sustained attention,

memory, and parent ratings were examined in multivariate models to evaluate similar measures together and to reduce the number of comparisons (e.g., CPT scores of detectability, variability, and reaction time were included in one multivariate model). Fixed effects included linear and quadratic terms for time since diagnosis; random effects included a linear term for time since diagnosis. We examined models that included 1–6 latent classes, and each model was run with 50 iterations to improve model fit. The best model was chosen based on lower values of Akaike information criterion (AIC) and Bayesian information criterion (BIC), entropy > 0.80 , and mean posterior probability > 0.80 for each class (Berlin et al., 2014; Nagin et al., 2018).

The derived classes were further analyzed using a two-step process to examine predictors of class membership. In the first step, differences between classes were examined using analysis of variance for continuous measures (i.e., age at diagnosis, baseline IQ) and likelihood ratio tests for categorical measures (i.e., sex, race, risk group, bacteremia group, pain group). A measure was carried forward to the second step if it was significant at $p < 0.10$. In the second step, significant measures from the first step were included as predictors of class membership using multinomial logistic regression (*brmultinom* function) (Kosmidis & Firth, 2011). All comparisons were two-tailed and were considered significant at $p < 0.05$; analyses were conducted in R-3.6.0 (R Core Team, Vienna, Austria). Results are reported according to phase name (i.e., induction, reinduction, end of therapy, 2 years post-therapy).

Results

Demographic and Clinical Information

Participants eligible for the current analysis were those who completed neurocognitive testing for two or more time points (Table 1; $n = 169$ performance measures; $n = 233$ parent ratings). Participants were younger (performance measures $p < 0.001$) and had higher baseline IQ (performance measures $p = 0.002$; ratings $p = 0.01$) than non-participants. The frequency of participants and non-participants differed by treatment risk group (standard/high-risk for performance measures: 56% vs. 44%, $p = 0.027$; standard/high-risk for ratings: 45% vs. 61%, $p = 0.006$).

Group Neurocognitive Performance

Results from cross-sectional comparisons are shown in Supplementary Table S1. The overall group performed significantly below age expectations in sustained attention across time points (CPT reaction time 2 years post, $p = 0.039$, $d = 0.19$; CPT detectability, $p < 0.05$, d range = 0.29–0.95; CPT variability,

Table I. Demographic and Clinical Information for Participants

Category	Variable	Performance measures (<i>n</i> = 169)		Parent ratings (<i>n</i> = 233)	
		<i>n</i>	%	<i>n</i>	%
Sex	Female	71	42.0	111	47.6
	Male	98	58.0	122	52.4
Race	White	131	77.5	182	78.1
	Black	32	18.9	42	18.0
	Mixed race/other	6	3.6	9	3.9
Treatment risk group	Low	74	43.8	128	54.9
	Standard/high	95	56.2	105	45.1
Bacteremia during treatment	None	127	75.1	168	72.1
	Without sepsis	28	16.6	50	21.5
	With sepsis	14	8.3	15	6.4
Neuropathic pain during treatment	No	95	56.2	134	57.5
	Yes	74	43.8	99	42.5
		M	SD	M	SD
Age at diagnosis	Years	9.1	4.1	6.6	3.7
Baseline IQ	Standard score	101.2	16.4	100.3	16.5

Note. Standard score—age-standardized norms from Wechsler scales (*M* = 100, *SD* = 15). Participants with ≥ 2 data points included in analysis.

$p < 0.05$, d range = 0.31–0.62). Learning and memory performance were significantly below age expectations at reinduction and subsequent time points (CVLT learning slope, $p < 0.05$, d range = 0.20–0.47). Parent ratings showed that participants had fewer attention and learning problems at induction (CPRS-R attention–learning, $p = 0.026$, $d = 0.21$), but significantly elevated problems at end of therapy and 2 years post-therapy (CPRS-R attention–learning, $p < 0.05$, d range = 0.16–0.17).

Latent Longitudinal Classes

Results from latent longitudinal modeling showed that the one-class model had the poorest fit (Supplementary Table S2), which suggests that one group or trajectory is not an adequate descriptor of patient performance or ratings. Rather, performance or ratings were classified into three longitudinal patterns, namely “Stable,” “Declining,” or “Improving” (Supplementary Table S3).

On the sustained attention measure, three classes showed the best model fit (CPT; AIC = 4,103.1, BIC = 4,164.6, entropy = 0.88, mean posterior probability = 0.92). Figure 1 illustrates performance over time on CPT detectability. The largest group was relatively stable with minimal declines over time [88.8%; linear (SE) = -0.61 (0.24), $p = 0.009$; Stable Class 1]. Smaller groups had better [6.3%; intercept (SE) = 6.17 (1.31), $p < 0.0001$; Declining Class 2] or similar [5.0%; intercept (SE) = -1.64 (1.09), $p = 0.13$; Declining Class 3] performance to the Stable group at baseline, but both showed declining performance over time [Class 2: linear (SE) = -2.75 (0.65), $p < 0.0001$; Class 3: linear (SE) = -4.22 (1.09), $p < 0.0001$]. Quadratic trends were shown in all groups ($p < 0.05$). In general, attention difficulties were greatest at the

end of therapy (i.e., 2.5–3 years after diagnosis), which was followed by improvement at 2 years post-therapy.

On learning and memory performance, four classes showed the best model fit (CVLT; AIC = 4,535.0, BIC = 4,618.2, entropy = 0.84, mean posterior probability = 0.92). Figure 2 illustrates performance over time on CVLT total recall. The largest group was stable over time [80.1%, linear (SE) = -0.04 (0.08), $p = 0.59$; quadratic (SE) = 0.01 (0.01), $p = 0.55$; Stable Class 1]. Other groups had poorer performance than the Stable group at baseline (intercept range = -1.03 to -4.26 , $p < 0.05$), but they had differing longitudinal trajectories. One group had improving scores [11.2%, linear (SE) = 1.20 (0.28), $p < 0.0001$; quadratic (SE) = -0.06 (0.04), $p = 0.14$; Improving Class 2], one group showed declining performance and most difficulties at end of therapy [3.1%, linear (SE) = -1.74 (0.36), $p < 0.0001$; quadratic (SE) = 0.24 (0.07), $p = 0.0002$; Declining Class 3], and the last group showed improving performance during treatment and then declining performance at 2 years post-therapy [5.6%, linear (SE) = 1.07 (0.30), $p = 0.0003$; quadratic (SE) = -0.29 (0.06), $p < 0.0001$; Declining Class 4].

Four classes had the best model fit for parent ratings of attention (CPRS-R; AIC = 9,492.7, BIC = 9,565.2, entropy = 0.88, mean posterior probability = 0.90). Figure 3 illustrates ratings over time for CPRS-R attention–learning. The largest group was stable over time [85.4%, linear (SE) = 0.04 (0.10), $p = 0.67$; quadratic (SE) = -0.01 (0.02), $p = 0.44$; Stable Class 1]. The other groups were variable, and ratings of problems were similar to or greater than expectations depending on the time point. Greatest problems were shown at: induction [3.4%, linear (SE) = 2.04 (0.69), $p = 0.0030$; quadratic (SE) = -0.61 (0.13), $p < 0.0001$; Improving Class 3], end of therapy [5.2%, linear (SE) = 4.03 (0.54), $p < 0.0001$, quadratic (SE) = -0.66 (0.10), $p < 0.0001$;

Figure 1. Sustained attention performance for longitudinal latent classes.

Note. CPT = Conners Continuous Performance Test; Z-score = lower score = worse performance ($M = 0$, $SD = 1$). CPT detectability shown as an example of longitudinal patterns. Lines illustrate quadratic change with 95% confidence interval in gray shaded area. The three classes were derived from longitudinal latent class modeling; all classes showed significant linear and quadratic changes ($p < 0.05$). Class 2 was older at diagnosis than Class 1 ($p = 0.020$); Class 3 were more likely to be female ($p = 0.018$) or have experienced sepsis during treatment ($p = 0.047$).

Figure 2. Learning and memory performance for longitudinal latent classes.

Note. CVLT: California Verbal Learning Test; Z-score = lower score = worse performance ($M = 0$, $SD = 1$). CVLT total recall shown as an example of longitudinal patterns. Lines illustrate quadratic change with 95% confidence interval in gray shaded area. The four classes were derived from longitudinal latent class modeling; Classes 2–4 showed linear changes; Classes 3 and 4 showed quadratic changes ($p < 0.05$). Class 2 had lower baseline IQ than Class 1 ($p = 0.035$); predictors of Classes 3 and 4 were not significant ($p > 0.05$).

Declining Class 4], or 2 years post-therapy [6.0%, linear (SE) = -1.20 (0.40), $p = 0.0028$, quadratic (SE) = -0.66 (0.10), $p < 0.0001$; Declining Class 2].

Predictors of Classes

On the sustained attention measure, differences between classes were shown for age at diagnosis

($F_{2,157} = 3.96$, $p = 0.021$), sex ($G_2 = 14.44$, $p = 0.0007$), and bacteremia group ($G_4 = 10.19$, $p = 0.037$). In logistic regression, patients in Declining Class 2 were older than Stable Class 1 [relative risk ratio (RR) = 1.20, $p = 0.020$; Table II], whereas Declining Class 3 were more likely to be female (RR = 0.04; $p = 0.018$) or to have experienced sepsis during treatment (RR = 6.80; $p = 0.047$) than Stable Class 1.

Figure 3. Parent ratings of attention–learning problems for longitudinal latent classes.

Note. CPRS-R: Conners Parent Rating Scale, Revised; T-score = higher score = ratings of problems ($M = 50$, $SD = 10$). CPRS-R shown as an example of longitudinal patterns. Lines illustrate quadratic change with 95% confidence interval in gray shaded area. The four classes were derived from longitudinal latent class modeling; Classes 2–4 showed significant linear and quadratic changes ($p < 0.05$). Class 3 had lower baseline IQ than Class 1 ($p = 0.013$); predictors of Classes 2 and 4 were not significant ($p > 0.05$).

Table II. Predictors of Longitudinal Latent Classes

Measure	Predictor	Class 2 versus Class 1			Class 3 versus Class 1			Class 4 versus Class 1		
		RR	SE	p^a	RR	SE	p^a	RR	SE	p^a
Sustained attention (CPT)	Age at diagnosis	1.20	0.08	0.020	1.18	0.09	0.077	–	–	–
	Male versus female	0.55	0.64	0.342	0.04	1.41	0.018	–	–	–
	Bacteremia versus no bacteremia	2.75	0.72	0.159	4.12	0.92	0.124	–	–	–
	Sepsis versus no bacteremia	0.46	1.58	0.624	6.80	0.97	0.047	–	–	–
Learning and memory (CVLT)	Age at diagnosis	0.99	0.07	0.921	1.17	0.12	0.164	1.05	0.09	0.623
	Baseline IQ	0.96	0.02	0.035	0.98	0.03	0.520	1.02	0.02	0.415
	Std/high risk versus low risk	0.43	0.60	0.158	3.13	1.47	0.439	0.44	0.77	0.285
Attention–learning problems (CPRS-R)	Age at diagnosis	0.81	0.14	0.118	1.09	0.10	0.421	0.92	0.09	0.352
	Baseline IQ	0.97	0.02	0.145	0.94	0.02	0.013	1.00	0.02	0.990

Note. CPRS-R = Conners Parent Rating Scale, Revised, CPT = Conners Continuous Performance Test, CVLT = California Verbal Learning Test; RR = relative risk ratio; SE = standard error. Exponentiated coefficients were used to calculate RRs. ^aResults from multinomial logistic regression ($p < 0.05$); (–) not applicable because only three classes identified on the sustained attention domain.

For the memory and learning measure, there were differences between classes for age at diagnosis ($F_{3,157} = 3.10$, $p = 0.028$), baseline IQ ($F_{3,144} = 2.13$, $p = 0.099$), and risk group ($G_3 = 8.63$, $p = 0.035$). Logistic regression showed that patients in Improving Class 2 had lower baseline IQ than Stable Class 1 ($RR = 0.96$; $p = 0.035$); predictors for Declining Classes 3 or 4 were not significant ($p > 0.15$). Age at diagnosis and risk group were not significant predictors across models.

On parent ratings of attention and learning, differences between classes were shown for age at diagnosis ($F_{3,229} = 2.33$, $p = 0.075$) and baseline IQ ($F_{3,180} = 2.97$, $p = 0.033$). In logistic regression, patients from Improving Class 3 had lower baseline IQ than Stable

Class 1 ($RR = 0.94$; $p = 0.013$); predictors for Declining Classes 2 or 4 were not significant ($p > 0.10$). Age at diagnosis was not a significant predictor in the logistic regression models.

Discussion

This study identified distinct trajectories of neurocognitive performance in pediatric ALL as well as unique risk factors associated with these trajectories. The majority (80–89%) of patients had stable neurocognitive skills over time, although individual performance was variable and ranged from below average to above average compared to age expectations. There were also smaller groups of patients who had variable

performance over time and deficits at different phases of treatment or survivorship. Declines in sustained attention were associated with experiencing bacteremic sepsis during treatment, female sex, or older age at diagnosis; whereas, improved memory performance and fewer attention or learning concerns were associated with lower baseline IQ. This research highlights that most patients with ALL have stable neurocognitive functioning and that impairments for smaller groups of patients may emerge at various time points. Therefore, it will be important to closely follow patients to ensure that interventions are implemented for those who have below-average performance or declining skills over time. For example, patients may need accommodations at school, one-to-one tutoring, or organizational strategies to support their attention or learning difficulties. Cognitive training (Conklin et al., 2015), aerobic exercise (Riggs et al., 2017), or pharmaceutical interventions (Conklin et al., 2010) have been examined as potential methods to improve neurocognitive functioning in childhood ALL or brain tumor survivors, although further studies are needed to determine whether these methods prevent cognitive decline before deficits are apparent.

Our results are consistent with previous studies that used group analyses, which showed that most childhood ALL survivors have intact neurocognitive performance following chemotherapy-only protocols. Several studies have implicated stable IQ (Jansen et al., 2006; Krull et al., 2013; Sleurs et al., 2017) and greater impairments in attention, working memory, executive functioning, or processing speed with increasing time from treatment (Espy et al., 2001; Jansen et al., 2008); however, there have also been contrasting results (Halsey et al., 2011; Harila et al., 2009). Also, some of the patients in the current study were included in previous longitudinal analyses, which showed that there were greater learning problems and improved attention scores from end of therapy to 2 years post-therapy, although attention performance was still within impaired ranges (Jacola et al., 2016b). Furthermore, 19 patients who experienced seizures during therapy had poorer neurocognitive performance compared to a matched sample, and they had declines in reading and fewer ratings of learning problems over time (Nassar et al., 2017). In the current study, individual trajectories were analyzed, and we included all four time points rather than two (Jacola et al., 2016b) or three (Nassar et al., 2017) time points in previous longitudinal studies. The current results provide evidence that various trajectories are possible, including stable, improving, or declining performance depending on the individual child. In addition to differences in study design, contrasting findings may be due to this variability in longitudinal trajectories. For instance, individually defined trajectories of

psychological distress have been identified in ALL during treatment (Hockenberry et al., 2017) and long-term survivorship (Brinkman et al., 2013). Our study adds to these findings by identifying individual trajectories of neurocognitive performance or behavioral ratings.

Sepsis is a life-threatening infection that requires immediate intervention; we showed that experiencing bacteremic sepsis during treatment was associated with declining attention performance. This occurred in a small number of patients but suggests that those who experience significant medical complications should be closely monitored for neurocognitive decline. Sepsis has been associated with poorer cognitive performance in long-term survivors of ALL (Cheung et al., 2018), some of whom were included in this study, as well as in general pediatric (Sidhu et al., 2015) and adult populations (Calsavara et al., 2018a). The mechanism for sepsis-related neurocognitive impairment is unknown, but could be related to glial activation, oxidative stress, or increased permeability of the blood–brain barrier (Annane & Sharshar, 2015). Factors such as premorbid depression or length of hospitalization (Calsavara et al., 2018b), or lower socioeconomic status (Als et al., 2013) may also contribute to cognitive impairment following sepsis. Prospective, longitudinal designs are needed to determine the causality between sepsis and cognitive decline. Interventions to prevent or mitigate cognitive changes, and sepsis prevention by antibiotics prophylaxis (Wolf et al., 2017) may be warranted.

We also found older age at diagnosis and female sex to be risk factors for cognitive decline. Previous reports have implicated young age as a risk factor for neurocognitive impairment (Halsey et al., 2011; Jacola et al., 2016b), which is independent from treatment risk arm (Hardy et al., 2017a). Older patients had better sustained attention performance at baseline and greater declines over time when compared to younger patients. This inconsistency could be due to representativeness from these relatively small subgroups of patients; test restrictions based on age (i.e., ≥ 6 years for performance measures); or differences in criteria for risk assignment (i.e., age ≥ 10 years at diagnosis is an NCI criterion for standard/high risk). We found that females had greater attention problems than males over time, suggesting greater susceptibility to neurocognitive weaknesses. Results from studies examining sex differences in neurocognitive functioning in survivors of ALL have been inconsistent (Baran et al., 2016; Brinkman et al., 2013; Espy et al., 2001; Halsey et al., 2011; Jacola et al., 2016b; Jansen et al., 2008). Further research will be needed to disentangle the effects of sex and medical complications. Our study showed that lower baseline IQ was associated with improved memory/learning performance and

fewer concerns of attention problems. This may relate to sample representativeness (i.e., participants had higher baseline IQ than non-participants).

The current findings show that although most patients have stable performance over time, some patients have clinically significant impairments (i.e., >1.5 SD from standardized mean) starting early in treatment. This result could be related to premorbid functioning, clinical or social-emotional factors, and highlights the importance of following individual children given the variability between and within groups. Other studies show that although *group* performance is within the average range early in treatment, *individual* children may have deficits when ranges or frequencies of impairment are considered (Brown et al., 1992; Meyers et al., 2005; Sleurs et al., 2017). Current guidelines suggest that patients who receive CNS-directed treatment should receive neuropsychological assessment starting in treatment (Annett et al., 2015). We provide additional evidence that some children treated for ALL with chemotherapy-only regimens are at risk of neurocognitive impairment, regardless of time point. Monitoring beginning during treatment using a tiered approach should be considered in this population; for example, Hardy et al. (2017b) outline a model that includes regular monitoring for all patients, targeted screening based on risk factors, and comprehensive assessments to inform remediation for existing difficulties. Monitoring batteries that are sensitive and time efficient will be necessary to follow these patients. Other studies have examined the utility of neurocognitive monitoring or screening batteries in long-term survivors of leukemia (Boulet-Craig et al., 2018; Viola et al., 2017) or various pediatric cancers (Balsamo et al., 2019; Bull et al., 2015; Krull et al., 2008), and batteries that are sensitive to impairments starting early in treatment and continuing into survivorship are an area of active research (Sands et al., 2017). Results from the current study and previous research suggest that performance measures and parent/teacher ratings of attention, executive functioning, and learning are important to include in a battery. Brief monitoring measures may be sensitive to emerging deficits, which sets the stage for earlier intervention. It is essential to consider how monitoring can be applied in settings with limited resources and in populations with primary languages other than English.

Data on school attendance during treatment and educational supports were not collected in our study. These factors may contribute to findings and should be systematically collected in future research studies. Some patients were unable to complete testing due to scheduling conflicts, illness, and age restrictions in the measures. Others were excluded based on neurodevelopmental conditions such as Down Syndrome, which could be an independent risk factor for neurocognitive

deficits (Roncadin et al., 2015). Future studies should include measures appropriate for younger children and those with premorbid cognitive vulnerability, particularly as those excluded from longitudinal analyses were younger at diagnosis and had lower baseline IQ. Differences in participation rate by treatment risk group are likely related to risk assignment criteria. Furthermore, intercepts were estimated from models for those patients who did not complete baseline testing due to reasons listed above; choosing an appropriate baseline continues to be an area of debate (Ris et al., 2017) and thus should be a consideration for future studies in ALL. Finally, we were unable to identify predictors of some group trajectories. Treatment risk group was a significant predictor of memory performance in univariate analyses, but not multinomial models, which suggests that other factors (i.e., IQ) were a stronger predictor of group membership. It is also possible that these trajectories may be predicted by depression, anxiety, stress, or family factors. Future research with large sample sizes may focus on the association of neurocognitive functioning with these factors.

This research highlights the stability of neurocognitive functioning in patients with ALL who received chemotherapy-only treatment. However, subgroups of patients have significant deficits at baseline or who show declines over time. These patients might have experienced medical complications, such as sepsis, and thus require close monitoring to ensure that they do not develop further impairments. Studies beginning during treatment and continuing into long-term survivorship are needed to clarify the trajectory of functioning over time. Monitoring batteries that are sensitive, efficient, and implemented systematically in the overall group are necessary.

Supplementary Data

Supplementary data can be found at: <https://academic.oup.com/jpepsy>.

Funding

This work was supported by the National Cancer Institute (Grants P30 CA21765; P30 GM9266) and the American Lebanese Syrian Associated Charities.

Conflicts of interest: None declared.

References

- Als, L. C., Nadel, S., Cooper, M., Pierce, C. M., Sahakian, B. J., & Garralda, M. E. (2013). Neuropsychologic function three to six months following admission to the PICU with meningoencephalitis, sepsis, and other disorders: a prospective study of school-aged children. *Critical Care*

- Medicine*, 41(4), 1094–1103. <https://doi.org/10.1097/CCM.0b013e318275d032>
- Anghelescu, D. L., Faughnan, L. G., Jeha, S., Relling, M. V., Hinds, P. S., Sandlund, J. T., Cheng, C., Pei, D., Hankins, G., Pauley, J. L., & Pui, C. H. (2011). Neuropathic pain during treatment for childhood acute lymphoblastic leukemia. *Pediatric Blood & Cancer*, 57(7), 1147–1153. <https://doi.org/10.1002/pbc.23039>
- Annane, D., & Sharshar, T. (2015). Cognitive decline after sepsis. *The Lancet Respiratory Medicine*, 3(1), 61–69. [https://doi.org/10.1016/s2213-2600\(14\)70246-2](https://doi.org/10.1016/s2213-2600(14)70246-2)
- Annett, R. D., Patel, S. K., & Phipps, S. (2015). Monitoring and assessment of neuropsychological outcomes as a standard of care in pediatric oncology. *Pediatric Blood & Cancer*, 62(S5), S460–S513. <https://doi.org/10.1002/pbc.25749>
- Balsamo, L. M., Mitchell, H. R., Ross, W., Metayer, C., Hardy, K. K., & Kadan-Lottick, N. S. (2019). Monitoring neurocognitive functioning in childhood cancer survivors: evaluation of Cogstate computerized assessment and the Behavior Rating Inventory of Executive Function (BRIEF). *BMC Psychology*, 7(1), 26. <https://doi.org/10.1186/s40359-019-0302-3>
- Baran, J. N., Noll, R. B., Jacola, L. M., Willard, V. W., Hardy, K. K., Embry, L., Kairalla, J. A., Saoji, N., Hullman, S. E., & Winick, N. (2016). General adaptive functioning in survivors of childhood acute lymphoblastic leukemia. Paper presentation. *Blood*, 128(22), 4770.
- Berlin, K. S., Parra, G. R., & Williams, N. A. (2014). An introduction to latent variable mixture modeling (part 2): longitudinal latent class growth analysis and growth mixture models. *Journal of Pediatric Psychology*, 39(2), 188–203. <https://doi.org/10.1093/jpepsy/jst085>
- Boulet-Craig, A., Robaey, P., Laniel, J., Bertout, L., Drouin, S., Krajcinovic, M., Laverdiere, C., Sinnett, D., Sultan, S., & Lippe, S. (2018). DIVERGT screening procedure predicts general cognitive functioning in adult long-term survivors of pediatric acute lymphoblastic leukemia: a PETALE study. *Pediatric Blood & Cancer*, 65(9), e27259. <https://doi.org/10.1002/pbc.27259>
- Brinkman, T. M., Zhu, L., Zeltzer, L. K., Recklitis, C. J., Kimberg, C., Zhang, N., Muriel, A. C., Stovall, M., Srivastava, D. K., Robison, L. L., & Krull, K. R. (2013). Longitudinal patterns of psychological distress in adult survivors of childhood cancer. *British Journal of Cancer*, 109(5), 1373–1381. <https://doi.org/10.1038/bjc.2013.428>
- Brown, R. T., Madan-Swain, A., Pais, R., Lambert, R. G., Baldwin, K., Casey, R., Frank, N., Sexson, S. B., Ragab, A., & Kamphaus, R. W. (1992). Cognitive status of children treated with central nervous system prophylactic chemotherapy for acute lymphoblastic leukemia. *Archives of Clinical Neuropsychology*, 7(6), 481–497.
- Bull, K. S., Liossi, C., Peacock, J. L., Yuen, H. M., Kennedy, C. R., & Children's Cancer and Leukaemia Group (CCLG). (2015). Screening for cognitive deficits in 8 to 14-year old children with cerebellar tumors using self-report measures of executive and behavioral functioning and health-related quality of life. *Neuro-Oncology*, 17(12), 1628–1636. <https://doi.org/10.1093/neuonc/nov129>
- Calsavara, A. J. C., Costa, P. A., Nobre, V., & Teixeira, A. L. (2018a). Factors associated with short and long term cognitive changes in patients with sepsis. *Scientific Reports*, 8(1), 4509. <https://doi.org/10.1038/s41598-018-22754-3>
- Calsavara, A. J. C., Nobre, V., Barichello, T., & Teixeira, A. L. (2018b). Post-sepsis cognitive impairment and associated risk factors: a systematic review. *Australian Critical Care*, 31(4), 242–253. <https://doi.org/10.1016/j.aucc.2017.06.001>
- Cheung, Y. T., Eskin, A., Inaba, H., Hudson, M. M., Pui, C. H., Krull, K. R., & Wolf, J. (2018). Association of bacteremic sepsis with long-term neurocognitive dysfunction in pediatric patients with acute lymphoblastic leukemia. *JAMA Pediatrics*, 172(11), 1092–1094.
- Common Terminology Criteria v2.0 (1999). National Cancer Institute, https://ctep.cancer.gov/protocolDevelopment/electronic_applications/ctc.htm Access date: May 1, 2020.
- Conklin, H. M., Krull, K. R., Reddick, W. E., Pei, D., Cheng, C., & Pui, C. H. (2012). Cognitive outcomes following contemporary treatment without cranial irradiation for childhood acute lymphoblastic leukemia. *Journal of the National Cancer Institute*, 104(18), 1386–1395. <https://doi.org/10.1093/jnci/djs344>
- Conklin, H. M., Ogg, R. J., Ashford, J. M., Scoggins, M. A., Zou, P., Clark, K. N., Martin-Elbahesh, K., Hardy, K. K., Merchant, T. E., Jeha, S., Huang, L., & Zhang, H. (2015). Computerized cognitive training for amelioration of cognitive late effects among childhood cancer survivors: a randomized controlled trial. *Journal of Clinical Oncology*, 33(33), 3894–3902. <https://doi.org/10.1200/JCO.2015.61.6672>
- Conklin, H. M., Reddick, W. E., Ashford, J., Ogg, S., Howard, S. C., Morris, E. B., Brown, R., Bonner, M., Christensen, R., Wu, S., Xiong, X., & Khan, R. B. (2010). Long-term efficacy of methylphenidate in enhancing attention regulation, social skills, and academic abilities of childhood cancer survivors. *Journal of Clinical Oncology*, 28(29), 4465–4472. <https://doi.org/10.1200/JCO.2010.28.4026>
- Conners, C. K. (1995). *Conners' Continuous Performance Test*. Multi-Health Systems.
- Conners, C. K. (1997). *The Conners Rating Scales—revised*. Multi-Health Systems.
- Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *The California Verbal Learning Test—adult version*. The Psychological Corporation.
- Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1994). *The California Verbal Learning Test—children's version*. The Psychological Corporation.
- Espy, K. A., Moore, I. M., Kaufmann, P. M., Kramer, J. H., Matthay, K., & Hutter, J. J. (2001). Chemotherapeutic CNS prophylaxis and neuropsychologic change in children with acute lymphoblastic leukemia: a prospective study. *Journal of Pediatric Psychology*, 26(1), 1–9.
- Halsey, C., Buck, G., Richards, S., Vargha-Khadem, F., Hill, F., & Gibson, B. (2011). The impact of therapy for childhood acute lymphoblastic leukaemia on intelligence quotients; results of the risk-stratified randomized central nervous system treatment trial MRC UKALL XI. *Journal of Hematology & Oncology*, 4(1), 42–12.

- Hardy, K. K., Embry, L., Kairalla, J. A., Helian, S., Devidas, M., Armstrong, D., Hunger, S., Carroll, W. L., Larsen, E., Raetz, E., Loh, M. L., Yang, W., Relling, M. V., Noll, R. B., & Winick, N. (2017a). Neurocognitive functioning of children treated for high-risk B-acute lymphoblastic leukemia randomly assigned to different methotrexate and corticosteroid treatment strategies: a report from the Children's Oncology Group. *Journal of Clinical Oncology*, 35(23), 2700–2707. <https://doi.org/10.1200/JCO.2016.71.7587>
- Hardy, K. K., Olson, K., Cox, S. M., Kennedy, T., & Walsh, K. S. (2017b). Systematic review: a prevention-based model of neuropsychological assessment for children with medical illness. *Journal of Pediatric Psychology*, 42(8), 815–822. <https://doi.org/10.1093/jpepsy/jsx060>
- Harila, M. J., Winqvist, S., Lanning, M., Bloigu, R., & Harila-Saari, A. H. (2009). Progressive neurocognitive impairment in young adult survivors of childhood acute lymphoblastic leukemia. *Pediatric Blood & Cancer*, 53(2), 156–161. <https://doi.org/10.1002/pbc.21992>
- Hockenberry, M. J., Hooke, M. C., Rodgers, C., Taylor, O., Koerner, K. M., Mitby, P., Moore, I., Scheurer, M. E., & Pan, W. (2017). Symptom trajectories in children receiving treatment for leukemia: a latent class growth analysis with multitrajectory modeling. *Journal of Pain and Symptom Management*, 54(1), 1–8. <https://doi.org/10.1016/j.jpain-symman.2017.03.002>
- Inaba, H., Pei, D., Wolf, J., Howard, S. C., Hayden, R. T., Go, M., Varchchouk, O., Hahn, T., Buaboonnam, J., Metzger, M. L., Rubnitz, J. E., Ribeiro, R. C., Sandlund, J. T., Jeha, S., Cheng, C., Evans, W. E., Relling, M. V., & Pui, C. H. (2017). Infection-related complications during treatment for childhood acute lymphoblastic leukemia. *Annals of Oncology*, 28(2), 386–392. <https://doi.org/10.1093/annonc/mdw557>
- Iyer, N. S., Balsamo, L. M., Bracken, M. B., & Kadan-Lottick, N. S. (2015). Chemotherapy-only treatment effects on long-term neurocognitive functioning in childhood ALL survivors: a review and meta-analysis. *Blood*, 126(3), 346–353. <https://doi.org/10.1182/blood-2015-02-627414>
- Jacola, L. M., Edelstein, K., Liu, W., Pui, C. H., Hayashi, R., Kadan-Lottick, N. S., Srivastava, D., Henderson, T., Leisenring, W., Robison, L. L., Armstrong, G. T., & Krull, K. R. (2016a). Cognitive, behaviour, and academic functioning in adolescent and young adult survivors of childhood acute lymphoblastic leukaemia: a report from the Childhood Cancer Survivor Study. *The Lancet Psychiatry*, 3(10), 965–972. [https://doi.org/10.1016/s2215-0366\(16\)30283-8](https://doi.org/10.1016/s2215-0366(16)30283-8)
- Jacola, L. M., Krull, K. R., Pui, C. H., Pei, D., Cheng, C., Reddick, W. E., & Conklin, H. M. (2016b). Longitudinal assessment of neurocognitive outcomes in survivors of childhood acute lymphoblastic leukemia treated on a contemporary chemotherapy protocol. *Journal of Clinical Oncology*, 34(11), 1239–1247. <https://doi.org/10.1200/JCO.2015.64.3205>
- Jansen, N. C., Kingma, A., Schuitema, A., Bouma, A., Huisman, J., Veerman, A. J., & Kamps, W. A. (2006). Post-treatment intellectual functioning in children treated for acute lymphoblastic leukaemia (ALL) with chemotherapy-only: a prospective, sibling-controlled study. *European Journal of Cancer*, 42(16), 2765–2772. <https://doi.org/10.1016/j.ejca.2006.06.014>
- Jansen, N. C., Kingma, A., Schuitema, A., Bouma, A., Veerman, A. J., & Kamps, W. A. (2008). Neuropsychological outcome in chemotherapy-only-treated children with acute lymphoblastic leukemia. *Journal of Clinical Oncology*, 26(18), 3025–3030. <https://doi.org/10.1200/JCO.2007.12.4149>
- Kosmidis, I., & Firth, D. (2011). Multinomial logit bias reduction via the Poisson log-linear model. *Biometrika*, 98(3), 755–759.
- Krull, K. R., Hockenberry, M. J., Miketova, P., Carey, M., & Moore, I. M. (2013). Chemotherapy-related changes in central nervous system phospholipids and neurocognitive function in childhood acute lymphoblastic leukemia. *Leukemia & Lymphoma*, 54(3), 535–540. <https://doi.org/10.3109/10428194.2012.717080>
- Krull, K. R., Okcu, M. F., Potter, B., Jain, N., Dreyer, Z., Kamdar, K., & Brouwers, P. (2008). Screening for neurocognitive impairment in pediatric cancer long-term survivors. *Journal of Clinical Oncology*, 26(25), 4138–4143. <https://doi.org/10.1200/JCO.2008.16.8864>
- Kunin-Batson, A., Kadan-Lottick, N., & Neglia, J. P. (2014). The contribution of neurocognitive functioning to quality of life after childhood acute lymphoblastic leukemia. *Psycho-Oncology*, 23(6), 692–699. <https://doi.org/10.1002/pon.3470>
- Meyers, C. A., Albitar, M., & Estey, E. (2005). Cognitive impairment, fatigue, and cytokine levels in patients with acute myelogenous leukemia or myelodysplastic syndrome. *Cancer*, 104(4), 788–793. <https://doi.org/10.1002/cncr.21234>
- Nagin, D. S., Jones, B. L., Passos, V. L., & Tremblay, R. E. (2018). Group-based multi-trajectory modeling. *Statistical Methods in Medical Research*, 27(7), 2015–2023. <https://doi.org/10.1177/0962280216673085>
- Nassar, S. L., Conklin, H. M., Zhou, Y., Ashford, J. M., Reddick, W. E., Glass, J. O., Laningham, F. H., Jeha, S., Cheng, C., & Pui, C. H. (2017). Neurocognitive outcomes among children who experienced seizures during treatment for acute lymphoblastic leukemia. *Pediatric Blood & Cancer*, 64(8), e26436. <https://doi.org/10.1002/pbc.26436>
- Partanen, M., Alberts, N., Conklin, H. M., Angheliescu, D. L., Krull, K. R., Pui, C. H., & Jacola, L. M. (2020). Association between neuropathic pain and neurocognitive functioning in survivors of childhood acute lymphoblastic leukemia. International Neuropsychological Society Annual Meeting, Denver, CO.
- Proust-Lima, C., Amieva, H., & Jacqmin-Gadda, H. (2013). Analysis of multivariate mixed longitudinal data: a flexible latent process approach. *British Journal of Mathematical and Statistical Psychology*, 66(3), 470–487. <https://doi.org/10.1111/bmsp.12000>
- Pui, C. H., Campana, D., Pei, D., Bowman, W. P., Sandlund, J. T., Kaste, S. C., Ribeiro, R. C., Rubnitz, J. E., Raimondi, S. C., Onciu, M., Coustan-Smith, E., Kun, L., Jeha, S., Cheng, C., Howard, S. C., Simmons, V., Bayles, A., Metzger, M. L., Boyett, J., ... Relling, M. V. (2009). Treating childhood acute lymphoblastic leukemia without

- cranial irradiation. *New England Journal of Medicine*, 360(26), 2730–2741.
- Pui, C. H., Nichols, K. E., & Yang, J. J. (2019). Somatic and germline genomics in paediatric acute lymphoblastic leukaemia. *Nature Reviews Clinical Oncology*, 16(4), 227–240. <https://doi.org/10.1038/s41571-018-0136-6>
- Riggs, L., Piscione, J., Laughlin, S., Cunningham, T., Timmons, B. W., Courneya, K. S., Bartels, U., Skocic, J., de Medeiros, C., Liu, F., Persadie, N., Scheinemann, K., Scantlebury, N., Szulc, K. U., Bouffet, E., & Mabbott, D. J. (2017). Exercise training for neural recovery in a restricted sample of pediatric brain tumor survivors: a controlled clinical trial with crossover of training versus no training. *Neuro Oncol*, 19(3), 440–450. <https://doi.org/10.1093/neuonc/now177>
- Ris, M. D., Grosch, M., Fletcher, J. M., Metah, P., & Kahalley, L. S. (2017). Measurement of neurodevelopmental changes in children treated with radiation for brain tumors: what is a true 'baseline? *The Clinical Neuropsychologist*, 31(2), 307–328. <https://doi.org/10.1080/13854046.2016.1216070>
- Roncadin, C., Hitzler, J., Downie, A., Montour-Proulx, I., Alyman, C., Cairney, E., & Spiegler, B. J. (2015). Neuropsychological late effects of treatment for acute leukemia in children with Down syndrome. *Pediatric Blood & Cancer*, 62(5), 854–858. <https://doi.org/10.1002/pbc.25362>
- Sands, S. A., Harel, B. T., Savone, M., Kelly, K., Vijayanathan, V., Welch, J. G., Vrooman, L., Silverman, L. B., & Cole, P. D. (2017). Feasibility of baseline neurocognitive assessment using Cogstate during the first month of therapy for childhood leukemia. *Supportive Care in Cancer*, 25(2), 449–457. <https://doi.org/10.1007/s00520-016-3422-9>
- Sidhu, N., Joffe, A. R., Doughty, P., Vatanpour, S., Dinu, I., Alton, G., Acton, B., Robertson, C. M., & Western Canadian Complex Pediatric Therapies Follow-Up Program (2015). Sepsis after cardiac surgery early in infancy and adverse 4.5-year neurocognitive outcomes. *Journal of the American Heart Association*, 4(8), e001954. <https://doi.org/10.1161/JAHA.115.001954>
- Sleuys, C., Lemiere, J., Vercruysse, T., Nolf, N., Van Calster, B., Deprez, S., Renard, M., Vandecruys, E., Benoit, Y., & Uytendaele, A. (2017). Intellectual development of childhood ALL patients: a multicenter longitudinal study. *Psycho-Oncology*, 26(4), 508–514. <https://doi.org/10.1002/pon.4186>
- Viola, A., Balsamo, L., Neglia, J. P., Brouwers, P., Ma, X., & Kadan-Lottick, N. (2017). The Behavior Rating Inventory of Executive Function (BRIEF) to identify pediatric acute lymphoblastic leukemia (ALL) survivors at risk for neurocognitive impairment. *J Pediatr Hematology/Oncology*, 39, 174–178.
- Wechsler, D. (1991). *Wechsler intelligence scale for children* (3rd edn). The Psychological Corporation.
- Wechsler, D. (1997). *Wechsler adult intelligence scale* (3rd edn). The Psychological Corporation.
- Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence* (3rd edn). The Psychological Corporation.
- Wolf, J., Tang, L., Flynn, P. M., Pui, C. H., Gaur, A. H., Sun, Y., Inaba, H., Stewart, T., Hayden, R. T., Hakim, H., & Jeha, S. (2017). Levofloxacin prophylaxis during induction therapy for pediatric acute lymphoblastic leukemia. *Clinical Infectious Diseases*, 65(11), 1790–1798. <https://doi.org/10.1093/cid/cix644>