1. Supplemental Methods

1.1 Data Collection and Definitions

Adverse events (AEs) were collected per NCI Cancer Therapy Evaluation Program (CTEP) AE reporting requirements using Common Terminology Criteria for Adverse Events (CTCAE) version 4 (NCT01593696 and NCT02315612) and version 5 (NCT03448393). AEs were collected at baseline (prior to initiation of lymphodepleting chemotherapy) and through at least day +28 (+/- 4 days) disease assessment, return to home institution, or off treatment (whichever came latest). AEs were captured and graded weekly by 2 consistent data managers using patient chart data using CTCAE definitions and grading schema (as below). AEs were then verified by research nurses and clinicians directly involved in the patients' care and given an attribution to research, investigational new drug (IND), and disease per the categories listed below. AE reports were then uploaded to the NIH Adverse Event Reporting System.

CTCAE Grades:
Grade 1 Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated.
Grade 2 Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL).
Grade 3 Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self care ADL.
Grade 4 Life-threatening consequences; urgent intervention indicated.
Grade 5 Death related to AE.

Adverse Event Attribution Categories:

Attribution	Description
Unrelated	The AE is clearly NOT related to the intervention
Unlikely	The AE is doubtfully related to the intervention
Possible	The AE may be related to the intervention
Probable	The AE is likely related to the intervention
Definite	The AE is clearly related to the intervention

1.2 Study Population

Standard protocol specific eligibility criteria (e.g., age, detectable antigen-positive disease, adequate performance status and baseline organ function) guided enrollment. Prior therapy with hematopoietic stem cell transplant (HSCT) and CAR T-cell infusion of a different product were not exclusionary, however, second infusions of the same product were not included.

1.3 Data Extraction

For retrospective analysis, AE data for all patients treated on the three above protocols up through data cut off of 12/31/2020 was extracted from the NIH Adverse Event Reporting System database. AEs data was edited to remove any events less than grade 3 or occurring outside the

timeframe of days 0-30. The following AE terms were eliminated: “Fever, Lymphocyte Count decreased, WBC decreased, CRS, Multiorgan Failure, CPK.” Only those AEs with at least one attribution (research, IND, disease) of possible or higher were included (in other words, any AEs with all attributions designated unlikely or unrelated were excluded). AEs that were dose limiting toxicities (DLTs) that met other criteria for inclusion in the analysis were included.

An excel file was generated for each patient, with one tab including desired demographic and disease data, and a second tab including extracted AEs meeting above criteria.

1.4 CAR Constructs

NCT01593696: CD19.28z CAR T-cell construct
NCT02315612: CD22.41BBz: CAR T-cell construct
NCT03448393: CD19/22.41BBz configured in a tandem structure

1.5 Bioinformatics & Data Analysis

AE data, patient demographics, and disease information were loaded from individual patient Excel files into Python 3.9 for preprocessing and statistical analysis. The ‘pandas’ library was used to read and load data, and ‘numpy’ library was used for subsequent numerical and statistical analyses (median, IQR, range, etc). Additional statistical analyses were done using SAS version 9.4 and GraphPad Prism 9.0. Figures were generated using GraphPad Prism 9.0 and BioRender.

1.6 Secondary “Focused” Analysis

1.6.1 Elimination of duplicates

Cumulative NC AE data was edited to focus the analysis to those NC AEs of the highest grade per AE term per patient to eliminate duplicates. If a patient had multiple of the same AE term at the same grade, the one that lasted the longest was chosen; if duration was the same, the event with the later onset was selected.

1.6.2 Elimination of terms with overlapping significance

Following elimination of duplicates, AEs were reviewed and any deemed to have overlapping significance for a patient at a given point in time were narrowed down to the single most representative term. For example, in one patient with hypoxia, pulmonary edema and respiratory failure, hypoxia and pulmonary edema were removed as it was felt that respiratory failure captured those terms. 12 patients were found to have overlapping terms requiring elimination of 1-4 terms.

1.6.3 Determination of Duration

Duration of each AE term was calculated using date of onset and date of resolution. If the AE end date was later than day +30 or if the AE had no end date, it was resolved as of day +30. Of note, 3 NC AEs were removed due to onset that occurred after day +30.

1.6.4 Univariate and Multivariable Analyses

The same set of univariate and multivariable analyses that were completed for the cumulative set of NC AEs as described in the methods of the primary manuscript were repeated in the focused set of NC AEs.

2. Supplemental Results

2.1 Supplemental Tables

Supplemental Table 1: Changes Between CTCAE versions 4 and 5 and Potential Impact on Findings

CTCAE Toxicity Term	**Changes between v4 and v5**	**Impact on AE assessment**
Hypophosphatemia	v4 was based on lab values whereas v5 defines grade 3 and 4 by medical significance, hospitalization required or prolonged or life-threatening consequences	May lead to a lower number of events captured by v5
Hyponatremia	Grade 3 in v4 was based on lab values alone (120 to <130) whereas grade 3 in v5 required symptoms for lab values between 125-129; no change in definition of grade 4	May lead to a lower number of grade 3 events captured by v5
Hyperglycemia	v4 was based on lab values whereas v5 defines grade 3 and 4 by insulin therapy initiated, hospitalization required or life-threatening consequences	May lead to lower capture of events on v5 if hyperglycemic without initiation of insulin, v4 if insulin initiated below lab threshold would still be captured by term “glucose intolerance”
Hyperuricemia	v4 was based on lab values grade 3 ULN to 10mg/dL, grade 4 > 10mg/dL plus physiologic consequences whereas v5 defines grade 3 by greater	May lead to a lower number of events captured by v5; grade 4 per v4 by lab value alone could be considered

	than ULN with physiologic consequences and grade 4 just by life-threatening consequences	grade 3 by v5 if value > 10mg/dL with physiologic but not life-threatening consequences
Alanine aminotransferase increased	Both v4 and 5 are based on lab values, however, v4 is based on the ULN, while v5 is based on either ULN or baseline value if abnormal	Patients with baseline abnormal (elevated) value may be captured by v4 but either not by v5 or at lower grade by v5
Aspartate aminotransferase increased	Both v4 and 5 are based on lab values, however, v4 is based on the ULN, while v5 is based on either ULN or baseline value if abnormal	Patients with baseline abnormal (elevated) value may be captured by v4 but either not by v5 or at lower grade by v5
Alkaline phosphatase increased	Both v4 and 5 are based on lab values, however, v4 is based on the ULN, while v5 is based on either ULN or baseline value if abnormal	Patients with baseline abnormal (elevated) value may be captured by v4 but either not by v5 or at lower grade by v5
Blood bilirubin increased	Both v4 and 5 are based on lab values, however, v4 is based on the ULN, while v5 is based on either ULN or baseline value if abnormal	Patients with baseline abnormal (elevated) value may be captured by v4 but either not by v5 or at lower grade by v5
GGT increased	Both v4 and 5 are based on lab values, however, v4 is based on the ULN, while v5 is based on either ULN or baseline value if abnormal	Patients with baseline abnormal (elevated) value may be captured by v4 but either not by v5 or at lower grade by v5
Electrocardiogram QT corrected interval prolonged	In v4, grade 3 is based on QTc $\geq$501 on at least 2 ECGs, whereas in v5 grade 3 is an average QTc $\geq$ 501 or >60ms change from baseline. In v4, grade 4 is QTc $\geq$ 501 or >60ms change from baseline and Torasade de pointes or polymorphic ventricular tachycardia or signs/symptoms of serious arrhythmia, whereas in v5 grade 4 is not based on QTc and only Torasade de pointes or polymorphic ventricular tachycardia or signs/symptoms of serious arrhythmia	May lead to a higher number of grade 3 events and a lower number of grade 4 events captured by v5
Hypertension	Grade 3 in v5 specifies a systolic/diastolic value for children/adolescents whereas v4 combines adults and	v4 may lead to fewer grade 3 events, missing the capture of young pediatric patients with blood pressures above

	children/adolescents, and no change in definition of grade 3 values for adults or grade 4 definitions	the 99th percentile but not above adult thresholds
Vascular access complication	In v5, pulmonary embolism is grade 3 whereas in v4 it was a grade 4 event; v5 also adds hemodynamic or neurologic instability to grade 4	Pulmonary embolism would be captured as grade 4 on v4 but grade 3 on v5 unless life threatening
INR increased	v4 is based on lab values to a degree above the ULN or baseline if on anticoagulation, whereas v5 uses distinct lab values or degree above baseline if on anticoagulation	May lead to a higher number of grade 3 events captured by v5 if defined ULN is >1
Serum amylase increased	v5 adds "with signs or symptoms", also v5 includes >5x ULN and asymptomatic as grade 3, whereas in v4 any value >5x ULN would be grade 4	Patients with value >2-5x ULN without symptoms would be grade 3 per v4 but grade 2 per v5. Patients with value >5x ULN without symptoms would be grade 4 per v4 but grade 3 per v5
Lipase increased	v5 adds "with signs or symptoms", also v5 includes >5x ULN and asymptomatic as grade 3, whereas in v4 any value >5x ULN would be grade 4	Patients with value >2-5x ULN without symptoms would be grade 3 per v4 but grade 2 per v5. Patients with value >5x ULN without symptoms would be grade 4 per v4 but grade 3 per v5
Diarrhea	Grade 3 in v4 includes "incontinence" whereas grade 3 in v5 does not, and no change in definition of grade 4 values	May lead to a lower number of grade 3 events captured by v5
Sepsis	There is no definition for grade 3 in v4 whereas v5 defines grade 3 by a positive blood culture and treatment indications, and no change in definition of grade 4 values	May lead to a higher number of grade 3 events captured by v5
Acute kidney injury	v4 incorporated creatinine in definition of AKI; whereas v5 separates out creatinine as a unique AE and identifies only $\geq$ Gr 3 AKI	As both would be captured in "renal" toxicity, there is no impact on analysis
Proteinuria	Grade 3 in v5 adds 4+ proteinuria in addition to $\geq$3.5g/24hrs to adult definition	May lead to a higher number of grade 3 events in adults captured by v5
Urine output decreased	v5 added specific child/infant definitions to both grades 3 (weight and body surface area based) and 4	Pediatric patients with lower weight may not have been

	(pediatric definition with no urine output over 12 hours), whereas v4 had one overall definition	captured in v4 but would be captured as grade 3 in v5
Intracranial hemorrhage	Grade 3 in v5 includes hospitalization and uses the term "invasive intervention" whereas v4 uses the term "operative intervention," and no change in definition of grade 4 values	May lead to a higher number of grade 3 events captured by v5 if hospitalized but no intervention indicated
Delirium	Both grades 3 and 4 in v4 include "hospitalization indicated" whereas both grades 3 and 4 in v5 instead state "urgent intervention indicated," Grade 3 in v5 includes "new onset" whereas grade 3 in v4 does not include "new onset"	Urgent intervention could mean medication rather than hospitalization, so v5 may capture more events; however, v4 could capture more grade 3 that aren't considered "new onset"
CTCAE v4 was used for CD19- and CD22-targeted CAR T-cell trials (n=118); v5 was used for CD19/22-targeted CAR T-cell trial (n=16) Abbreviations: CTCAE = Common Terminology Criteria for Adverse Events, v4= Version 4, v5 = Version 5, AE = adverse event, ULN = upper limit of normal, AKI = acute kidney injury		

Supplemental Table 2: Adverse Event Categories with Included CTCAE Terms

AE Category	**CTCAE Terms included**
Thrombocytopenia	Platelet count decreased
Neutropenia	Neutrophil count decreased
Anemia	Anemia
Metabolic	Hypophosphatemia, Hypokalemia, Hypocalcemia, Hypoalbuminemia, Hyponatremia, Hypomagnesemia, Hyperglycemia, Hypermagnesemia, Hypertriglyceridemia, Hyperkalemia, Hypercalcemia, Hypernatremia, Hyperuricemia, Acidosis, Anorexia, Weight gain, Tumor lysis syndrome
Hepatic	Alanine aminotransferase increased, Aspartate aminotransferase increased, Alkaline phosphatase increased, Blood bilirubin increased, GGT increased
Febrile Neutropenia	Febrile neutropenia
Cardiovascular	Cardiac arrest, Capillary leak syndrome, Ejection fraction decreased, Electrocardiogram QT corrected interval prolonged, Hypertension, Hypotension, Left ventricular systolic dysfunction, Mitral valve disease, Sinus tachycardia, Tricuspid valve disease, Vascular access complication
Coagulopathy	Activated partial thromboplastin time prolonged, Disseminated intravascular coagulation, Fibrinogen decreased, INR increased
Respiratory	Adult respiratory distress syndrome, Apnea, Dyspnea, Hypoxia, Pulmonary edema, respiratory failure, Respiratory thoracic and mediastinal disorders-Other
GI	Abdominal distension, Serum amylase increased, Lipase increased, Diarrhea

Infection	Catheter related infection, Enterocolitis infectious, Infections and infestations-Other, Lung infection, Sepsis, Typhlitis
Renal	Acute kidney injury, Creatinine increased, Hemolytic uremic syndrome, Proteinuria, Urine output decreased
Pain	Abdominal pain, Back pain, Bone pain, Pain, Pain in extremity
Neurologic	Dysphagia, Encephalopathy, Headache, Hydrocephalus, Intracranial hemorrhage, Muscle weakness left-sided, Somnolence
Immunologic	Immune system disorders - Other
Psychiatric	Agitation, Delirium
Musculoskeletal	Joint effusion, Musculoskeletal and connective tissue disorder-Other
Abbreviations: AE = adverse event, CTCAE = Common Terminology Criteria for Adverse Events, GGT = gamma-glutamyl transferase, INR = international normalized ratio, GI = gastrointestinal	

Supplemental Table 3: Number of Patients with AEs Per System

System	# Patients with 3+ event (Based on cumulative analysis)	# Patients with 3+ event (Based on restricted analysis)
All	109	108
Metabolic	84	83
Febrile neutropenia	79	76
Hepatic	42	42
CV	32	32
Coagulopathy	22	21
Respiratory	13	13
Infection	10	10
Pain	8	8
GI	7	7
Renal	5	5
Neuro	5	5
Immuno	3	3
Psych	2	1
MSK	1	1
Abbreviations: CV = Cardiovascular, GI = Gastrointestinal, Neuro = Neurologic, Immuno = Immunologic, Psych = Psychiatric, MSK = Musculoskeletal		

Supplemental Table 4: Dose Limiting and Treatment Limiting Toxicities (DLTs & TLTs)

Patient	**# DLTs**	**CTCAE Term**	**Grade**	**Number of Events**	**CR (Y/N)**	**CRS (Y/N)**
1907	2	Hypotension Hypotension	3 4	1 1	Y	Y
1910	1	Dysphagia	3	1	Y	Y
1914	12	Cardiac arrest^	4 3	1 2	Y	Y

		Electrocardiogram QT corrected interval prolonged Hypotension Hypotension Hypoxia Hypoxia Left ventricular systolic dysfunction Left ventricular systolic dysfunction Pulmonary edema Pulmonary edema Respiratory failure	3 4 3 4 3 4 3 4 4	1 1 1 1 1 1 1 1 1		
1916	1	Hypotension^	3	1	Y	Y
2202	2	Diarrhea^ Hypophosphatemia	3 4	1 1	Y	Y
2210	3	Hypoxia Hypoxia^	3 4	2 1	N	Y
2216	25	Alanine aminotransferase increased Aspartate aminotransferase increased Aspartate aminotransferase increased Blood bilirubin increased Cardiac arrest Catheter related infection Ejection fraction decreased Ejection fraction decreased Electrocardiogram QT corrected interval prolonged GGT increased Hypertriglyceridemia Hypertriglyceridemia Hypotension Lipase increased Lipase increased	4 3 4 3 4 4 3 4 3 4 4 4 3 4 3 4 3	1 1 2 1 1 1 1 2 1 1 1 2 1 1 1 1 1	Y	Y

		Respiratory failure Sepsis Serum amylase increased Serum amylase increased Typhlitis Urine output decreased Weight gain				
2219	3	Hypertriglyceridemia Hypertriglyceridemia	3 4	2 1	N	Y
2221	2	Hypertriglyceridemia	3	2	Y	Y
2224	6	Acidosis Adult respiratory distress syndrome Dyspnea Hypotension Hypoxia	3 5 3 3 4	1 1 1 1 1	N/A*	Y
2231	1	Hypokalemia	4	1	Y	Y
2232	8	Alanine aminotransferase increased Aspartate aminotransferase increased Aspartate aminotransferase increased Disseminated intravascular coagulation Hypertension Hypertriglyceridemia Hypokalemia	3 3 4 3 3 4 4	1 1 2 1 1 1 1	Y	Y
2233	3	Alanine aminotransferase increased Disseminated intravascular coagulation Immune system disorder - Other	4 3 3	1 1 1	Y	Y
2242	1	Hypoxia	3	1	N	Y
2244	1	Hypoxia	3	1	Y	Y
2251	2	Acute kidney injury	4	1	Y	Y

		Hypoxia	3	1		
2263		Intracranial hemorrhage	4	1	Y	Y
192214		Encephalopathy^	3	1	Y	Y

^Dose Limiting Toxicity (DLT)
*Patient unable to undergo disease assessment due to early death
Abbreviations: CTCAE = Common Terminology Criteria for Adverse Events, DLT = Dose Limiting Toxicity, TLT = Treatment Limiting Toxicity, CR = Complete Response, CRS = Cytokine Release Syndrome

Supplemental Table 5: Comparison of Cumulative NC AE vs Focused NC AE Analyses

Measure	Cumulative Analysis	Focused Analysis	Notes
Incidence			
Incidence All	All: 1719 Patients: 133 (99.3%) Median: 10 IQR: 4-19	All: 702 events Patients: 133 (99.3%) Median: 4.5 IQR: 2-7	1017 fewer AEs (18 AEs from cumulative analysis removed in focused analysis for reasons other than duplication: onset after day +3, not attributed to research or disease)
Incidence NC	NC: 737 Patients: 109 (81.3%) Median: 3 IQR: 1-8	NC: 457 Patients: 108 (80.6%) Median: 3 IQR: 1-5	280 fewer AEs (6 AEs from cumulative analysis removed in focused analysis due to reasons other than duplication: onset after day +3, not attributed to research or disease) 1 fewer patient with NC AEs in focused analysis as 1 patient was found to have NC AEs with onset after day +30 so these were removed
Across trials	CD19: 193 events Median 3, IQR 0-6 CD22: 512 events Median: 5, IQR 2-10	CD19: 109 events Median 2, IQR 0-3 CD22: 324 events Median 4, IQR 2-6	No major change, same degree of significant differences across trials

	CD1922: 32 events Median: 1, IQR 0-3.75 Kruskal-Wallis: **p < 0.0001** Mann-Whitney: CD19 vs CD22: **p =0.0010** CD1922 vs CD22: **p = 0.0002** CD19 vs CD1922: p=0.22	CD1922: 24 events Median 1, IQR 0-2.75 Kruskal-Wallis: **p < 0.0001** Mann-Whitney: CD19 vs CD22: **p <0.0001** CD1922 vs CD22: **p = 0.0002** CD19 vs CD1922: p= 0.31	
Outcomes			
CRS vs No CRS	CRS N=104 NC events: 685 Median: 4.5 IQR: 1.25-9 No CRS N=30 NC events: 52 Median: 1 IQR: 0-3 CRS vs No CRS: Mann-Whitney: **p <0.0001**	CRS N=104 NC events: 423 Median: 3 IQR: 1-5.75 No CRS N=30 NC events:34 Median: 1 IQR: 0-2 CRS vs No CRS: Mann-Whitney: **p <0.0001**	Similar, no major difference
CRS Max Grade	Max CRS 1-2 N=81 NC events: 417 Median: 3 IQR: 1-8.5 Max CRS 3-4 N=23 NC events: 268 Median: 7 IQR: 5-12 Max CRS 1-2 vs 3-4 Mann-Whitney: **p=0.0002**	Max CRS 1-2 N=81 NC events: 263 Median: 3 IQR: 1-5 Max CRS 3-4 N=23 NC events: 160 Median: 5 IQR: 4-8 Max CRS 1-2 vs 3-4 Mann-Whitney: **p=0.0001**	Similar, no major difference
CR vs No CR	CR N=88	CR N=88	Focused analysis demonstrated a more

	NC events: 538 Median: 4 IQR: 1-8 No CR: N=45 NC events: 189 Median: 3 IQR: 0-5.5 CR vs no CR Mann-Whitney: p=0.10	NC events: 334 Median: 3 IQR: 1-5 No CR: N=45 NC events: 115 Median: 2 IQR: 0-4 CR vs no CR Mann-Whitney: **p=0.037**	significant increase in NC AEs in patients who achieved a CR versus those who did not compared to cumulative analysis
CR and CRS	+CR+CRS N=86 NC events: 537 Median: 4 IQR: 1-8.25 +CR-CRS N=2 NC events: 1 Median: 0.5 IQR: 0-1 -CR+CRS N=17 NC events: 138 Median: 7 IQR: 4-11 -CR-CRS N=28 NC events: 51 Median: 1 IQR: 0-3 Kruskal-Wallis: **p <0.0001** Mann Whitney +CR+CRS vs -CR+CRS: p = 0.071 Mann Whitney -CR+CRS vs -CR-CRS: **p<0.0001**	+CR+CRS N=86 NC events: 333 Median: 3 IQR: 1-5 +CR-CRS N=2 NC events: 1 Median: 0.5 IQR: 0-1 -CR+CRS N=17 NC events: 82 Median: 4 IQR: 3-6.5 -CR-CRS N=28 NC events: 33 Median: 1 IQR: 0-2 Kruskal-Wallis: **p <0.0001** Mann Whitney +CR+CRS vs -CR+CRS: p = 0.091 Mann Whitney -CR+CRS vs -CR-CRS: **p<0.0001**	Similar, no major difference
Disease Burden			

Bone Marrow Disease	M1 <5% N=43 NC events: 113 Median: 1 IQR: 0-3 M2 5-25% N=23 NC events: 92 Median: 2 IQR: 1-4 M3 >25% N=68 NC events: 532 Median: 5.5 IQR: 3-10 Kruskal-Wallis: **p<0.0001** Mann-Whitney M1 vs M2: p=0.21 Mann-Whitney M1 vs M3: **p<0.0001** Mann-Whitney M2 vs M3: **p=0.0055**	M1 <5%: N=43 NC events: 76 Median: 1 IQR: 0-3 M2 5-25% N=23 NC events: 60 Median: 2 IQR: 1-4 M3 >25% N=68 NC events: 321 Median: 4 IQR: 2-6 Kruskal-Wallis: **p <0.0001** Mann Whitney M1 vs M2: p=0.31 Mann Whitney M1 vs M3: **p<0.0001** Mann Whitney M2 vs M3: **p=0.0066**	Same findings, significantly higher NC AEs in patients with higher BM disease; particularly those with M3 compared to M2 and M1
CNS Disease	CNS2+ N=7 NC events: 47 Median: 5 IQR: 1-14 CNS1 N=127 NC events: 690 Median: 3 IQR: 1-7 Mann-Whitney: p=0.43	CNS2+ N=7 NC events: 24 Median: 3 IQR: 1-6 CNS1 N=127 NC events: 433 Median: 3 IQR: 1-5 Mann-Whitney: p=0.65	No difference, still not significant
EMD	EMD N= 24 NC events: 160 Median: 5 IQR: 1.5-10.5 No EMD N=110	EMD N= 24 NC events: 88 Median: 3.5 IQR: 1-5 No EMD N=110	No difference, still not significant

	NC events: 577 Median: 3 IQR: 1-7 Mann-Whitney: p=0.14	NC events: 369 Median: 2.5 IQR: 1-4.25 Mann-Whitney: p=0.31	
Baseline Demographics			
Age	<12 years N=45 NC events: 197 Median: 3 IQR: 1-6 12-17 years N=34 NC events: 157 Median 2 IQR: 1-5.25 18+ years N=55 NC events: 383 Median: 5 IQR: 1-9 Kruskal-Wallis: p=0.13 Mann-Whitney <12 vs 12-17: p=0.42 Mann-Whitney <12 vs 18+: p=0.15 Mann-Whitney 12-17 vs 18+: p=0.075	<12 years N=45 NC events: 141 Median: 3 IQR: 1-4.5 12-17 years N=34 NC events: 95 Median 2 IQR: 1-4 18+ years N=55 NC events: 221 Median: 3 IQR: 1-5 Kruskal-Wallis: p=0.39 Mann-Whitney <12 vs 12-17: p=0.33 Mann-Whitney <12 vs 18+: p=0.58 Mann-Whitney 12-17 vs 18+: p=0.20	No difference, still not significant
Ethnicity	Non-Hispanic N=94 NC events: 570 Median: 3 IQR: 1-9 Hispanic N=40 NC events: 167 Median: 3.5 IQR: 1-5.75 Mann-Whitney: p=0.67	Non-Hispanic N=94 NC events: 349 Median: 3 IQR: 1-6 Hispanic N=40 NC events: 108 Median: 2 IQR: 1-4 Mann-Whitney: p=0.47	No difference, still not significant
Race	White N=96	White N=96	No difference, still not significant

	NC events: 545 Median: 3 IQR: 1-7.75 Black N=5 NC events: 22 Median: 3 IQR: 0.5-9 Asian N=10 NC events: 50 Median: 4.5 IQR: 1.5-9.25 Hawaiian/PI N=2 NC events: 8 Median: 4 IQR: 0-8 Multi-Race N=6 NC events: 45 Median: 6 IQR: 1.5-12.25 Unknown N=15 NC events: 67 Median: 3 IQR: 1-4 Kruskal-Wallis: p=0.95 *Race by White vs Non-White* White N=96 NC events: 545 Median: 3 IQR: 1-7.75 Non-White N=23	NC events: 336 Median: 3 IQR: 1-4.75 Black N=5 NC events: 14 Median: 2 IQR: 0.5-5.5 Asian N=10 NC events: 37 Median: 3.5 IQR: 1.5-5.5 Hawaiian/PI N=2 NC events: 6 Median: 3 IQR: 0-6 Multi-Race N=6 NC events: 25 Median: 3.5 IQR: 0.75-7.75 Unknown N=15 NC events: 39 Median: 2 IQR: 1-3 Kruskal-Wallis: p=0.86 *Race by White vs Non-White* White N=96 NC events: 336 Median: 3 IQR: 1-4.75 Non-White N=23	

	NC events: 125 Median: 5 IQR: 1-9 Mann Whitney p=0.58	NC events: 82 Median: 3 IQR: 1-6 Mann Whitney p=0.50	
Sex	Female N=37 NC events: 206 Median: 3 IQR: 1-6.5 Male N=97 NC events: 531 Median: 4 IQR: 1-8 Mann-Whitney: p=0.55	Female N=37 NC events: 131 Median: 3 IQR: 1-4.5 Male N=97 NC events: 326 Median: 3 IQR: 1-5 Mann-Whitney: p=0.82	No difference, still not significant
Prior Treatments			
# Prior Therapies	≤2 Regimens N=19 NC events: 68 Median: 2 IQR: 0-6 3-5 Regimens N=60 NC events: 361 Median: 3 IQR: 1-8.75 >5 Regimens N=55 NC events: 308 Median: 4 IQR: 1-7 Kruskal-Wallis: p=0.39 Mann-Whitney ≤2 vs 3-5: p=0.30 Mann-Whitney ≤2 vs >5: p=0.18 Mann-Whitney 3-5 vs >5: p=0.63	≤2 Regimens N=19 NC events: 46 Median: 2 IQR: 0-4 3-5 Regimens N=60 NC events: 223 Median: 2 IQR: 1-4.75 >5 Regimens N=55 NC events: 188 Median: 3 IQR: 1-5 Kruskal-Wallis: p=0.40 Mann-Whitney ≤2 vs 3-5: p=0.42 Mann-Whitney ≤2 vs >5: p=0.17 Mann-Whitney 3-5 vs >5: p=0.50	No difference, still not significant
Prior HSCT	No Prior HSCT N=58	No Prior HSCT N=58	No difference, still not significant

	NC events: 274 Median: 3 IQR: 1-6.25 Prior HSCT N=76 NC events: 463 Median: 4 IQR: 1-8.75 Mann-Whitney: p=0.29	NC events: 187 Median: 2 IQR: 1-4.25 Prior HSCT N=76 NC events: 270 Median: 3 IQR: 1-5 Mann-Whitney: p=0.52	
Prior CAR	No Prior CAR N=80 NC events: 448 Median: 3 IQR: 0.25-8 Prior CAR N=54 NC events: 289 Median: 4 IQR: 1-7 Mann-Whitney: p=0.36	No Prior CAR N=80 NC events: 267 Median: 2 IQR: 0.25-4.75 Prior CAR N=54 NC events: 190 Median: 3 IQR: 1-5 Mann-Whitney: p=0.14	No difference, still not significant
Prior Blina	No Prior Blina N=90 NC events: 472 Median: 3 IQR: 1-7 Prior Blina N=44 NC events: 265 Median: 4 IQR: 1-9.75 Mann-Whitney: p=0.11	No Prior Blina N=90 NC events: 295 Median: 2 IQR: 1-4 Prior Blina N=44 NC events: 162 Median: 3 IQR: 1-5 Mann-Whitney: p=0.11	No difference, still not significant
Prior Ino	No Prior Ino N=111 NC events: 546 Median: 3 IQR: 1-6 Prior Ino N=23 NC events: 191	No Prior Ino N=111 NC events: 345 Median: 2 IQR: 1-4 Prior Ino N=23 NC events: 112	Still significant difference

	Median: 6 IQR: 2-14 Mann-Whitney: **p=0.012**	Median: 4 IQR: 2-7 Mann-Whitney: **p=0.0076**	
Abbreviations: NC = Non-cytopenic, AE = Adverse Event, IQR = Interquartile Range, CRS = Cytokine Release Syndrome, CR = Complete Response, M1 <5%, M2 5-25%, M3 >25%, CNS = Central Nervous System, EMD = Extramedullary Disease, PI = Pacific Islander, HSCT = Hematopoietic Stem Cell Transplant, CAR = Chimeric Antigen Receptor, Blina = Blinatumomab, Ino = Inotuzumab			

Supplementary Table 6: Specific AE Terms Per Clinical System in Focused Analysis

Hepatic	
Alanine aminotransferase increased	28
Aspartate aminotransferase increased	38
Alkaline phosphatase increased	3
Blood bilirubin increased	7
GGT increased	1
Cardiovascular	
Capillary leak syndrome	1
Cardiac arrest	2
Ejection fraction decreased	3
Electrocardiogram QT corrected interval prolonged	4
Hypertension	9
Hypotension	19
Mitral valve disease	1
Sinus tachycarida	2
Tricuspid valve disease	1
Vascular access complication	1
Respiratory	
Adult respiratory distress syndrome	1
Apnea	1
Dyspnea	2
Hypoxia	8
Respiratory failure	2
Respiratory, thoracic and mediastinal disorders - Other, specify	1
Neurologic	
Dysphagia	1

Encephalopathy	1
Headache	2
Intracranial hemorrhage	1
Infection	
Catheter related infection	3
Enterocolitis infectious	1
Infections and infestations - Other, specify	1
Lung infection	2
Sepsis	2
Typhlitis	3
Renal	
Acute kidney injury	3
Hemolytic uremic syndrome	1
Proteinuria	1
Urine output decreased	3

2.2 Supplemental Figure Legends

Supplemental Figure 1. Additional Across Trial Details. (A) Graph of number of patients treated per year on each of the 3 trials. (B) Dot plot of number of ≥Gr3 AEs per system on each of the 3 trials. Median represented by horizontal black line; if no line present, the median is equal to 0. Abbreviations: Gr = grade, NC = non-cytopenia, AE = adverse event, GI = gastrointestinal.

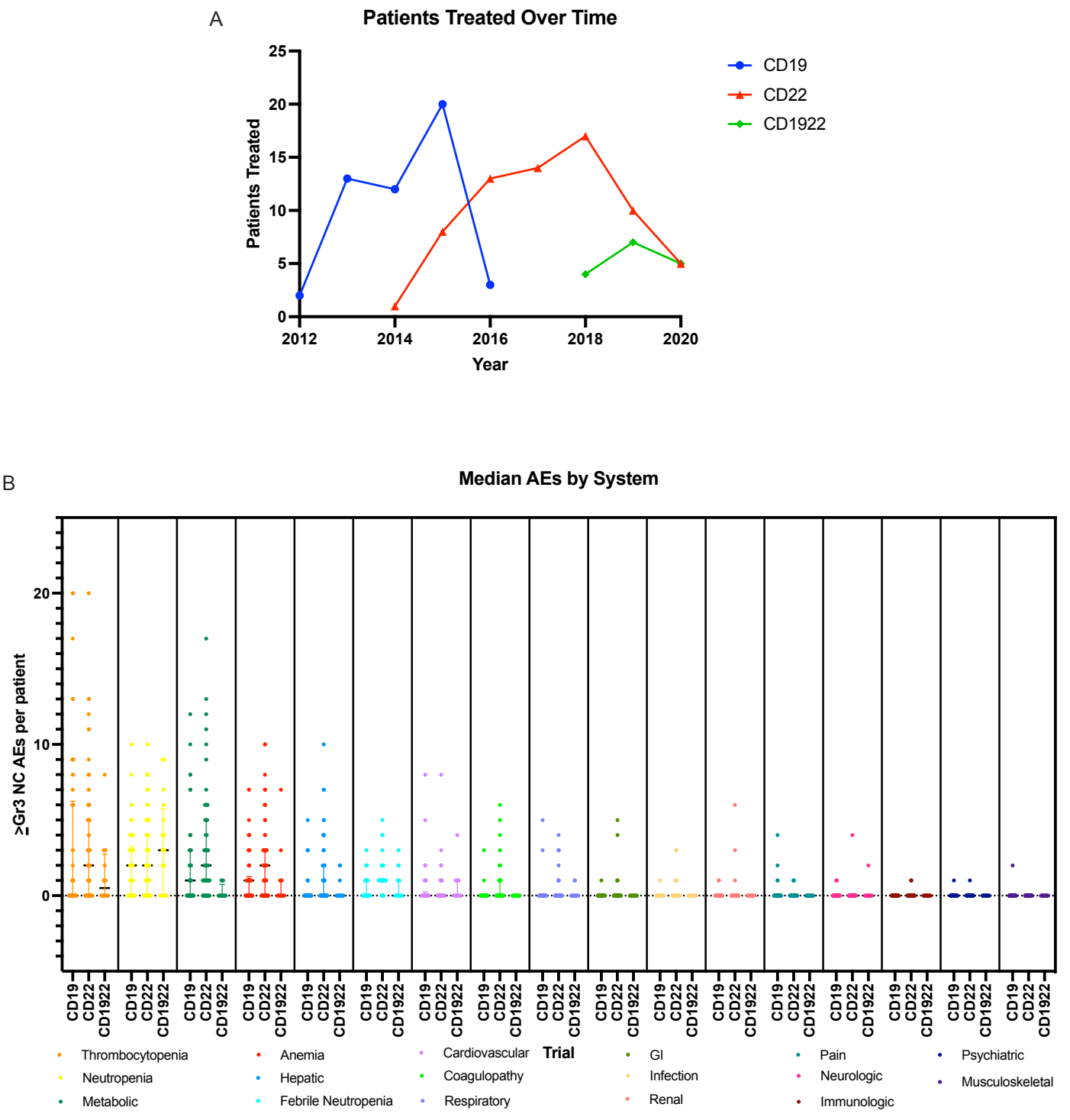
Supplemental Figure 2. System Based Analysis in Specific Sub-groups. (A) Pie graph of ≥Gr3 AEs in patients who received prior inotuzumab. (B) Pie graph of ≥Gr3 NC AEs attributed to disease only—AEs deemed unlikely due to CAR T-cell therapy and likely due to underlying leukemia. Abbreviations: Gr = grade, AE = adverse event, GI = gastrointestinal

Supplemental Figure 3. Multivariable Analysis of Pre-CAR Factors. (A) Linear regression predictive model for patients on CD19 trial. Graph demonstrates predicted versus actual values of ≥Gr3 NC AEs based on formula 0.20* age + 0.034 * actual (0-99) marrow disease level; adjusted r2 value 0.46. (B) Linear regression predictive model for patients on CD22 trial. Graph demonstrates predicted versus actual values of ≥Gr3 NC AEs based on formula -3.87 (if prior CAR) + 4.07*(1, 2, or 3 depending on marrow disease).; adjusted r2 value 0.51.

Supplemental Figure 4. Incidence of Focused >Gr3 NC AEs Across Trials and based on CRS, Disease Burden and Response. (A) & (B) Dot plot of the number of >Gr3 AEs (and >Gr3 NC AEs) per patient by trial with horizontal line representing the median per patient. Pair-wise rank comparisons by Mann-Whitney are represented by p values at the top of graphs, with Kruskal-Wallis comparisons across all 3 trials to the right. (C) Dot plot of >Gr3 NC AEs in patients based on presence of CRS. (D) Dot plot of >Gr3 NC AEs in patients based on severity of CRS max grade 1-2 vs 3-4 as graded per ASTCT guidelines. (E) Dot plot of >Gr3 NC AEs in patients based on baseline bone marrow disease; M1 <5%, M2 5-25%, M3 >25%. (F) Dot plot of >Gr3

NC AEs in patients based on complete response (CR) compared to those without complete response (No CR). (G) Dot plot of >Gr3 NC AEs in patients based on both CRS and response. +CR/+CRS = complete response with CRS, -CR/+CRS= no complete response with CRS, -CR/-CRS= no complete response without CRS, +CR/-CRS= complete response without CRS. Abbreviations: Gr = grade, NC = non-cytopenia, AE = adverse event, CRS = cytokine release syndrome, CR = complete response

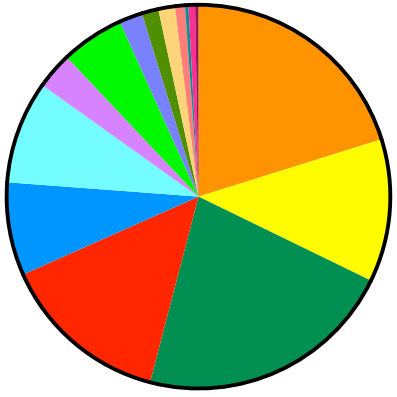
Supplemental Figure 5. Incidence of >Gr3 NC AEs in patients Based on Pre-CAR Factors in Focused Analysis. (A) Dot plot of >Gr3 NC AEs in patients by presence of extramedullary disease (EMD). (B) Dot plot of >Gr3 NC AEs in patients by presence of central nervous system (CNS) disease. CNS1=0 blasts on cytospin, CNS2+ = some degree of CNS disease, either CNS2 (< 5/µl WBCs, cytospin positive for blasts) or CNS3 (≥ 5/µl WBCs, cytospin positive for blasts). (C) Dot plot of >Gr3 NC AEs in patients by age. (D) Dot plot of >Gr3 NC AEs in patients by Sex. (E) Dot plot of >Gr3 NC AEs in patients by Ethnicity. (F) Dot plot of >Gr3 NC AEs in patients by Race; due to small numbers analysis was grouped into White vs Non-White, with Non-White including: African American, Asian, Hawaiian/PI, Multi-Race; 15 patients with Unknown race were excluded from this analysis. (G) Dot plot of >Gr3 NC AEs in patients by number of prior treatment regimens excluding prior HSCT. (H) Dot plot of >Gr3 NC AEs in patients by prior HSCT. (I) Dot plot of >Gr3 NC AEs in patients by prior CAR T-Cell Therapy. (J) Dot plot of >Gr3 NC AEs in patients by prior receipt of blinatumomab. (K) Dot plot of >Gr3 NC AEs in patients by prior receipt of inotuzumab. Abbreviations: Gr = grade, NC = non-cytopenia, AE = adverse event, EMD = extramedullary disease, CNS = central nervous system, HSCT = hematopoietic stem cell transplant, CAR = chimeric antigen receptor



Supplementary Figure 1

A

≥ Gr 3 AEs in Patients Receiving Prior Inotuzumab

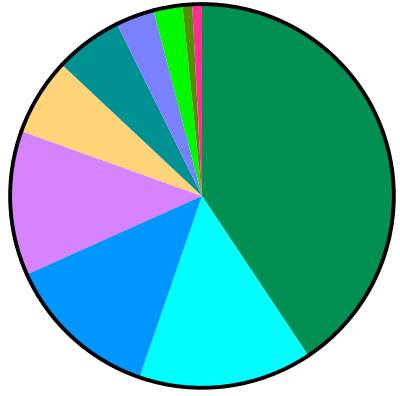


Total=357

- 20.17% Thrombocytopenia
- 12.04% Neutropenia
- 21.85% Metabolic
- 14.29% Anemia
- 7.84% Hepatic
- 8.68% Febrile Neutropenia
- 3.08% Cardiovascular
- 5.32% Coagulopathy
- 1.96% Respiratory
- 1.40% GI
- 1.40% Infection
- 0.84% Renal
- 0.28% Pain
- 0.56% Neurologic
- 0.28% Immunologic

B

Non-Cytopenia ≥ Gr 3 AEs Attributed to Disease Only

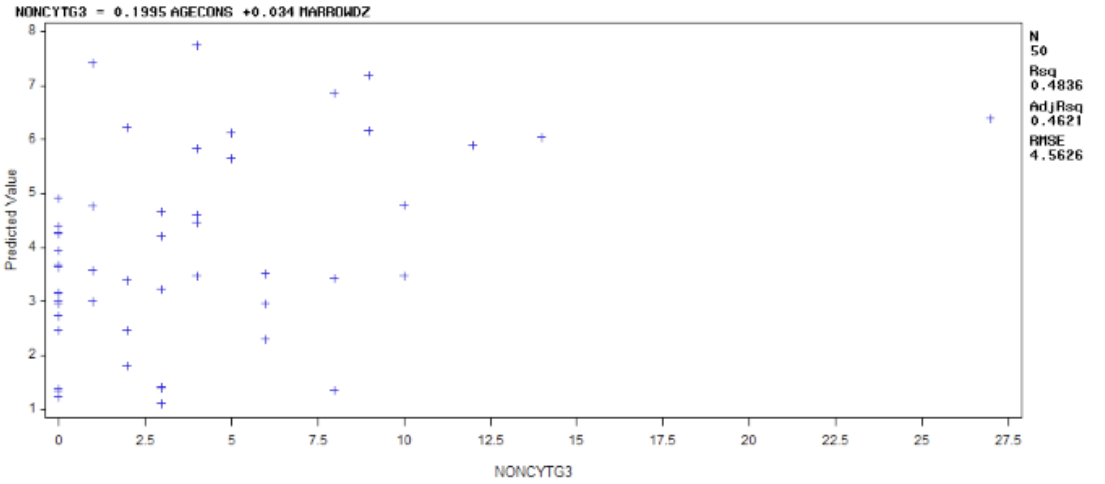


Total=123

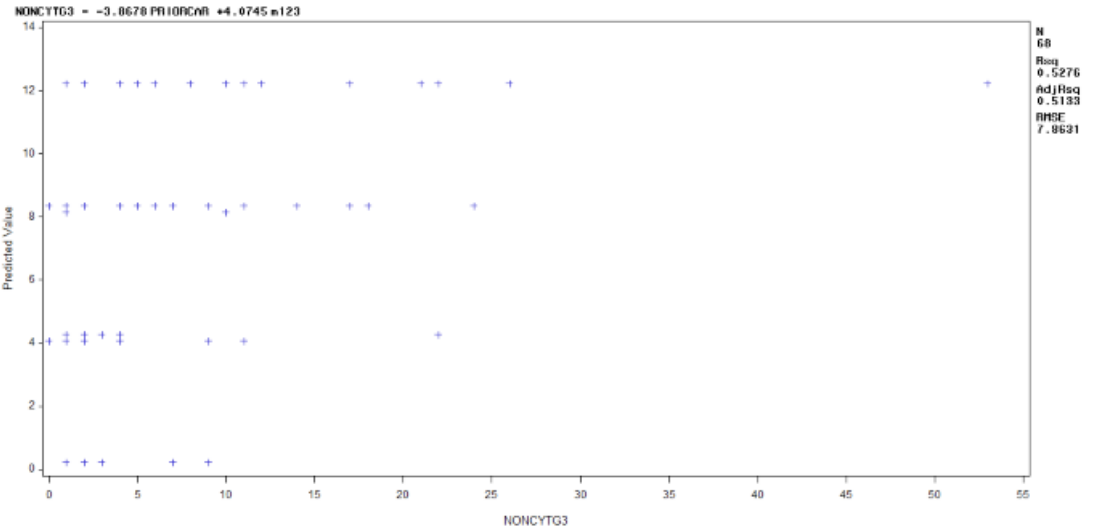
- 40.65% Metabolic
- 14.63% Febrile Neutropenia
- 13.01% Hepatic
- 12.20% Cardiovascular
- 6.50% Infection
- 5.69% Pain
- 3.25% Respiratory
- 2.44% Coagulopathy
- 0.81% GI
- 0.81% Neurologic

Supplementary Figure 2

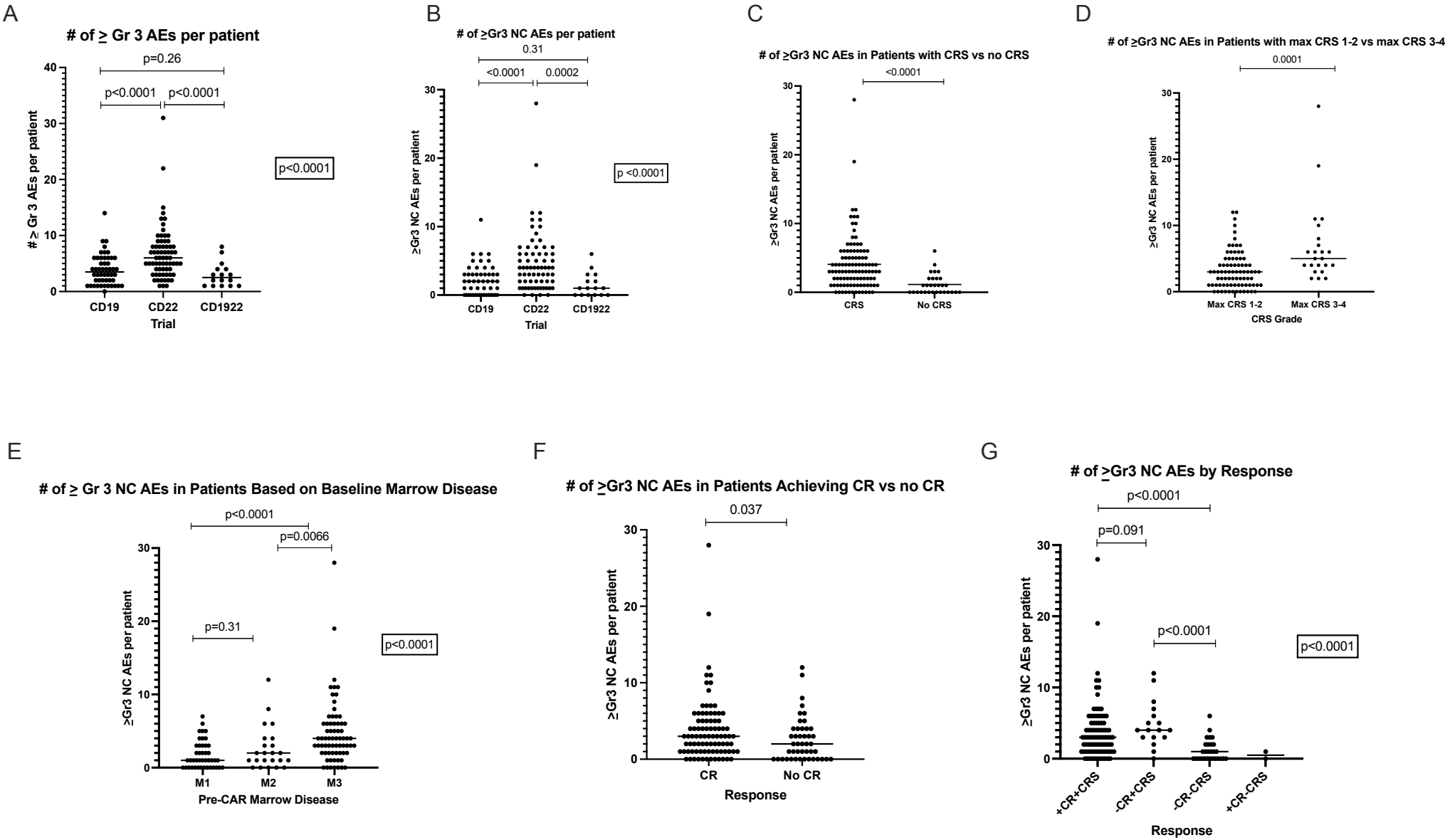
A. Linear Regression Predictive Model, CD19



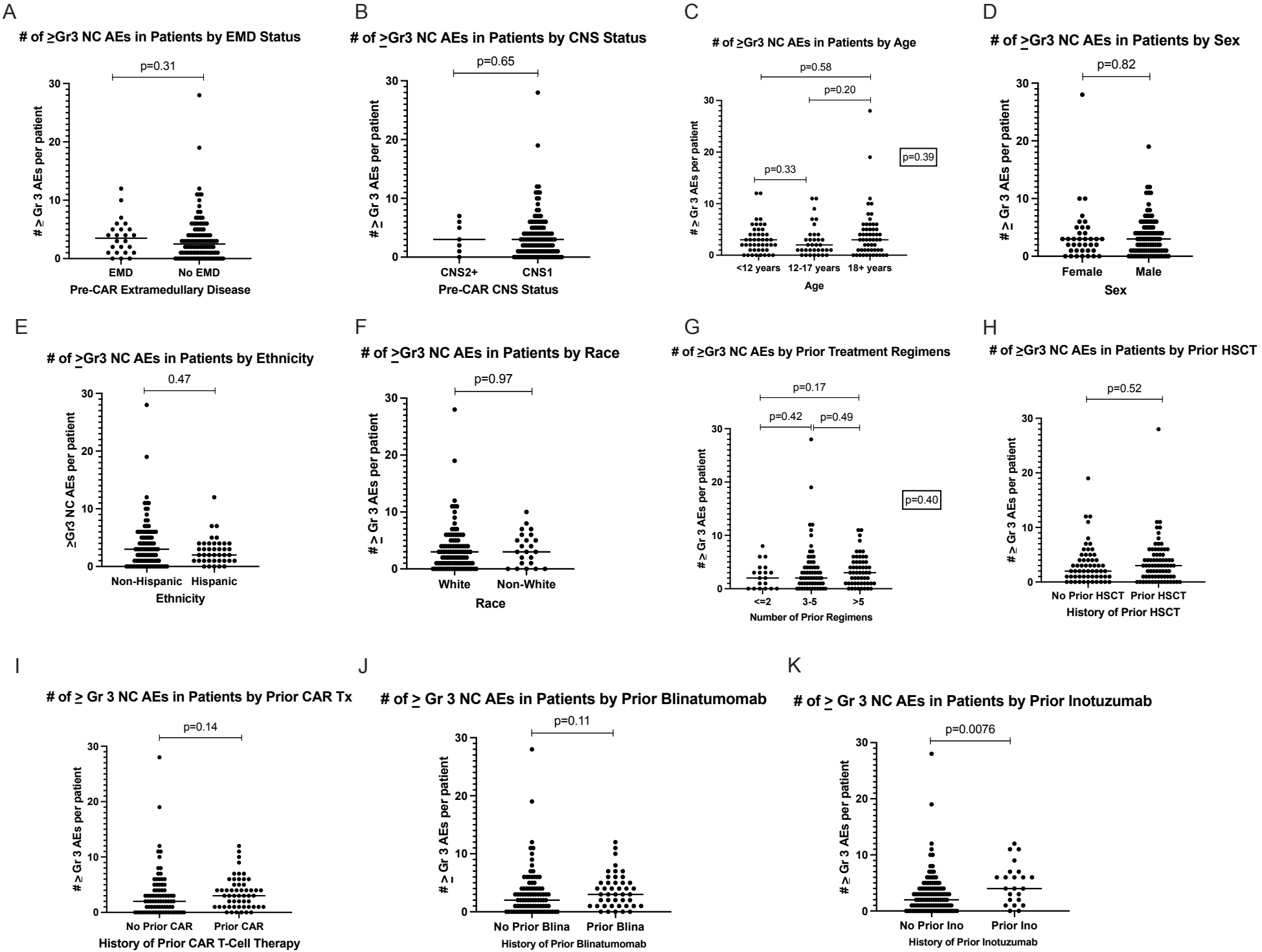
B. Linear Regression Predictive Model, CD22



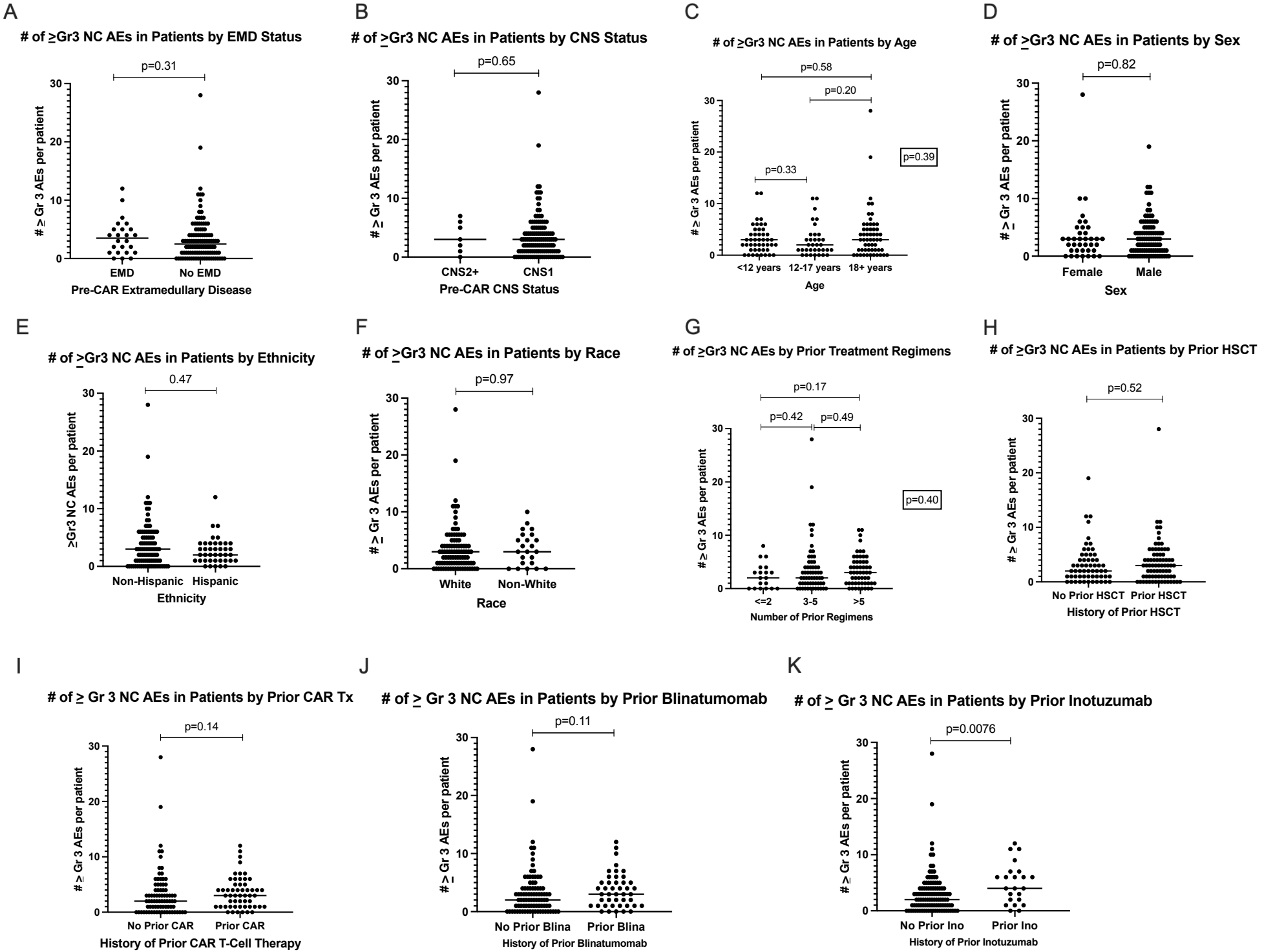
Supplementary Figure 3



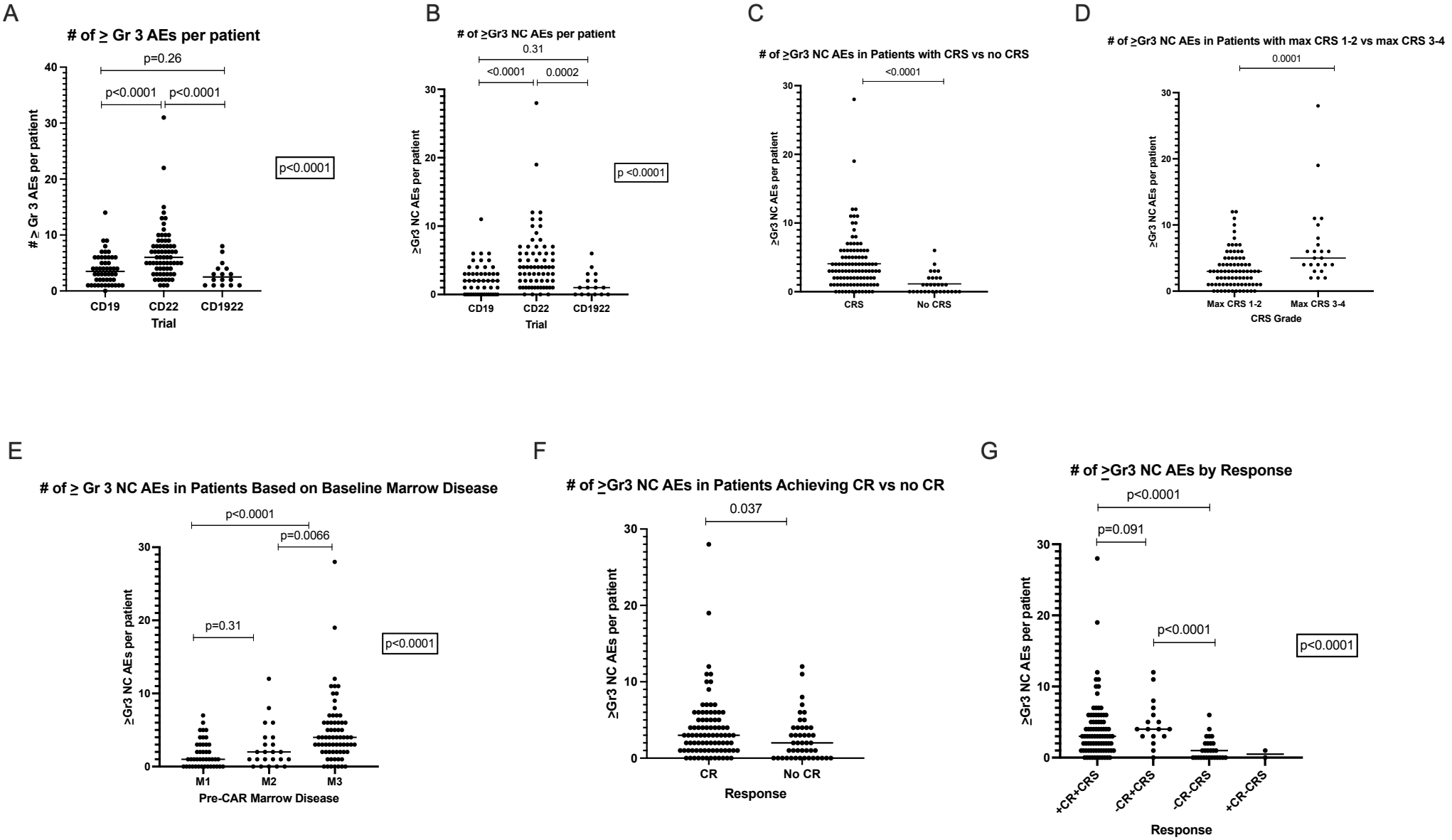
Supplementary Figure 4



Supplementary Figure 5



Supplementary Figure 5



Supplementary Figure 4