SPECIFIC AIMS

CD19-directed chimeric antigen receptor (**CAR**)-T cell therapy[39,40] has revolutionized treatment for CD19+ B-cell acute lymphoblastic leukemia (**ALL**), inducing complete remission in 70-90% in patients who are either refractory to chemotherapy or have relapsed after prior successful treatments.[41-45] Unfortunately, CAR-T cell therapy causes a potentially life-threatening immune effector cell-associated neurotoxicity syndrome (**ICANS**) in approximately 40% of patients, which symptomatically presents as confusion, attention and memory impairments, hallucinations, delirium, and seizures.[46-49] ICANS commonly manifests 5-8 days post CAR-T cell infusion, and once present, the median time to resolution of its most florid symptoms is 9 days.[50,51] The brain and immune mediators of ICANS remain poorly understood,[52] and identifying them is essential for minimizing or treating ICANS while concurrently maximizing the anti-tumor efficacy of CAR-T cell therapy.[53,54]

We will conduct a prospective, longitudinal study and collect 380 MRI scans, cognitive assessments, and blood and cerebrospinal fluid (**CSF**) samples in 80 consecutive children and adolescents (5-21 years) with CD19+ B-cell ALL who will receive CAR-T cell therapy.[39,40] We will comprehensively assess all patients at (1) a baseline, prior to lymphodepleting chemotherapy and CAR-T cell infusion; and (2) on post-infusion day 10, during the risk period for ICANS; (3) on day 28, after resolution of ICANS; and (4) at month 6, for intermediate-term follow-up. We will collect data at least once during ICANS, and therefore will repeat data collection in patients who present neurological symptoms between days 10 and 28 post infusion. In patients, we will collect: (**a**) multimodal MRI to quantify brain structure, function, and metabolism; (**b**) detailed neuropsychological assessments; (**c**) clinical assessment of cytokine release syndrome (**CRS**) and ICANS; and (**d**) CSF when clinically indicated and blood samples for measuring cytokine levels, endothelial activation, and blood brain barrier (**BBB**) permeability, and for immunophenotyping using **CyTOF** (mass Cytometry by Time-Of-Flight). We also will recruit 20 healthy controls matched to patients on age, sex, ethnicity, and socio economic status (**SES**). In controls, we will acquire MRI and neuropsychological data at 2 time points 6 months apart.

Aim 1: We aim to understand whether the risk for developing ICANS can be predicted by pretreatment brain imaging in combination with immunological markers, whether changes in those measures accompany the development and resolution of ICANS, and whether those changes are present even in patients who do not develop clinically florid ICANS. We also aim to study whether those abnormalities persist and affect long-term neuropsychological functioning. Our preliminary data showed that at baseline due to illness and prior therapies patients relative to healthy controls have thinner cortices, lower brain perfusion, higher indices of tissue turnover, reduced density of healthy neurons, and decreased white matter integrity. These abnormalities increased by post-infusion day 10 and reverted to baseline values by day 28 in most, but not all, brain regions. **Hypotheses H1a**: Based on prior evidence that greater pretreatment ALL disease burden is associated with more severe ICANS,[50] we hypothesize that baseline brain abnormalities will be associated with ICANS severity. **H1b**: Post-infusion changes in brain abnormalities in somatosensory cortex, Broca's area, internal and external capsule, thalamus, and striatum will be associated with the severity of ICANS, and treatment-induced brain changes will significantly mediate treatment-induced impairments in cognitive functioning. **H1c**: Although brain abnormalities will attenuate following the resolution of ICANS, more subtle brain abnormalities and psychological impairments will persist and not fully resolve by 6 months post infusion.

Aim 2: Brain abnormalities will serve as objective, measurable endophenotypes[55] that will aid in identifying the inflammatory and immunological determinants of ICANS. Using brain endophenotypes as intermediate outcomes, we will assess whether ICANS is mediated by cytokines and whether cytokine levels are associated with changes in the frequency and intracellular phosphorylation of proteins in specific immune cells. Our nonhuman primate model[56] and others[50,57] suggest that cytokines mediate ICANS by activating endothelial cells and disrupting BBB. CyTOF findings in the pilot study showed greater expansion and phosphorylation of key proteins in monocytes (**MCs**) and myeloid-derived suppressor cells (**MDSCs**), suggesting these immune cells produce the elevated IL-1 and IL-6 levels implicated previously in the generation of ICANS.[46,50,58] **Hypotheses H2a**: Prior studies associated cytokine levels and BBB permeability with ICANS.[59] We therefore hypothesize that changes in IL-1 and IL-6 cytokines levels, BBB permeability, and immune cell profiles will correlate with increasing brain abnormalities during ICANS.[46,50,59] **H2b**: We also hypothesize that cytokine levels will mediate the association of immune cell activation (from CyTOF) with increasing brain abnormalities.

Aim 3: We aim to advance the mechanistic understanding for how the immune system response to CAR-T cell infusion causes ICANS and neuropsychological impairments. We will do so by conducting an exploratory, multi-group, longitudinal path analysis using a cross-lagged, structural equation model that evaluates predictive and associative relationships among longitudinally assessed brain abnormalities, neuropsychological impairments, cytokine and BBB permeability measures, and immune cell activation.

SIGNIFICANCE Previous studies of ICANS have not evaluated neuropsychiatric symptoms or systematically and prospectively characterized cognitive, behavioral, and emotional functioning, thereby likely underestimating the prevalence rates for ICANS. To our knowledge, potentially enduring sequelae of ICANS have not been assessed past two months post-infusion and have been limited to assessing the most florid ICANS symptoms, such as delirium and hallucinations, rather than more subtle behavioral and cognitive impairments. No study has assessed whether ICANS produces long-term alterations in brain structure and functioning. Understanding better the pathophysiology of ICANS and its long-term outcomes is critically important for its early detection and better clinical management. *Identifying predictive biomarkers for the development of severe ICANS will permit stratification of patients according to their risk for development of this complication during CAR-T cell therapy.* Several risk factors associated with ICANS have been identified in prior clinical trials, but few of these have been prospectively validated, and none have assessed brain characteristics that make individual patients more vulnerable to developing severe ICANS. Predicting who will develop severe ICANS or identifying the adverse brain effects of CAR-T cell therapy early in the course of treatment will permit patients to be monitored more closely and even be treated prophylactically with preemptive anti-cytokine therapies.[60,61] Furthermore, identifying the brain alterations and its associated immunological profiles that produce severe ICANS will guide future clinical trials that aim to mitigate ICANS severity and neuropsychological impairments by developing effective treatments,[62] management plans,[63-65] and novel CAR constructs[66,67] that do not interfere with the efficacy of the CAR-T cell therapy.

INNOVATION This prospective, longitudinal study of children with ALL is innovative both in terms of the study design and the technology that we will bring to this area of investigation, which together will provide previously unrealized inferential power about the nature and extent of neurotoxicities associated with immunotherapy. First, to the best of our knowledge, this is the first study of CD19 CAR-T cell immunotherapy that will prospectively acquire state-of-the-art, multimodal brain MRI data using the most advanced MRI technologies available and that will quantify abnormalities in structure, function, and metabolism across the entire brain, from a pretreatment baseline, through the course of treatment, and into long-term follow-up. These brain-based endophenotypes will increase our sensitivity for identifying the cytokine and immune cell determinants of ICANS with CAR-T cell therapy. Second, this is the first study that will acquire comprehensive, research-grade assessments of neuropsychiatric symptoms and cognitive and behavioral functioning in CAR-T cell patients. The use of comprehensive and valid assessment tools will allow us to evaluate precisely the nature of cognitive, behavioral, and emotional disturbances as a consequence of immunotherapy. For example, knowing whether memory impairment is due to poor attention or information retrieval, or whether impairment in executive functioning is due to poor processing speed, planning, or response inhibition, can guide caregivers to the most appropriate cognitive intervention for those impairments. Long-term follow-up will allow us to ascertain whether acute disturbances persist upon completion of immunotherapy. This study, therefore, will allow us to determine how brain disturbances associated with immunotherapy mediate long-term cognitive, emotional, and behavioral disturbances following immunotherapy. Third, we will acquire identical multimodal MRI and comprehensive neuropsychological data in healthy controls who are age-, sex-, ethnicity-, and SES-matched to patients. These data will allow us to identify baseline brain abnormalities in patients, whether patients with greater baseline abnormalities are at increased risk for severe neurotoxicity, whether those abnormalities deviate further away from healthy values during neurotoxicity, and whether by month 6 the brain normalizes toward healthy values or elicits compensatory responses that help patients to cope with neuropsychological deficits. Fourth, we will recruit consecutive children and adolescents with ALL eligible to receive immunotherapy, thereby minimizing, if not eliminating, ascertainment bias inherent in retrospective studies of patients who have had neurotoxicity. Minimizing ascertainment bias will generalize our findings to a larger patient population. Fifth, we will use each individual as its own control in this longitudinal study design, thereby permitting us to assess true trajectories of brain changes during and after immunotherapy. Sixth, quantitative MRI will allow us to determine rigorously the effects of pretreatment biomarkers and post infusion cytokines on brain outcomes. Furthermore, multimodal MRI in the same set of individuals will allow us to constrain better the cellular interpretations of MRI findings and their associations. Finally, integrating high-dimensional, single-cell, mass cytometry with imaging and neuropsychological measures will allow us to implicate the immunological profiles of individual cells in acute neurotoxicity. CyTOF-based mass cytometry can simultaneously measure 32 or more biological signatures in single cells, including protein epitopes where antibodies bind,[68] cellular responses to stimulation as measured by phosphorylation of intercellular signaling molecules,[69] and regulators of proliferation-specific mRNA transcripts.[70] These simultaneous, high-dimensional

investigations in the same cell will greatly reduce variability and increase the resolution for discriminating functional states and metabolism across immune cells, even for rare cell populations that otherwise may be missed.[71] Therefore, CyTOF-based profiling of immune cells and associating those profiles with serum cytokines and brain endophenotypes will markedly improve our ability to implicate specific immune cells in ICANS.

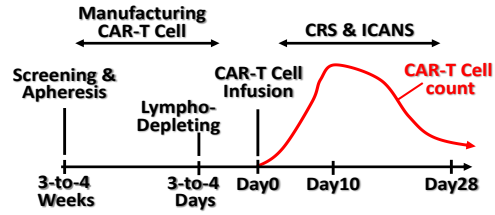

Figure 1: Time Course for Immunotherapy and CAR-T Cell Proliferation CAR-T cell manufacturing takes 2–4 weeks after leukapheresis. Patients receive lymphodepletion chemotherapy 3 to 4 days before CAR-T cell infusion. CAR-T cells are activated and proliferate *in vivo* in the first two weeks then contract, leaving a fraction of persistent CAR-Tcells.[1]

BACKGROUND

Acute Lymphoblastic Leukemia (ALL) is the most common type of childhood cancer. Although multi-agent chemotherapy regimens have improved the 5-year, event-free survival rates to 85-90%, 10-15% of patients fail to respond or suffer a relapse. Long-term outcomes are dismal for relapsed/refractory patients: only 30-50% of these patients survive following intensive chemotherapy and HSCT, thereby necessitating new treatment modalities.

CD19 Chimeric Antigen Receptor (CAR) T Cell Therapy has revolutionized the treatment for patients with relapsed/refractory ALL. CAR-T cell therapy involves autologous collection of the patient's lymphocytes via apheresis followed by lentiviral or retroviral transduction of lymphocytes to induce CD19 CAR expression and *in vitro* expansion of resultant CD19 CAR-T cells (**Fig.1**). Prior to infusion of CD19 CAR-T cells, patients receive lymphodepleting chemotherapy, such as fludarabine and cyclophosphamide, to minimize host lymphocyte rejection of CAR T cells and to thereby improve *in vivo* expansion and persistence of CAR T-cells.[72] Recent clinical trials have shown that CAR T-cell therapy induces complete remission in >80% of patients and produces *minimal residual disease* (MRD) *negative* rates in ≥ 70% of patients.[72-76] Because of these promising results, the FDA in 2017 approved two CD19-directed CAR T cell therapies (Tisagenlecleucel, Kymriah® [Novartis] and axicabtagene lisoleucel, Yescarta® [Kite]), as treatment options in pediatric and young adult patients with relapsed and refractory ALL, and adult patisents with relapsed or refractory NHL.[77]

CD19 CAR T Cell Toxicities: CD19 CAR-T cell-based therapy has significant side effects, the two most common being CRS and ICANS. CRS manifests clinically as fever, myalgias, fatigue, malaise, nausea, but also as hypoxia, hypotension, coagulopathy, and capillary leak, resulting in multi-organ dysfunction. CRS of any grade occurs in 30–94% of patients and grade ≥3 CRS occurs in 19–48% of patients.[78-86] The onset of CRS is variable, ranging from 1-22 days (median 2–3 days) after CAR-T cell infusion, with time to resolution that ranges from 1-60 (median 7) days post-infusion. The clinical symptoms of ICANS include headache, tremor, dysphasia, ataxia, apraxia, facial nerve palsy, dysmetria, and delirium with preserved alertness. More severe ICANS symptoms include focal neurological deficits, as well as delirium with a depressed level of consciousness, generalized tonic/clonic seizures, and—in the most severe cases— rapidly progressive, diffuse cerebral edema.[57,79,81,82,87-90] The majority of previous studies on ICANS has been descriptive and did not include systematic neuropsychological assessment before or after CAR-T cell infusion. Furthermore, those studies used different grading systems for evaluating neurotoxicity, resulting in highly variable rates, with neurotoxicity of all grades reported in 20–64% of patients, and grades ≥3 neurotoxicity reported in 13–52% of patients.[57,79,81,82,87-91] The first global clinical trial, ELIANA, of pediatric and young adult patients treated with tisagenlecleucel reported that 40% patients developed ICANS, with 13% patients developing grade ≥ 3 ICANS.[75] The median onset of ICANS was 5-8 days post-CAR infusion, concurrent with CRS, but occasionally beginning after CRS resolution or independent of it, with the resolution of symptoms within four weeks.[57,82,87]

To the best of our knowledge, only one small, preliminary study has systematically assessed ICANS symptoms and cognitive functioning prior to and after B-cell directed CAR-T cell infusion. In 22 pediatric and young adult patients (ages 7–30; mean 17.9 years) consecutively enrolled in a phase I dose-escalation study of anti-CD22 CAR-T therapy for relapsed or refractory B cell malignancies, cognition was assessed primarily for frontal lobe functioning in 4 domains (attention, working memory, cognitive flexibility, and processing speed), and some, but not all, ICANS symptoms were assessed using a checklist, completed by the parents, at days 0, 7, and 21-28 post-CAR therapy.[92] Nineteen patients had prior allogeneic HSCT, 18 had received either total body or whole brain radiation, 19 had received prior CD19 CAR-T cell therapy, and 3 prior blinatumomab. Baseline brain MRI in 20 subjects confirmed the absence of active CNS disease but showed cerebral atrophy in 6 subjects and vascular anomalies in 2. Of the 15 patients with cognitive testing at multiple time points, 12 had stable or improved scores on all four tests, 2 had stable or improved scores on 3 tests and a decrement on 1, and 1 showed clinical decline on more tests. Scores did not change significantly from pre- to post-infusion

on 3 of the tests and improved significantly for processing speed (paired t=2.15, P<0.05). These findings suggest the likelihood of practice effects in this sample, which could not be tested formally because a control group was not included. The parent checklist yielded new ICANS symptoms at any post-infusion time point for depressed mood (n=3), visual hallucinations (n=2), auditory hallucinations (n=1), unresponsiveness (n=1), and disorientation (n=2). ICANS symptoms, excluding fever and pain, was reported in 10 of 22 patients (45%), and *undoubtedly would have been more frequent if additional ICANS symptoms had been surveyed*. Importantly, only 1 of these 10 ICANS presentations reported by caregivers was captured clinically by the medical team, *suggesting strongly that ICANS prevalence and severity is likely underreported in published studies*. A repeated measures ANOVA showed a significant change in mean total symptom score over time (F=11.03; P=0.0002), with a quadratic effect (F=15.63; P=0.001) showing an initial significant increase from baseline to day 14 (F=14.68; P=0.001) and then significant decline from day 14 to day 21–28 (F=12.81; P=0.002). Serum IL-2, IL-6, IL-8, IL-10, IL-13, IL-15, TNF-α, and GM-CSF were significantly higher in subjects with more severe ICANS (hallucinations or delirium) than in those without (P<0.05). Most symptoms appeared by day 14 post-infusion, but 2 patients with extensive disease burden who had later CAR-T-cell expansion and delayed CRS onset at 21–28 days post-infusion experienced later-onset ICANS.

No studies have acquired brain imaging systematically before and after CAR-T cell infusion, so the rates of gross radiological abnormalities associated with CAR-T cell therapy are unknown. The most comprehensive study of CD19 CAR-T cell therapy thus far has reported grossly visible abnormalities on anatomical brain MRI in 23 of the 53 (43%) patients manifesting ICANS.[50] An abnormal scan was associated with a higher risk of poor outcome, suggesting that MRI may be a useful tool for patients with ICANS to predict progression of CNS lesions and clinical deterioration.[87] Vasogenic edema, leptomeningeal enhancement, and multifocal microhemorrhages were detected on T2/FLAIR images in severe ICANS cases. No studies of ICANS have been conducted using more sensitive, quantitative brain imaging techniques that can identify subtle brain abnormalities that are undetected with the naked eye used for neuroradiological readings.

Pathophysiology: The pathophysiologies of CRS and ICANS are related but distinct. CRS is a consequence of supraphysiological immune activation, resulting in excessive levels of inflammatory cytokines. High levels of interferon (INF)-γ are thought to activate macrophages, which then enhance the immune response and increase the likelihood of CRS.[93-95] INF-γ, IL-6, and TNF-α released by hyperactivation of CAR-T cells are also thought to activate and damage vascular endothelial cells, which in turn produce more IL-6, resulting in vascular instability, capillary leakage, and consumptive coagulopathy associated with more severe CRS.[84,96,97]

The mechanisms leading to ICANS are inadequately understood. CAR-T cells cross the BBB, even in instances without CNS disease[75,98-102] and therefore, could produce neurological toxicity directly by interacting with a previously undetected expression of targeted antigens in the CNS. However, there is no evidence for the expression of CD19 in brain tissue.[103] Moreover, ICANS manifests even in patients treated with CAR-T cells targeting other B-cell antigens, such as CD22,[90,92] and neurotoxicity can manifest even in the absence of CAR-T cells in the CNS,[98] suggesting that a mechanism other than direct CAR-T cell interactions with targeted antigens in the CNS. An alternative theory of the pathogenesis for ICANS involves cytokine-induced neuroinflammation and endothelial dysfunction. This leads to disruption of the BBB, which in turn produces a feed-forward loop of continued endothelial cell and pericyte activation, and subsequent breakdown of the parenchymal basement membrane and vascular disruption, cerebral edema, hemorrhage, infarction, necrosis, and neuronal death.[50] Evidence in support of this mechanism includes a profound increase in cytokines within CSF post CAR-T cell infusion relative to plasma concentrations in a large sample of 133 adults treated with CD19 CAR T cell therapy for relapsed/refractory (r/r) B-cell malignancies.[87] IL-6 in this study was identified as an especially important trigger of ICANS, as its elevation was particularly strongly associated with the development of high-grade ICANS.

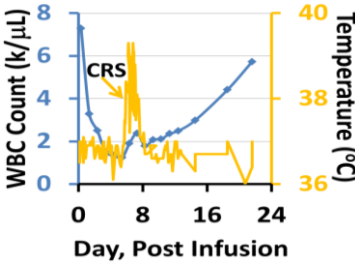

Figure 2: Post-Infusion White Blood Cell (WBC) Count and Temperature in a patient treated with CAR-T cell therapy. WBC count (blue line) declined following lympho-depletion chemotherapy but rebounded on day 5. On day 6, the patient exhibited CRS, indicated by body temperature>38°C (yellow line).

Animal models have elaborated these pathophysiological models for CRS and ICANS. A humanized mouse model with high leukemia burden showed that human monocytes were the primary source of IL-1 and IL-6 during CRS and that blockade of the IL-6 receptor prevented CRS but not ICANS, whereas blockade of the IL-1 receptor prevented both CRS and ICANS.[104] A rhesus macaque model of CRS and ICANS after CD20 CAR-T cell infusion reported disproportionately high CSF levels of IL-6, IL-2, granulocyte-macrophage colony stimulating factor (**GM-CSF**), and VEGF in

animals that developed ICANS,[56] accompanied by a pan-T cell encephalitis with multifocal perivascular T cell cuffing,[56] similar to human autopsy data after fatal ICANS.[87,105]

Predicting Toxicities Even though ICANS can occasionally occur alone, it generally occurs along with CRS, and their severities tend to correlate with one another and with any factors that increase CAR-T cell numbers,[106] including enhanced CAR-T cell expansion,[87,107] higher infused CAR-T cell dose, high disease burden, lymphodepletion chemotherapy, and patient characteristics related to the pathophysiology of CRS and ICANS, such as preexisting endothelial activation and severe thrombocytopenia.[87,107] A recent large prospective study of adults found that elevated serum levels of IL-5 and IL-2 at day 3 post-CAR T cell infusion were specific for predicting the development of severe ICANS.[106] In another study of adult patients with ALL, serum biomarkers associated with ICANS included IL-2, GM-CSF, and ferritin.[108] Whether any brain-based biomarkers are predictors of ICANS is at present unknown.

PRELIMINARY DATA

Pilot Study We conducted a pilot study (CHLA IRB # 16-00282) for generating hypotheses, establishing the collaborative clinical and research team, and ascertaining feasibility for recruiting patients, maintaining cohort, and acquiring and processing complex dataset longitudinally in patients receiving CAR-T cell therapy. In this pilot study, we approached 17 patients over 4 months who were eligible to receive CD19 CAR-T cell therapy, 9 (53%) of whom consented to participate. Following the baseline assessments, one patient dropped out of the study, and another did not respond to CAR-T cell therapy. We have obtained multimodal MRI, detailed neuropsychological assessments, and blood and CSF samples at baseline, day 10, and day 28 in 7 patients. Five patients (56%) developed CRS (**Fig.2**) and 4 (44%) developed grade 1 neurotoxicity. We also have acquired multimodal MRI data at baseline in 9 healthy controls age- and sex-matched to patients.

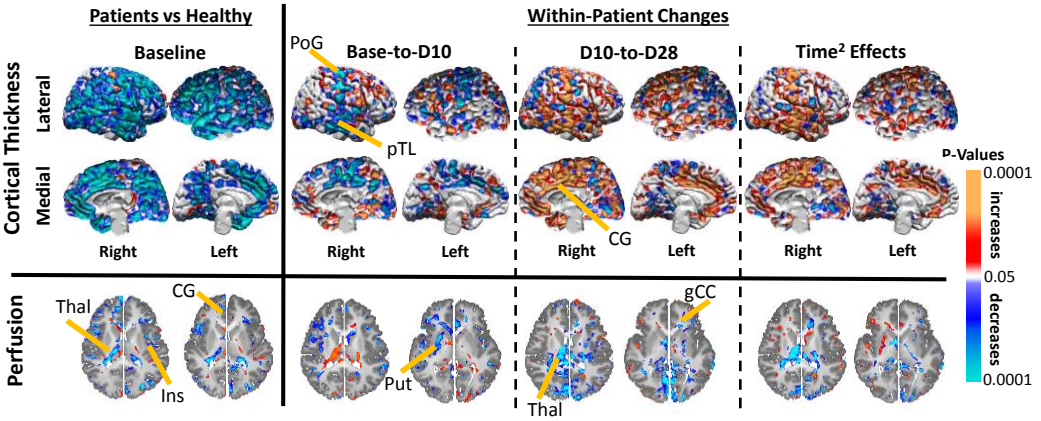

Figure 3: Baseline Abnormalities in the Cortex and Blood Perfusion and the Effects of CAR-T Cell Infusion *First Panel*: at baseline, patients relative to controls had thinner cortex across most of the brain, decreased regional cerebral blood flow (**rCBF**) in the thalamus (**Thal**), insular cortex (**Ins**), and cingulate gyrus (**CG**), and increased rCBF in the genu of the corpus callosum (**gCC**). *Second Panel*: within patients from baseline to D10, cortex thinned in the post central gyrus (**PoG**), posterior temporal lobe (**pTL**), and cingulate gyrus (**CG**), and rCBF increased in the Thal and decreased in the putamen (**Put**), globus pallidus (**GP**), and gCC. *Third Panel*: from D10 to D28, cortex to baseline values, except for certain brain regions, including Broca's region, where thinning continued; and rCBF decreased in the Thal and gCC. *Fourth Panel*: significant curvilinear changes in cortical thickness and blood perfusion across large portions of the brain. We controlled for false positives using a method for false discovery rate (FDR). Blue & violet show decreases; red & orange showing increases in MRI-derived brain measures.

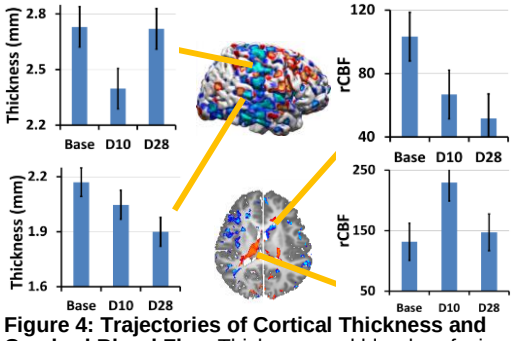

Figure 4: Trajectories of Cortical Thickness and Cerebral Blood Flow Thickness and blood perfusion first deviated away and then reverted to baseline values across most brain regions. However, in some regions, thickness and perfusion continued to deviate away from baseline values.

Relation to this proposal This hypotheses-generating pilot study demonstrates the feasibility of the proposed, definitive study. In particular, the study demonstrated that (a) our collaborative clinical and research teams can consent and recruit consecutive patients; (b) we can successfully maintain the cohort and acquire multimodal MRI, comprehensive neuropsychological assessments, and blood and CSF samples longitudinally; and (c) we can acquire motion free, multimodal MRI data with superb tissue contrast without the use of sedation in this patient population.

Trajectories of Brain Abnormalities in Patients We analyzed multimodal MRI and performed mass cytometry on blood samples for generating hypotheses of the proposed project. Anatomical MRI At baseline, patients relative to healthy controls had thinner cortices across most of the brain (**Fig.3**). Repeated measures analyses in patients only showed the cortex thinned and then reverted to baseline values by day 28, with a significant superimposed curvilinear (quadratic) component over time in many brain regions (**Figs.3&4**). Blood Perfusion Patients relative to controls at baseline had hypoperfusion of regional cerebral blood flow (**rCBF**) in

the thalamus, insular cortex, and anterior cingulate gyrus (**Fig.3**). Repeated measures analyses within patients only showed curvilinear rCBF changes in the thalamus, striatum, cingulate gyrus, and insular cortex, and linear decreases in the genu of the corpus callosum (**Figs.3&4**). Spectroscopy Brain metabolite analyses showed that patients had elevated choline (**CH**) concentrations in the striatum and low N-acetyl aspartate (**NAA**) concentrations in the insular cortex and striatum (**Fig.5**). Reduced NAA suggests a reduced density of healthy neurons,[109] whereas Increased choline, thought to be an index of membrane turnover and cellular density,[110] suggests greater tissue turnover in patients. During immunotherapy, repeated measure analyses within patients showed an initial decline in NAA concentrations, which then returned toward baseline concentrations (**Fig.5**). DTI showed that patients at baseline had reduced fractional anisotropy (FA) and increase average diffusivity (ADC) values throughout most of white matter, indicating widespread disruptions in white matter integrity, especially in the genu and splenium of the corpus callosum, and external and internal capsules (**Fig.6**). During immunotherapy, FA values initially increased and then decreased towards baseline values in patients (**Fig.6**). Day-28 measures, however, did not return completely to baseline.

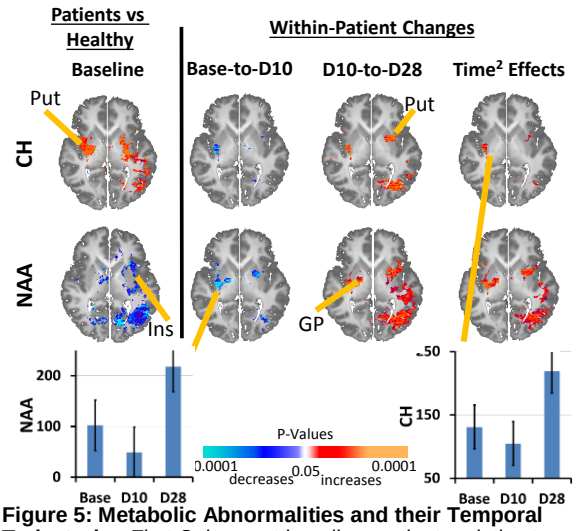

Figure 5: Metabolic Abnormalities and their Temporal Trajectories *First Column*: at baseline, patients relative to healthy controls had significantly elevated levels of choline (**CH**) in the putamen (**Put**) and globus pallidus (**GP**) and decreased levels of N-acetyl aspartate (**NAA**) in the insular cortex (**Ins**) and Put. *Second Column*: within patients from baseline to day 10 (**D10**), CH and NAA levels decreased in the striatum. *Third Column*: within patients from D10 to day 28 (**D28**), CH and NAA levels increased in Put, GP, and Ins. *Fourth Column*: changes in metabolite levels showed significant curvilinear effects, with metabolite levels first decreasing and then reverting to or increasing beyond baseline levels.

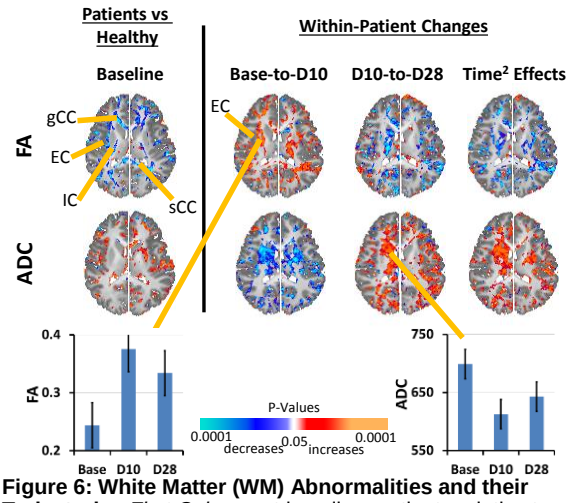

Figure 6: White Matter (WM) Abnormalities and their Trajectories *First Column*: at baseline, patients relative to healthy controls had reduced fractional anisotropy (**FA**) in the genu of corpus callosum (**gCC**), exterior capsule (**EC**), interior capsule (**IC**), and splenium of corpus callosum (**sCC**), and average diffusivity coefficient (**ADC**) was increased in EC and across gray matter (**GM**). *Second Column*: within patients from baseline to day 10 (**D10**), FA increased across majority of WM fiber tracts whereas ADC decreased in both WM and GM regions. *Third Column*: within patients from D10 to day 28 (**D28**), FA values decreased in IC and gCC, whereas ADC values increase in WM and GM regions. *Fourth Column*: there were significant curvilinear effects, with FA and ADC values first deviating away and then reverting towards baseline values.

Relation to this proposal These preliminary findings show patients have widespread abnormalities in brain structure, metabolism, and white matter integrity prior to treatment. These baseline abnormalities are not unexpected, as these patients all had been treated previously with multiple prior chemotherapy regimens that are thought to affect brain tissue.[111-118] Brain abnormalities in patients worsened by day 10 post-infusion and then generally normalized toward baseline values by day 28, sometimes not reaching baseline values, and sometimes surpassing them (as with NAA levels). These brain changes over time are generally consistent with clinical reports for the timing of ICANS, which develops by day 10 and then largely resolves by day 28 post-infusion. Brain measures in some regions, however, progressively deviated away from baseline and control values through day 28, suggesting that some treatment-induced abnormalities may persist long-term.

CyTOF-Based Immuno-Phenotyping: We used CyTOF to immunophenotype blood samples that were collected at baseline and post-infusion day 10 in 4 patients. Although all patients received lymphodepletion chemotherapy with fludarabine and cyclophosphamide, lymphocytes increased by day 5 or 6 post-CAR T cell infusion and were present in sufficient numbers for CyTOF analyses (**Fig.2**). Those analyses showed that comparing baseline and day 10 post-CAR infusion samples averaged across 4 patients the number of B cells declined as expected. In contrast, the number of classical monocytes, MDSCc, dendritic cells, effector memory T cells, effector T cells, and T cells increased during the *in vivo* activation and expansion of CAR-T cells. Monocytes differentiate into macrophages or dendritic cells, effectors of the immune response that participate in phagocytosis, antigen presentation, and cytokine production. MDSCs, in contrast, suppress the immune response by inhibiting proliferation and activation of T-cells and NK cells and may contribute to the contraction of CAR-T cells after

day 10. Furthermore, phospho-CyTOF on cytokine-stimulated cells demonstrated increased levels of phosphorylated ERK1/2 in monocytes and their derived cells, as well as in MDSCs, suggesting that these cells, especially on day 10 post-CAR infusion, have a greater capacity to produce cytokines in response to stimulation with GM-CSF or cocktail (**Fig.7***).* pSTAT3 and pSTAT5 (not shown) levels stayed high at both times in MDSCs, dendritic cells, and monocytes. *Relation to this application* These analyses show that sufficient numbers of leukocytes are present in the blood samples at baseline and on day 10 post-CAR infusion for CyTOF analyses. They also suggest that an expansion of monocytes and MDSCs occurs by day 10 post-infusion and, in addition, phosphorylation of key signaling proteins was greater in these cells in response to external stimulation, thereby implicating monocytes as a source of the elevated cytokines following CAR-T cell infusion.

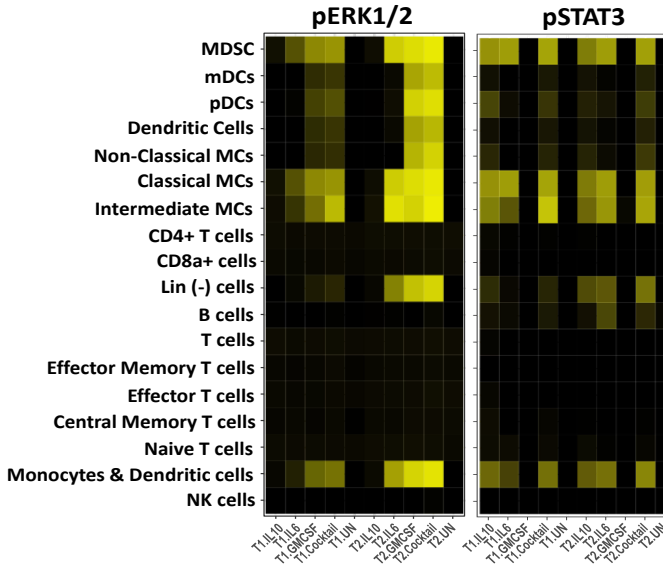

Figure 7: Phosphorylation of the Extracellular-Signal Regulated Kinase (ERK) and Signal Transducer and Activator of Transcription 3 (STAT3) Signaling Pathways that communicate cell surface signals into DNA transcription. Activation of these pathways leads to greater phosphorylation of signaling proteins. We analyzed peripheral blood samples collected at baseline (**T1**) and day 10 (**T2**) post infusion in 4 patients being treated with immunotherapy. Samples were analyzed in 5 conditions: (1) unstimulated (**UN**), and stimulated with either (2) interleukin-10 (**IL-10**), (3) interleukin-6 (**IL-6**), (4) granulocyte-macrophage colony stimulating factor (**GM-CSF**), or (5) a cocktail of IL-6, IL-10, GM-CSF, INFα). CyTOF signal was log-transformed and then color encoded. *Left*: stimulation by IL-6, GM-CSF, or the cocktail dramatically increased phosphorylation of ERK1/2 in specific cell types, including MDSCs, monocytes, and dendritic cells. Post infusion phosphorylation was greater for pERK1/2. Right: *Right*: stimulation by IL-6, IL-10, and cocktail increased phosphorylation of STAT3 in MDSCs, dendritic cells, and monocytes.

Blood-Brain Barrier Disruption Dr. Taraseviciute, a Co-I on the application, developed the first nonhuman primate (NHP) model for neurotoxicity via adoptive transfer of autologous CD20 CAR-T cells.[56] The NHP model showed that CRS and neurotoxicity were mediated by the accumulation of proinflammatory cytokines in the CSF following activation and expansion of CAR-T cells. The increased cytokine levels and T cells in the brain likely are a consequence of increased influx due to disrupted BBB, reduced efflux of cytokines, and *in situ* production of cytokines by activated microglia, astrocytes, and endothelial cells of BBB.[50,119] Activated microglia, for example, have been shown to produce proinflammatory cytokines INFγ, TNFα, and ILIβ, as well as chemokine MCP-1.[119] These findings suggest that the severity of neurotoxicity can be mitigated by (1) reducing leukocyte infiltration into the CNS, (2) reducing cytokine levels in serum, and (3) targeting activated endothelial cells or microglia. *Relation to this application*: These findings suggest that neurotoxicity is mediated by serum cytokines, which then activate endothelial cells and microglial cells to disrupt the BBB. We therefore hypothesize that cytokine levels and increased BBB permeability will mediate CyTOF-based profiles of immune cells on brain abnormalities during neurotoxicity in our sample.

METHODS

Study Design We will recruit 80 consecutive patients who will receive CAR-T cell therapy to identify brain-based and immunological predictors and mediators of ICANS as well as temporal the trajectory of those biomarkers as ICANS presents and then resolves post infusion. This, therefore, is a within-patient longitudinal study that will assess patients at four key time points: before, immediately following, and 6 months post infusion with CAR-T cells (**Fig.8**). CSF will be collected at day 10 if clinically indicated. Based on previous studies of immunotherapy,[50,76] we expect 50% (N= 40) patients to develop ICANS and of those half (N=20) to develop moderate-to-severe ICANS. Although this study focuses on within-patient changes post infusion, we will also recruit 20 healthy controls age-, sex-, ethnicity-, and socioeconomic status (SES) matched to the patients and acquire multimodal MRI and neuropsychological assessments at baseline and month 6 after baseline. We will

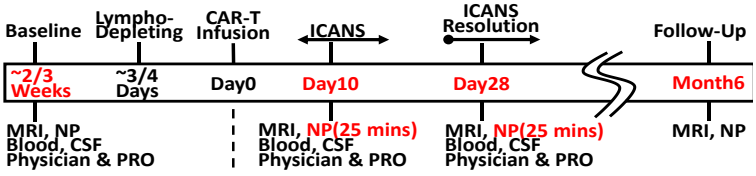

Figure 8: Patient Assessments We will collect multimodal MRI, research grade neuropsychological assessments (**NP**), physician assessed CRS and physician & patient-reported outcomes (**PRO**) for ICANS, and blood and CSF at (**1**) a baseline, and post infusion (**2**) day 10, (**3**) day 28, and (**4**) month 6. We will collect data at least once during ICANS. We will administer only a subset of NP assessments at Day10 and Day28 to minimize practice effects and reduce patient burden. The four assessment time points are shown in red.

recruit 20 healthy controls only because their data will be used for ascertaining baseline brain abnormalities and neuropsychological impairments, and for controlling time effects for when ascertaining post-infusion changes in the brain and neuropsychological functioning of patients. MRI scans for all participants will be acquired without sedation. Patients receiving intravenous fluids at the time of the scan will be nurse-supervised. We will record all clinical data, including the timing and severity of CRS and ICANS. We therefore will acquire a total of 380 brain MRI scans (4 scans at the predefined time points in 80 patients, an estimated 20 repeat scans if ICANS develops after day 10 scan, and 2 scans 6 months apart in 20 healthy controls). MRI will be processed blind to clinical status and time order of scan acquisition.

Eligibility We will recruit patients from an annual pool of about 50 patients who are treated with CAR-T cell therapy at the Children's Center for Cancer and Blood Diseases (CCCBD) of Children's Hospital Los Angeles (CHLA). Inclusion Criteria: Age: 5-21 years. Diagnosis: CD19+ B-cell ALL referred to CHLA for CD19-directed CAR-T cell therapy. Controls: Age-matched controls in good general health, with no history of malignancy who will be group-matched to patients based on sex, ethnicity, and SES. Exclusion Criteria for healthy controls include any previous or current psychiatric or neurological disorder. Additional exclusion criteria for both groups include a diagnosis of prior neurological illness or Autism Spectrum Disorder, Down syndrome, an IQ less than 70, a serious or uncontrolled chronic medical illness, or contraindication to MRI scanning.

Recruitment Eligible patients will be approached, recruited, and consented immediately after their identification at CHLA. We will recruit healthy controls by purchasing from a telemarketing company 1000 contacts for families with children whose ages, zip codes, and ethnicities are the same as our patients. We will use a random number generator to select those from the contacts who we will send introductory letters followed by screening phone calls. Based on our 20-year experience using this recruitment strategy, we expect 3-4% of those contacted will be eligible and agree to participate in the study.[120]

Feasibility For our pilot study over the last 4 months, we offered enrollment to 17 patients referred for CD19 CAR-T cell therapy, of whom 9 (53%) consented to participate. After baseline assessments, one patient dropped out and another patient did not respond to immunotherapy. We have acquired the complete dataset until day 28 in the other 7 patients. That is, patients either drop out early in the study after baseline assessment or remain in the study and complete all assessments (see Preliminary Findings). Thus, expecting that approximately 50% of patients will consent and a 20% attrition rate, we will recruit 100 consecutive patients from a pool of 200 patients over the first 4 years of the project to have a complete data in 80 patients.

Cohort Retention We will collect addresses and phone numbers of parents and relatives who maintain close relationships with the family. We will contact families with appointment reminders; if a visit is missed, we will reschedule the visit within 3 days. Frequent hospital visits for CAR-T cell therapy and follow-up monitoring will facilitate our retention efforts. We also will send greeting cards and a small gift on birthdays and holidays.

Demographics and Pretreatment Disease Burden We will collect demographic information at baseline for all participants, include age, sex, weight, and body mass index (BMI). For patients, we also will record: disease characteristics (leucocyte count, molecular subtype, CNS status); the number, type, duration, and dosing of prior therapies; prior history of HSCT; and prior disease burden, defined as the percentage of bone marrow and CSF blasts and radiographic evidence for extramedullary disease. Patients with ≥5% bone marrow blasts or those with the extramedullary disease will be considered as having a high pretreatment disease burden.

CAR-T Cell Therapy Drs. Pulsipher (Co-I) and Taraseviciute, MD, PhD, (Co-I) will supervise CAR-T cell therapy and patient-reported assessments of CRS and ICANS. Patients will receive CD19 CAR-T cell therapy either with a commercially available, FDA-approved product (e.g., Kymriah) or as part of a clinical trial. Patients will undergo standardized lymphodepleting chemotherapy with fludarabine and cyclophosphamide, administered within 1 week prior to CAR-T cell infusion. Patients will receive prophylactic therapy for infectious disease complications and levetiracetam therapy for seizure prophylaxis, beginning with the initiation of lymphodepleting chemotherapy, which will be continued through day +30 post-CAR-T cell therapy or longer as clinically indicated.

CRS Assessments Patients will be monitored closely for CRS by conducting daily assessments for inpatients and at least 3-times weekly assessments during scheduled clinical visits for outpatients. CRS will be graded using the National Cancer Institute's Common Terminology Criteria for Adverse Events version 5.0.[121]

ICANS Assessments The clinical care team will conduct once daily ICANS assessments and neurologic physical exams for inpatients or 3 times weekly during scheduled outpatient visits. We will use the Immune effector Cell-associated Encephalopathy (ICE) assessment tool (for patients >12 years) and the Cornell Assessment of Pediatric Delirium (CAPD) assessment tools (for children ≤12 years), which will be incorporated into the ASBMT ICANS consensus grading for children (Grades 1 to 4).[122] Neurology consults will be obtained

if clinically indicated. We will evaluate the severity and duration of the following: headache, delirium, hallucinations, or convulsive seizures, language disturbances or aphasia, writing ability, and orientation to place, person, and time. Previous studies have shown that observer reports provide valid assessments in children.[123,124]

Cytokine Measures We will collect 30 ml of peripheral blood for research purposes at the time of each assessment in patients, and only at baseline in healthy controls. CSF is obtained routinely in patients for clinical purposes at baseline, day 28, and, if clinically indicated, during ICANS; at those times we will also collect 3-5 ml CSF for research purposes. Blood and CSF will be frozen and stored at -80ºC. We will batch process these samples at a later time using the human cytokine/chemokine panels (see the letter from Eve Technologies, Calgary). Cytokines and inflammatory molecules measured will include IL-1α, IL-1β, IL-2, IL-3, IL-5, IL-6, IL-10, IL-12p40, IL-12p70, IL-13, IL-17A, granulocyte colony stimulating factor (G-CSF), granulocyte macrophage colony-stimulating factor (GM-CSF), interferon-γ, macrophage inflammatory protein-1α, MIP-1β, monocyte chemoattractant protein 1, regulated on activation normal T cell expressed and secreted (RANTES), and TNF-α.[125] We will analyze CSF for the presence of WBCs and cytokine and albumin levels.

Endothelial Cell Activation We will measure the activation status of endothelial cells as the ratio of ANG 2:ANG1 angiopoietin in serum.[126] ANG 1, produced primarily by vascular pericytes and platelets, binds to the TIE2 receptor, producing quiescence and stabilization of endothelial cells. ANG2, stored in endothelial Weibel-Palade bodies, is released when endothelial cells are activated by inflammatory cytokines, causing increased activation and permeability of microvasculature. We will confirm endothelial cell activation by evaluating concentrations of von Willebrand factor (vWF), a glycoprotein stored in endothelial Weibel-Palade bodies that is released when endothelial cells are activated (Eve Technologies, Calgary).[127,128] We will assess whether the ANG 2:ANG1 ratio and vWF concentrations during acute ICANS are higher in patients with grade ≥3 compared with grade <3 ICANS. We also will assess whether these serum-based biomarkers pre-infusion predict ICANS grade and MRI brain abnormalities.

BBB Permeability Neuron-specific proteins such as glial fibrillary acidic protein (GFAP), neuron-specific enolase (NSE), and S100B are biomarkers used to assess BBB integrity in other brain injuries (Alzheimer's, stroke, traumatic brain injury).[129,130] Under normal conditions, these should not be detected in serum. Commercially available kits (Eve Technologies, Calgary) will be used to assay these proteins. In addition, CSF to serum albumin ratio is a reliable and commonly used index of BBB integrity.[131-134] CSF collected at baseline, day 28, and during ICANS will permit us to measure this ratio and correlate it with other biomarkers, cytokines, and imaging measures.

Cytometry by Time Of Flight (CyTOF) We will conduct mass cytometry using the Helios system of Fluidigm, Inc., which permit interrogation of more than 40 markers simultaneously on individual cells. This groundbreaking technology, therefore, reveals network status within the immune system at a single-cell resolution. We will assess the functional states of a subset of immune cells in response to specific cytokines (IL-6 or a cocktail comprising of IL-2, IL-6, GM-CSF, and INFα) and, in addition, assess alterations in signaling pathways via phospho-CyTOF in peripheral blood samples.[135] The resulting mass spectrometry data are analyzed using unsupervised computational tools in Cytobank[136] to comprehensively explore these immune perturbations with minimal bias.[135,137,138] These analyses will identify populations of immune cells with high accuracy using a 32-antibody panel for cell surface markers as well as 6 phosphoproteins: pSTAT1, pSTAT3, pSTAT5, pP38, pERK1/2, and pS6. Immunophenotyping via CyTOF will allow us to test the hypothesis that the functional states and capabilities of circulating immune cells of CAR-T cell patients are distinct before and after CAR-T cell therapy, and we will gain insights into which cells specifically mediate ICANS. CyTOF technology will therefore help dissect the pathogenesis of immune-mediated toxicities by implicating specific immune cells as the source for serum IL-1, IL-6, and other cytokines associated with ICANS.

Neuropsychological Assessments Dr. Joey Trampush, Ph.D. (Co-I & a senior developmental neuropsychologist) will oversee the neuropsychological protocol, which will evaluate the cognitive, behavioral, and emotional disturbances associated with immunotherapy. *Assessments for Spanish-speaking parents will be conducted by native Spanish speakers. Our experience over the past 5 years has been that children are bilingual and fluent in English and therefore they will be tested in English. The selected instruments are valid in the age range of participants in this study.* We will obtain standardized reports of demographics, medical history, handedness, and socioeconomic status. Comprehensive cognitive assessments will be obtained at baseline and month 6 (*total assessment time: 3 hours*). However, post infusion day 10 and 28, participants will complete only a subset of assessments marked with asterisks below, which have minimal practice effects (*total assessment time: 25 minutes*). We will further reduce practice effects by using an alternate, but equivalent,

instruments when available. Additionally, Reliable Change Indices will be calculated using both null hypothesis[139] and estimation interval methods.[140,141] We will asses reliability, test bias, and drift on measures using 10 randomly selected videotaped assessments. Children will be permitted 15-minute breaks as needed to maintain adequate stamina and performance for the assessments. We will assess the following neurocognitive and behavioral domains: *Premorbid Ability/Reading Screen:* Wide Range Achievement Test, 5th ed. (WRAT5) Word Reading; *General Intelligence, Fluid Reasoning, and Crystallized Knowledge:* Wechsler Preschool and Primary Scale of Intelligence-IV (30min) or Wechsler Abbreviated Scale of Intelligence (30min) depending on age; *Sustained Attention/Vigilance:* Kiddie Conners' Continuous Performance Test–3rd Edition[142] (7min) or Conners' Continuous Performance Test[143] (14min); *Working Memory:* Wechsler Intelligence Scale for Children-5th Edition Digit Span (7min) and Wechsler Nonverbal Scale of Ability[144] Spatial Span (7min); *Executive Functioning:* NIH Toolbox *Dimensional Card Sort Change (5min), Reward-Based Go-no-Go Task (15min), and Simon Spatial Incompatibility Task[145] (5min); *Processing Speed:* WPPSI-IV Animal Coding and Bug Search or WISC-V Coding and Symbol Search, and *NIH Toolbox: Pattern Comparison Processing Speed (4min); *Visuomotor Integration & Fine Motor Coordination:* Beery-Buktenica Test of Visual-Motor Integration 6th Edition[146] (15min) and NIH Toolbox: 9-Hole Dexterity Pegboard (5min); *Learning & Memory:* California Verbal Learning Test-Children's Version (20min), Wide Range Assessment of Memory and Learning Design Copying[147] (10min), and NIH Toolbox: Picture Sequence Memory (7min); *Parent-on-Child Reports of Emotion & Behavior:* Kiddie Schedule for Affective Disorders and Schizophrenia, Present and Lifetime Version[148] (60min), Life Events and PTSD Module[149] (10min), *Screen for Child Anxiety Related Disorders-Parent Version (7min), Social Responsiveness Scale 2[150] (15min), Behavior Assessment System for Children, 3rd ed (BASC-3) Parent Rating Scale[151] (15min) and NIH Toolbox for Emotion Parent Proxy Battery (15min); *Child-Self-Reports of Affect & Behavior:* *Children's Depression Rating Scale, Revised[152] (5min), *The Screen for Child Anxiety Related Disorders-Child Version[153] (7min), and Personal Experiences Checklist[154] (5min); *Teacher-on-Child Report of Emotion & Behavior:* BASC-3 Teacher Rating Scale.[151]

MRI Scanning MRI data will be acquired on our 3.0T Philips Achieva MRI scanner equipped with a 32-channel, phased array head coil. Participants will be familiarized with scanning procedures in a mock scanner on the morning of scan. Frequent praise and reminders to remain still during scan will help to acquire motion-free images. We will acquire high resolution T1- and T2-weighted anatomical MRI,[155] high angular resolution DTI,[156] perfusion MRI,[157] MPCSI, and resting state fMRI data for myelin content mapping data using multiband sequences (**Table 1**). These sequences, including the ASL-based perfusion imaging, do not require administration of contrast agents. Moreover, all MRI data will be acquired without sedation. MRI is entirely safe in children, even when scanning children repeatedly in longitudinal studies, as it is noninvasive and does not use ionizing radiation.[158] Below, we briefly describe the processing of MRI datasets, which are amply documented in detail in our publications.[120,155,159-164] Total scan time is 55 mins.

	Pulse Sequence	Resolution (mm^3)	TR/TE/FA (ms/ms/degree)	Slices	Sequence Specific Parameters	Time
T1w	3D gradient echo	0.8 x 0.8 x 0.8	8.9/3.4/8°	WB	TSE = 128, NEX=2	10 min, 18 sec
Multi-Shell DTI	Single shot, spin echo	1.89 x 1.89 x 2.0	5300,89/78°	WB	b=500, 1000, 2000, & 3000 s/mm^2; 100 DWIs, 2 B0 images, EPI factor = 119	7 min, 30 sec
rsFMRI	Echoplanar	2 x 2 x 2	2000/30/90°	WB	250 imaging volumes x 2 runs; multiband=6; SENSE=1; 10 dummy scans/run	7 min
Perfusion	2D pCASL	2.5 x 2.5 x 5	4000/14/90°	WB	PLD=1525ms; Labeling duration=1800ms; 2 background suppression pulses	10 min
MPCSI	MPCSI, spin echo	10 x 10 x 10	2300/144/90°	6	CHESS water suppression; OVS	20 min

Table 1: MRI Pulse Sequence Parameters **T1w** = T1-weighted; **PLD** = post labeling delay; **TSE** = turbo spin echo; **pCASL** = pseudo-continuous arterial spin labeling; **MPCSI** = multiplanar chemical shift imaging; **OVS**=outer volume suppression; WB = whole brain

Post-Acquisition Quality Control is conducted within 48 hours of every scan. For some sequences (rs-fMRI, MRS), motion is not always readily visible on the console and is detected only during post-processing, after the participant has left the scanning suite. For *fMRI*, we quantify head motion for each participant using 2 summary statistics, Root Mean Squared (RMS)[37] and Mean Frame-wise Displacement (FD) metric, both of which sum differentiated realignment estimates[38] derived from 3 translational (x,y,z) and 3 angular rotation (roll, pitch, yaw) variables. We remove images with >0.5 mm motion from further preprocessing; if >10% of images have this much motion, that dataset is excluded from further processing. For *MRS*, we reconstruct the data, assessing it for excess noise, and visualizing the spectrum in each voxel for baseline distortions, signal contamination by lipid signal from the scalp, incorrect placement of suppression bands, or broadening of linewidth. Sequences containing any of these artifacts are not processed further. *Remediation Plans for Data*

Loss If any sequences fail quality control, the participant will be invited back for a repeat of the problematic sequence(s). Plans also include overrecruiting to replace participants with failed scans, in addition to retaining enrollment of those who fail a scan or sequence at any time point, but estimating and accounting for missing data in the data analytic plan (see Biostatistics).

I**mage Processing Methods** Though our image processing methods are state-of-the-art, they are also quite standard, have been described in peer-reviewed publications in more detail than possible here,[7,21,35,36,157,161,165-174] and have been implemented in our laboratory for many years. Therefore, we have placed the description of those methods in the Facilities section under, "Image Processing within the Pediatric Imaging Laboratory."

STATISTICAL ANALYSES Co-I/Biostatistician Dr. Ji Hoon Ryoo, PhD, will oversee all utilizing linear mixed-effects models (**LMM**)[175,176] and structural equation modeling (**SEM**)[177,178]. All data will be scored and rescored by separate staff members and checked for accuracy prior to double entry into a REDCap[179] database.[180] Imaging, behavioral, clinical, and immunological data will backed up daily. Prior to statistical modeling, we will perform descriptive analyses and data visualizations to examine the distributions of all clinical, brain imaging, neuropsychological, and cytokine measures. Our extensive preliminary data indicate that our dependent measures are normally distributed and that covariance matrices are similar across groups. Appropriate transformations to obtain approximate normality will be performed; where necessary, additional nonparametric tests will confirm results. MRI and neurodevelopmental outcome measures will be continuous variables. General linear regression and linear mixed models will be performed at each voxel of the image, with brain measure (either local brain volumes, rCBF, metabolite levels, or diffusion tensor vectors) serving as the dependent variable and immunological profiles as the independent variable to test our *a priori* hypotheses. We will account appropriately for the temporal correlation of repeated measures across time, the correlation among multivariate measures, and the spatial correlation and smoothness of imaging measures. Covariates will include age, sex, and socioeconomic status.**[181]** Multiple comparisons across all the voxels of the image will be corrected using a procedure for false discovery rate**[182]** and Gaussian Random Fields (GRF),[170] and hypothesis testing will be adjusted for multiple comparisons. We will examine parameter estimates, 95% confidence intervals, and p-values of the component terms. Least squares means and standard errors will be calculated in the models and then plotted to assist in interpretation of significant interactions.

Testing H1a: We hypothesized that pretreatment brain abnormalities will predict severity of clinical neurotoxicity following infusion with CAR-T cells. We will first conduct between-group analysis at baseline to assess whether patients relative healthy controls have abnormal brain structure, function, and metabolism by applying a LMM. We then will conduct a within-patient analysis to assess whether those baseline brain abnormalities predict severity of clinical neurotoxicity sCN_i at day 10 by using a GLM $sCN_i = b_0 + b_1 * Age_i + b_2 * Sex_i + b_2 * Brain_i^b + \epsilon_i$, where for patient i, $Brain_i^b$ is the baseline brain measure at day 10.

Testing H1b: We hypothesized that patients with greater increase in brain abnormalities will have more severe neurotoxicity. To test this hypothesis, we first will evaluate whether brain abnormalities are associated with severity of neurotoxicity at day 10 using a LMM. We subsequently will test whether patients with more severe neurotoxicity have greater increase in brain abnormalities by using a LMM: $Brain_i^{10} = b_0 + b_1 * Age_i + b_2 * Sex_i + b_3 * Brain_i^b + b_4 * sCN_i + \epsilon_i$, where $Brain_i^{10}$ is the brain measure at day 10. We also hypothesized that treatment-induced brain abnormalities will mediate the effects of neurotoxicity on treatment-induced cognitive and behavioral impairments. We will test this hypothesis by using a two wave, longitudinal mediation model that assesses a causal relation between changes in brain measures with changes in neuropsychological functioning in patients with and without clinical neurotoxicity[183-185] and use the Sobel test[186] to assess statistical significance of the cross-time mediating effects of brain abnormalities.[187]

Testing H1c: We hypothesized that although brain abnormalities will attenuate by day 28, subtle brain and cognitive changes will persist into month 6 post infusion. We will test whether brain abnormalities in patients attenuate by day 28 by applying a LMM for brain imaging data collected at baseline, day 10, and day 28. We expect significant curvilinear effects, which will demonstrate that brain abnormalities attenuate by day 28 post infusion. We further hypothesized that subtle brain abnormalities in patients will persist into month 6 post infusion. We will test this hypothesis by applying a linear mixed-model (LMM)[188] to data collected 5 months apart in patients and healthy controls. Finally, we will test whether in patients long-term brain changes are associated with changes in neuropsychological impairments by applying a repeated-measures analysis[189] $Brain_i^j = \beta_0 + \beta_1 * Age_i + \beta_2 * Sex_i + \beta_3 * Time_i^j + \beta_4 * \widehat{NP_i} + \beta_5 * \left(NP_i^j - \widehat{NP_i}\right) + b_{0i} + b_{1i} * Time_i^j + \epsilon_i^j$, where $\widehat{NP_i} = \frac{1}{2}\left(NP_i^{m6} + NP_i^{28}\right)$ is the average of the neuropsychological measures at month 6 and day 28, β_4 is the between-patient correlation of brain measures with neuropsychological measure, and $\beta_5$is the

longitudinal, within-patient brain change accompanying the change in neuropsychological performance.

Testing H2a: We hypothesized that baseline cytokine levels and BBB permeability will predict increase in brain abnormalities from baseline to day 10. We will test this hypothesis by using a LMM: $Brain_i^{10} = b_0 + b_1 * Age_i + b_2 * Sex_i + b_3 * Brain_i^b + b_4 * Blood_i^b + \epsilon_i$, where $Brain_i^b$ is the baseline brain measure, $Brain_i^{10}$ is the brain measure at day 10, and $Blood_i^b$ is the baseline cytokine levels or BBB permeability measure.

Testing H2b: We hypothesized that serum cytokines partially mediate the associations of CyTOF-derived immunological profiles of activated monocytes and endothelial cells with brain abnormalities during neurotoxicity at day 10. We will test this hypothesis by applying a mediation analyses. Let X=immunological profiles, M=serum cytokine levels, and Y=brain measures. The mediation analyses examines the association between: (1) immune cell profiles and brain outcome measures, $Y = c_1X + e_1$; (2) immune cell profiles and cytokine levels, $M = aX + e_2$; and (3) brain outcomes and immune cell profiles while adjusting for the cytokine levels, $Y = c_2X + bM + e_3$. [187]

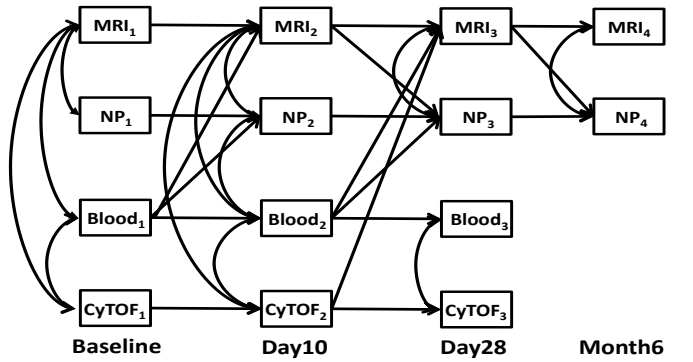

Figure 9: A Hypothesized Cross-Lagged, Structural Equation Model relating (1) MRI-derived brain measures (**MRI**); (2) assessments of attention and executive functioning or visuomotor integration (**NP**); (3) measures of IL-1 or IL-6 cytokines or BBB permeability (**Blood**); and (4) CyTOF-based profiles of monocytes (**CyTOF**). At baseline, blood- and CyTOF-based markers will be associated with brain measures, which in turn will be associated with neuropsychological functioning. Furthermore, baseline blood markers will predict brain abnormalities and neuropsychological impairments at day 10. At day 10, brain abnormalities will be associated with neuropsychological impairments, blood markers, and immunological markers. The blood- and CyTOF-based markers at day 10 will predict brain abnormalities and neuropsychological impairments at day 28. Furthermore, brain abnormalities at day 10 will predict neuropsychological impairments at day 28, whereas abnormalities at day 28 will predict neuropsychological impairments at month 6.

Mechanistic, Exploratory Analyses We will apply a multi-group, cross-lagged, SEM (**Fig.9**)[190,191] to test simultaneously the associative and predictive relationships among the longitudinally acquired brain measures, neuropsychological assessments, serum measures of cytokines and BBB permeability, and immune cell profiles. Based on the reported findings, we will specify longitudinal SEMs using MRI derived measures of blood perfusion, NAA concentration, or cortical thickness; neuropsychological assessments of attention and executive functioning or visuomotor integration; serum measures of IL-1, IL-6, or BBB permeability; and single cell measures for activity and metabolic profiles in macrophages or antigen presenting cells. Instead of a confirmatory approach that tests whether a prespecified model explains the data, we will use a nested model approach wherein we will first describe alternate models that use the same data and conduct a difference-in-chi-squared test. We further will delete or add paths by assessing statistical significance of each individual path in the initially specified model to arrive at the final best fit model. We will conduct a single statistical test across all nested models but report parameters for each of the SEMs for different single cell profiles. Finally, we will carefully assess whether the estimated magnitudes and directions of path coefficients are biologically plausible. The SEM models will be compared at a p-value <0.05 for the χ^2 difference computed using the Satorra-Bentler scaled χ^2 value.[192-194] If two models fit the data equally well, then we will select the most parsimonious one. Finally, we will omit nonsignificant paths in the best fit model to determine the most parsimonious, best fit model.[195] We subsequently will evaluate how well the final model fits the data by computing the widely reported, root-mean-square error of approximation (RMSEA)[196-198] and require that RMSEA <0.06 and comparative fit index (CFI) >0.95 for accepting it as a preliminary mechanistic model for neurotoxicity.[199]

Statistical Power Analyses We performed power analyses in Mplus v.8.2 using Monte Carlo simulation that implemented 10,000 replications with random seed values. Variances ratios, attrition patterns, and effect sizes were estimated from our pilot study of treatment with CD19 CAR-T cell therapy. The statistical models in hypotheses testing were used for both data generation and analyses. <u>LMMs in H1a and H2a:</u> For a Type-I error rate $\alpha = 0.05$ and assuming the residual variance estimates ($\mathrm{var}_{\mathrm{ba}} = 1$ and $\mathrm{var}_{\mathrm{ICANS}} = 1$ for the brain abnormalities (ba) and ICANS severity in H1a, and $\mathrm{var}_{\mathrm{blood}} = 1$ and $\mathrm{var}_{\mathrm{ba10}} = 1$ for the blood cytokines and brain abnormalities at day 10 (ba10) in H2a) and path coefficients ($\mathrm{b}_{\mathrm{ba2ICANS}} = 0.375$ for the path from ba to ICANS in H1a, and $\mathrm{b}_{\mathrm{blood2ba10}} = 0.575$ for the path from blood to ba10 in H2a), 80 patients provide a statistical power of 0.910 and 0.999 for testing H1a and H2a, respectively. <u>Mediation Analyses in H1b and H2b:</u> We considered

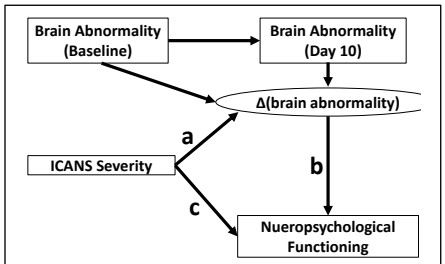

Figure 10: Path Diagram for power analyses in H1b

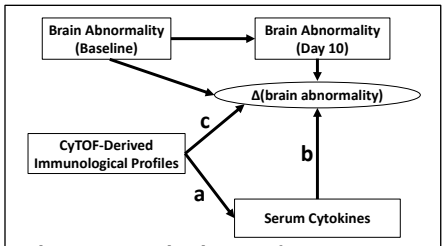

Figure 11: Path Diagram for power analyses in H2b

10% missing data for mediator and 10% missing data for final outcome in H1b and H2b. We assumed data is missing at random. That is, missing data is from family circumstances rather than from study-related variables,[200] and therefore, we will applied the full information maximum likelihood (FIML). We assumed residual variances ($\mathrm{var}_{\mathrm{np10}} = 1$ for neuropsychological measures on day 10 (np10) in H2a, and $\mathrm{var}_{\mathrm{cytof}} = 1$ for CyTOF-drived immunological profiles on day 10 in H2b), and path coefficients ($\mathrm{b}_{\mathrm{ICANS2ba10}} = 0.375$ from ICANS to ba10 and $\mathrm{b}_{\mathrm{ba102np10}} = 0.5$ from ba10 to np10, representing small-to-medium and medium-to-large effects,[201] respectively, in H1b, **Fig.10**; and $\mathrm{b}_{\mathrm{cytof2blood}} = 0.375$ from CyTOF to blood and $\mathrm{b}_{\mathrm{blood2ba10}} = 0.575$ from blood to ba10, representing small-to-medium and medium-to-large effects,[201] respectively, in H2b, **Fig.11**). Statistical analyses showed that 80 patients provide a statistical power of 0.824 and 0.811 for testing mediations in H1b and H2b.

Potential Pitfalls and Alternative Strategies Prior Treatments Patients with relapsed/refractory ALL will have had received multiple prior therapies, including chemotherapy, radiotherapy, and HSCT, which are expected to affect baseline brain measures and neuropsychological functioning. Our use of longitudinal, repeated measures study design treats each patient as its own control, and therefore controls for prior treatment effects and ascertains true temporal trajectories. Baseline abnormalities, however, may moderate changes in the brain and neuropsychological measures. We therefore will conduct post hoc analyses to evaluate whether the number and type of prior therapies moderate changes in the brain or neuropsychological functioning.

Non-Responders and Post-Immunotherapy Treatments Our preliminary data show that approximately 20% of patients may drop out or not respond to immunotherapy. We will account for the 20% attrition by recruiting 100 patients, which will provide us with a full cohort of 80 responders for testing our *a priori* hypotheses. Patients who do not respond to immunotherapy will likely be identified by day 14 post-infusion because of their lack of T-cell expansion or persistent leukemia cells in the blood. We will continue to follow nonresponding patients and will conduct *post hoc* analyses to assess the trajectories of their brain abnormalities and neuropsychological functioning. In addition, some patients who respond to CAR-T cell therapy receive HSCT within 1 to 3 months post-infusion and prior to our 6-month assessment. Some patients may relapse post-immunotherapy and HSCT and will receive other salvage regimens. We will incorporate any post-CAR-T-treatments as modifiers in assessing the long-term outcomes in patients. We expect that patients who receive post-immunotherapy treatment, relative to those who do not, will have poorer neuropsychological outcomes.

Psychological Trauma of having a life-threatening illness may affect neuropsychological functioning in patients. Prior studies and our preliminary data thus far suggest that the instruments that we will use generate valid measures of the cognitive and behavioral functioning in this patient population.[202] Nevertheless, we will evaluate the severity of depression, anxiety, and stressful life events, and conduct *post hoc* analyses to discern whether these symptoms moderate the association of brain with neuropsychological measures.

Missing Data Some patients may miss one or more assessments. If the data are assumed to be missing at random then full information maximum likelihood (FIML) methods can be applied.[203,204] Available evidence indicates that FIML estimates match those that are estimated using full data and have lower standard error than the other common missing data approaches such as multiple imputation in unbiased standard error estimation.[203,204] In contrast, data missing not at random (MNAR) may cause bias in the analyses, loss of power, or both. For MNAR data, we will use logistic regression to identify the factors related to the probability of missing data and include these factors as covariates to minimize bias. Furthermore, we will calculate probability weights from a logistic regression model with the outcome being 'has complete data'. The fitted values will estimate the probability of having missing data and we will assess influence by comparing results from the weighted and unweighted analyses.

Age Range of Patients The participants in our study will be 5 to 21 years of age. Therefore, we will covary for age in all analyses to control for age-related variations in brain measures. Moreover, and most importantly, the longitudinal study design permits each patient to serve as her own control in linear mixed modeling, and we will be able to identify within-subject brain changes associated with CAR T-cell treatment. Nevertheless, these brain and cognitive effects of treatment may differ in children (5-12 years) compared with adolescents (13-21 years). We will therefore conduct *post hoc* subgroup analyses within (a) children only and (b) adolescents only, to determine whether brain and cognitive abnormalities differ across the two age groups. In the absence of significant group effects, we will use simplified statistical models that pool patients in the two age subgroups.