

Supplementary material

Table S1. States with US healthcare institutions receiving services from Arkana Laboratories

State
Alabama
Alaska
Arizona
Arkansas
California
Colorado
Connecticut
Florida
Georgia
Hawaii
Idaho
Illinois
Indiana
Iowa
Kansas
Kentucky
Louisiana
Maine
Maryland
Michigan
Minnesota
Mississippi
Missouri
Montana
Nebraska
Nevada
New Jersey
New Mexico
New York
North Carolina
Ohio
Oklahoma
Oregon
Pennsylvania
Rhode Island
South Carolina
Tennessee
Texas
Utah
Virginia
Washington
West Virginia
Wisconsin

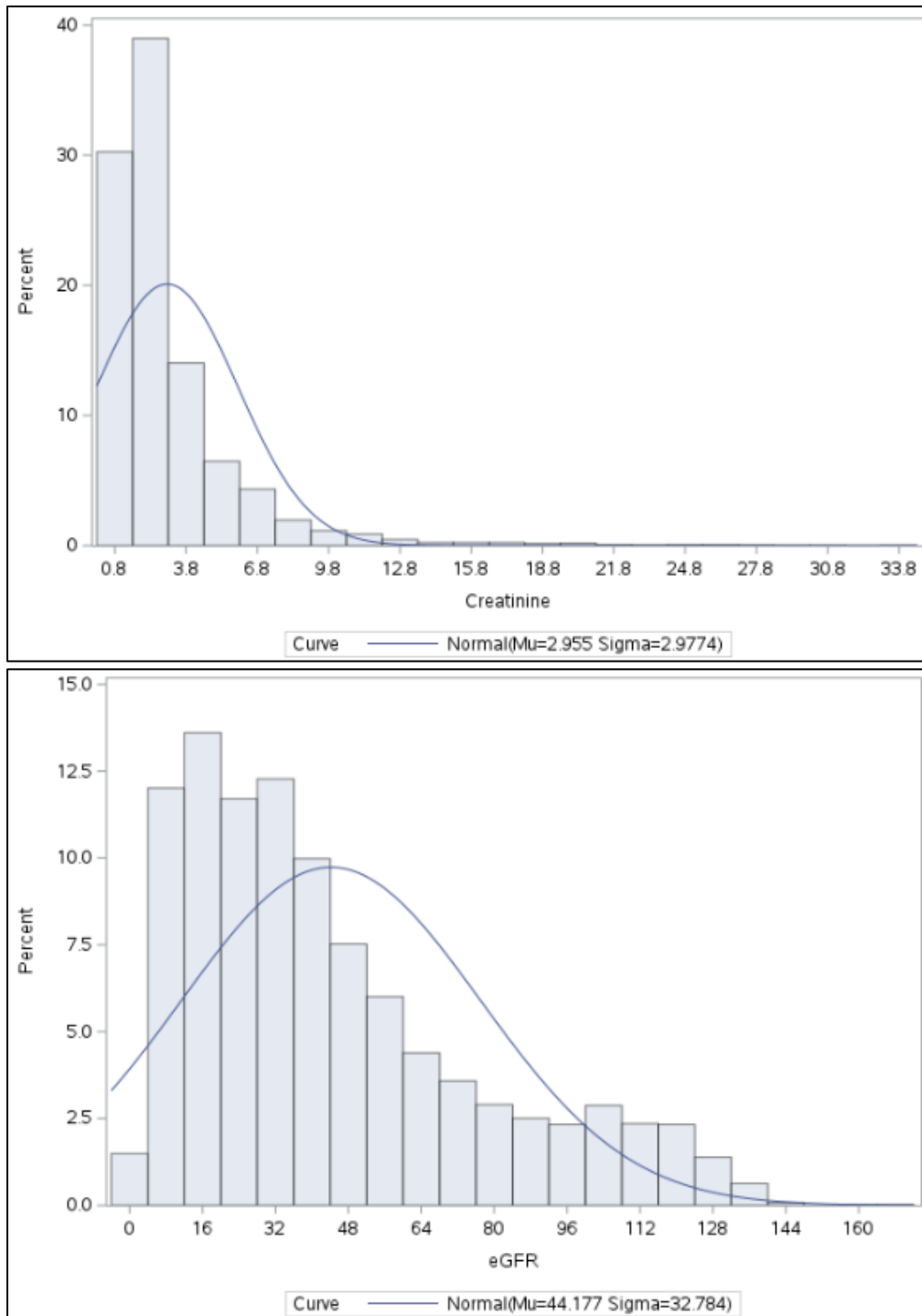


Figure S1. Histograms showing the distribution of patient creatinine (top) and eGFR (bottom) measurements at baseline

Table S2. Histological characteristics for overall cohort determined using immunofluorescence microscopy

Histological characteristic (N=4 375)	Frequency (N)	Percentage (%)
IgA		
1+	53	1.2%
2+	1 376	31.5%
3+	2 945	67.3%
Unknown ^a	1	0.0%
IgG		
Negative/Trace	3 743	85.6%
1+	358	8.2%
2+	217	5.0%
3+	56	1.3%
Unknown ^a	1	0.0%
IgM		
Negative/Trace	3 741	85.5%
1+	430	9.8%
2+	179	4.1%
3+	24	0.6%
Unknown ^a	1	0.0%
Cq1		
Negative/Trace	4 347	99.4%
1+	15	0.3%
2+	11	0.3%
3+	1	0.0%
Unknown ^a	1	0.0%
C3		
Negative/Trace	1 120	25.6%
1+	1 064	24.3%
2+	1 435	32.8%
3+	755	17.3%
Unknown ^a	1	0.0%
Kappa		
Negative/Trace	608	13.9%
1+	906	20.7%
2+	1 582	36.2%
3+	1 278	29.2%
Unknown ^a	1	0.0%
Lambda		
Negative/Trace	172	3.9%
1+	274	6.3%
2+	1 388	31.7%
3+	2 539	58.0%
Unknown ^a	2	0.1%
Fibrinogen		
Negative/Trace	3 656	83.6%
1+	158	3.6%

Histological characteristic (N=4 375)	Frequency (N)	Percentage (%)
2+	358	8.2%
3+	202	4.6%
Unknown ^a	1	0.0%

^aRemoved prior to statistical testing.

Abbreviations: **IgA**, immunoglobulin A; **IgG**, immunoglobulin G; **IgM**, immunoglobulin M.

Table S3. Demographic and clinical characteristics stratified by age and race

Category	Age						Race				
	18–44 years		45–64 years		65+ years		Caucasian		Other ^b		P-value ^a
	N	%	N	%	N	%	N	%	N	%	
	1 995	100.0%	1 566	100.0%	814	100.0%	2 172	100.0%	2 203	100.0%	
Age (years)											
Mean (SD)	32.6 (7.3)		54.0 (5.8)		72.5 (5.9)		48.8 (17.0)		46.5 (16.1)		< 0.001
Median (Q1, Q3)	33 (27, 39)		54 (49, 59)		71 (68, 76)		48 (35, 62)		45 (34, 59)		
Sex											
Female	806	40.4%	558	35.6%	270	33.2%	689	31.7%	944	42.9%	< 0.001
Male	1 190	59.7%	1 008	64.4%	544	66.8%	1 483	68.3%	1 259	57.2%	
Race/Ethnicity											
Caucasian	929	46.6%	790	50.5%	453	55.7%					< 0.001
African American	109	5.5%	82	5.2%	34	4.2%					
Hispanic	117	5.9%	88	5.6%	25	3.1%					
Asian	135	6.8%	79	5.0%	33	4.1%					
Other	63	3.2%	51	3.3%	10	1.2%					
Unknown	642	32.2%	476	30.4%	259	31.8%					
Hypertension	1 384	69.4%	1 184	75.6%	662	81.3%	1 616	74.4%	1 614	73.3%	
Yes	1 112	80.3%	1 103	93.2%	637	96.2%	1423	88.1%	1 429	88.5%	0.671
No	272	19.7%	81	6.8%	25	3.8%	193	11.9%	185	11.5%	
Hematuria	1 433	71.8%	1 033	66.0%	492	60.4%	1480	68.1%	1 478	67.1%	
Yes	1 359	94.8%	956	92.5%	446	90.7%	1394	94.2%	1 367	92.5%	0.064
No	74	5.2%	77	7.5%	46	9.3%	86	5.8%	111	7.5%	
Creatinine (mg/dL)											
Patients with available creatinine measurements	1 726	86.5%	1 366	87.2%	736	90.4%	1 937	89.2%	1 891	85.8%	
Mean (SD)	3.1 (3.7)		2.8 (2.3)		3.0 (2.0)		2.9 (2.8)		3.0 (3.2)		0.091
Median (Q1, Q3)	1.8 (1.1, 3.4)		2.0 (1.5, 3.3)		2.4 (1.7, 3.6)		2.0 (1.3, 3.4)		2.0 (1.3, 3.5)		

eGFR ^c , mean, ml/min/1.73 m ² (SD)	54.5 (38.6)		39.0 (25.6)		29.6 (19.4)		< 0.001	44.4 (32.0)		44.0 (33.6)		0.723
eGFR ^c , median, ml/min/1.73 m ² (Q1, Q3)	45.8 (20.9, 84.6)		34.1 (19.7, 51.7)		25.3 (14.6, 39.0)			36.0 (19.3, 61.0)		34.0 (18.2, 60.2)		
Stage 1	392	22.7%	85	6.2%	14	1.9%	< 0.001	244	12.6%	247	13.1%	0.015
Stage 2	279	16.2%	158	11.6%	44	6.0%		251	13.0%	230	12.2%	
Stage 3	459	26.6%	552	40.4%	239	32.5%		650	33.6%	600	31.7%	
Stage 3A	201	11.6%	210	15.4%	70	9.5%		257	13.3%	224	11.8%	
Stage 3B	258	14.9%	342	25.0%	169	23.0%		393	20.3%	376	19.9%	
Stage 4	287	16.6%	331	24.2%	247	33.6%		438	22.6%	427	22.6%	
Stage 5	309	17.9%	240	17.6%	192	26.1%		354	18.3%	387	20.5%	
Proteinuria												
24-hour test results (g/d)	1 101	55.2%	798	51.0%	381	46.8%		1 169	53.8%	1 111	50.4%	
Mean (SD)	4.0 (3.9)		4.2 (4.0)		4.9 (4.3)		0.001	4.1 (3.7)		4.3 (4.3)		0.107
Median (Q1, Q3)	2.9 (1.3, 5.6)		3.0 (1.5, 5.5)		3.6 (1.6, 7.1)			3.0 (1.4, 5.7)		3.0 (1.5, 5.9)		
≥3.5 g/d	472	42.9%	349	43.7%	197	51.7%	0.009	505	43.2%	513	46.2%	0.153

^aANOVA/Kruskal-Wallis test and Chi-square test were used for continuous and categorical variables, respectively.

^bInclusive of African American, Hispanic, Asian, Other, and Unknown

^cCalculated for patients with available creatinine data using 2021 CKD-EPI creatinine calculation.

Abbreviations: eGFR, estimated glomerular filtration rate; N/A, not applicable; SD, standard deviation

Table S4. Demographic and clinical characteristics stratified by availability of proteinuria data

Category	With available proteinuria data		Without available proteinuria data	
	Frequency (N)	Percentage (%)	Frequency (N)	Percentage (%)
Overall	2 280	100.0%	2 095	100.0%
Age (years)				
Mean (SD)	46.9 (16.2)		48.5 (16.9)	
Median (Q1, Q3)	45 (34, 59)		48 (35, 62)	
Sex				
Female	891	39.1%	742	35.4%
Male	1 389	60.9%	1 353	64.6%
Race/Ethnicity				
Caucasian	1 169	51.3%	1 003	47.9%
African American	123	5.4%	102	4.9%
Hispanic	123	5.4%	107	5.1%
Asian	155	6.8%	92	4.4%
Other	57	2.5%	67	3.2%
Unknown	653	28.6%	724	34.6%
Hypertension	1 759	77.1%	1 471	70.2%
Yes	1 533	87.2%	1 319	89.7%
No	226	12.8%	152	10.3%
Hematuria	1 699	74.5%	1 259	60.1%
Present	1 568	92.3%	1 193	94.8%
Absent	131	7.7%	66	5.2%
Creatinine (mg/dL)				
Patients with available creatinine measurements	2 172	95.3%	1 656	79.1%
Mean (SD)	2.6 (2.5)		3.5 (3.5)	
Median (Q1, Q3)	1.8 (1.2, 2.9)		2.4 (1.5, 4.2)	
eGFR^a (ml/min/1.73m²) *	2 172	95.3%	1 656	79.1%
Mean (SD)	49.0 (33.5)		37.8 (30.6)	
Median (Q1, Q3)	40.0 (22.7, 68.7)		29.0 (14.6, 50.8)	
Stage 1	333	15.3%	158	9.5%
Stage 2	339	15.6%	142	8.6%
Stage 3	742	34.2%	508	30.7%
Stage 3A	290	13.4%	191	11.5%
Stage 3B	452	20.8%	317	19.1%
Stage 4	454	20.9%	411	24.8%
Stage 5	304	14.0%	437	26.4%
Proteinuria				
Using 24-hour results (g/d)				
Mean (SD)	4.2 (4.0)			
Median (Q1, Q3)	3.0 (1.5, 5.8)			
≥1 g/d	1 937	85.0%		
<1 g/d	343	15.0%		
≥3.5 g/d	1 018	44.7%		

<3.5 g/d	1 262	55.4%	
----------	-------	-------	--

^aCalculated for patients with available creatinine data using 2021 CKD-EPI creatinine calculation.

Abbreviations: eGFR, estimated glomerular filtration rate; SD, standard deviation; g/d, grams per day, Q, quartile; mg/DL, milligrams per deciliter.

Table S5. Patient Count by MEST-C Characteristic/CKD stage

Model	Overall	CKD stage					
		Stage 1	Stage 2	Stage 3A	Stage 3B	Stage 4	Stage 5
Mesangial Hypercellularity	3 828	491	481	481	769	865	741
No significant mesangial hypercellularity	2 003	296	258	251	396	387	415
Mild to moderate mesangial hypercellularity	1 825	195	223	230	373	478	326
Endocapillary Hypercellularity	3 828	491	481	481	769	865	741
No endocapillary proliferation	3 168	406	410	411	648	691	602
Minimal endocapillary hypercellularity	660	85	71	70	121	174	139
Segmental Sclerosis^a	3 828	491	481	481	769	865	741
Absence of	1 335	208	161	136	237	276	317
Presence of	2 493	283	320	345	532	589	424
Tubular Atrophy^b	3 802	489	477	480	763	860	733
≤25%	1 551	446	324	223	224	171	163
26–50%	1 151	37	127	205	357	290	135
>50%	1 100	6	26	52	182	399	435
Crescents	3 828	491	481	481	769	865	741
C0 (no crescents)	3 113	405	393	396	644	712	563
C1 (crescent present in at least one glomerulus)	579	79	79	74	103	118	126
C2 (present in >25% of the glomeruli)	136	7	9	11	22	35	52

^a. Presence of segmental sclerosis and/or adhesion of tuft to Bowman capsule

^b Degree of tubular atrophy or interstitial fibrosis. 'Unknown' was not included in models.

Abbreviations: CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; SD, standard deviation.

Table S6. Patient Count by MEST-C Characteristic/Proteinuria Level

Model	Overall	Proteinuria level	
		<1 g/d	≥1 g/d
Mesangial Hypercellularity	2 280	343	1 937
No significant mesangial hypercellularity	1 136	206	933
Mild to moderate mesangial hypercellularity	1 144	137	1 007
Endocapillary Hypercellularity	2 280	343	1 937
No endocapillary proliferation	1 890	309	1 581
Minimal endocapillary hypercellularity	390	34	356
Segmental Sclerosis^a	2 280	343	1 937
Absence of	687	153	534
Presence of	1 593	190	1 403
Tubular Atrophy^b	2 266	339	1 927
≤25%	956	233	723
26–50%	691	75	616
>50%	619	31	588
Crescents	2 280	343	1 937
C0 (no crescents)	1 842	304	1 538
C1 (crescent present in at least one glomerulus)	370	36	334
C2 (present in >25% of the glomeruli)	68	3	65

^a Presence of segmental sclerosis and/or adhesion of tuft to Bowman capsule

^b Degree of tubular atrophy or interstitial fibrosis. 'Unknown' was not included in models.

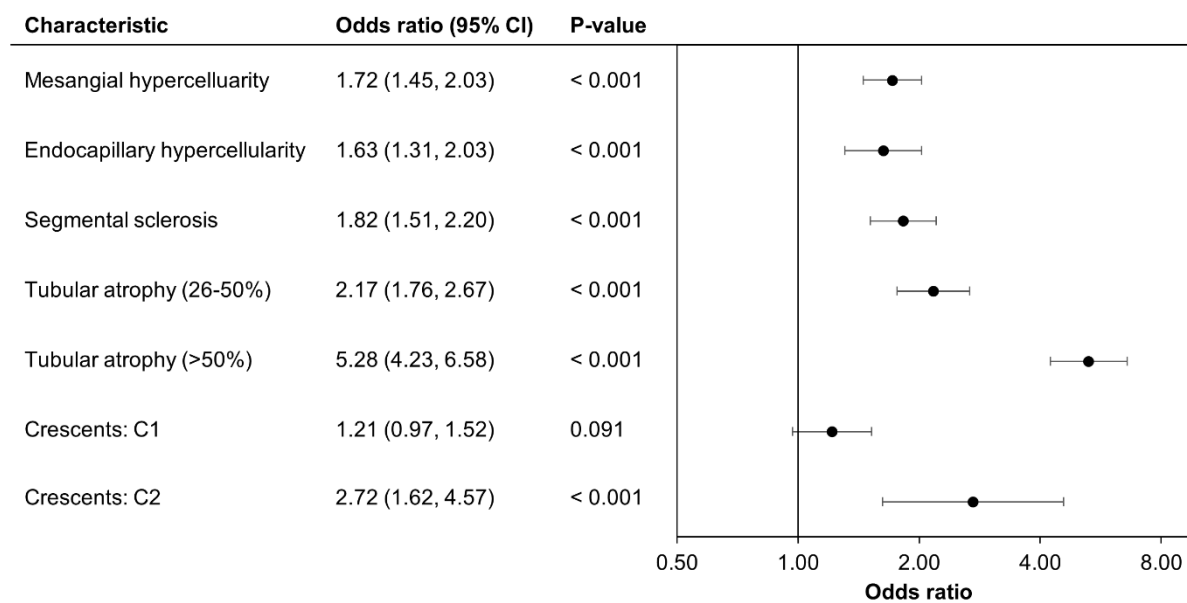


Figure S2. Associations between histological characteristics and proteinuria (≥ 3.5 vs < 3.5 g/d)

C1, crescent present in 1–24% of glomeruli; C2, present in >25% of glomeruli.

Reference groups:

Proteinuria (Ref: < 3.5 g/d)

Mesangial Hypercellularity (Ref: No significant mesangial hypercellularity)

Endocapillary Hypercellularity (Ref: No endocapillary proliferation)

Segmental Sclerosis (Ref: Absence)

Tubular Atrophy (Ref: $\leq 25\%$)

Crescents (Ref: C0 (no crescents))

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No.	Recommendation	Page No.	Relevant text from manuscript
Title and abstract	1	(a) Indicate the study’s design with a commonly used term in the title or the abstract	1	Adult IgA nephropathy in the United States: a retrospective cohort study assessing clinicopathological characteristics
		(b) Provide in the abstract an informative and balanced summary of what was done and what was found	2-3	<p>Introduction IgA nephropathy (IgAN) is a progressive autoimmune kidney disease and a leading cause of glomerular disease that can result in end stage kidney disease (KF). Median age at diagnosis is 35-37 years and close to 50% of patients will progress to KF within 20 years. We aimed to enhance the understanding of renal histology and chronic kidney disease (CKD) stage at the time of IgAN diagnosis using a large real-world biopsy cohort.</p> <p>Methods This retrospective cohort study evaluated biopsy data and clinical characteristics from adult patients within the US diagnosed with IgAN between 1 January 2016 to 31 May 2020. Descriptive statistics were summarized and relationship(s) between each Oxford Classification (MEST-C) component score with 24-H proteinuria or CKD stage were examined using regression analysis.</p> <p>Results A total of 4 375 patients (mean age 47.7 years, 62.7% male) met eligibility criteria. Mild to moderate mesangial hypercellularity (47.3%), segmental sclerosis (65.0%), tubular atrophy \geq 25% (57.4%) and crescents (18.5%) were identified and 74.6% of patients were at CKD stage \geq 3. Proteinuria \geq 1 g/d was associated with higher MEST-C scores, and odds of mesangial hypercellularity, segmental sclerosis, tubular atrophy and crescents increased with CKD stage.</p> <p>Conclusion Most patients with IgAN in our US cohort were diagnosed at CKD stage \geq3 and had high MEST-C scores and proteinuria suggestive of significant disease burden at the time of kidney biopsy. Strategies are required to raise awareness and promote</p>

earlier detection of asymptomatic urinary abnormalities before extensive irreversible kidney damage has occurred.

Introduction

Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	4-5	<p>IgA nephropathy (IgAN), a progressive autoimmune kidney disease, is the most common primary glomerular disease worldwide and a leading cause of end stage kidney disease (KF). Estimated incidence is between 1.29/100 000 (all ages) and 2.5/100 000 (adults) per year in the US population. Globally, prevalence rates tend to be lower in Africa, moderate in Northern Europe, and higher in East and Pacific Asian regions; however, the extent to which these variations in reported rates may be a reflection of differences in biopsy practices is unclear. IgAN is characterized by IgA deposition in the glomerular mesangium, which in turn triggers release of inflammatory cytokines and complement activation. Progressive glomerular injury and tubulointerstitial fibrosis are potential long-term consequences. Glomerular changes are diverse, ranging from no changes by light microscopy to crescentic glomerulonephritis. The extent of interstitial fibrosis and tubular atrophy in the biopsy sample has been demonstrated to be the strongest histological predictor of disease progression and renal survival. IgAN frequently progresses slowly without symptoms and is usually diagnosed when concern is raised due to gross or more commonly microscopic, hematuria and/or proteinuria with or without abnormal renal function. By the time of IgAN diagnosis, 50% of adult patients may already have chronic kidney disease (CKD) stage 3 or greater. This delay in diagnosis may expose the kidneys to sustained periods of elevated proteinuria which has been shown to be the strongest clinical predictor of progression to KF. Diagnosis, which requires a renal biopsy, often occurs between 25 and 40 years of age, and up to 50% of patients will progress to KF within approximately 20 years from diagnosis. Patients with IgAN may therefore require kidney replacement therapy (KRT) at a relatively young age. As the clinical course of IgAN is so variable, potential parameters that may be predictive of outcome have been studied. Histological findings of renal biopsy samples graded using the Oxford Classification (MEST-C) have been shown to predict</p>
----------------------	---	--	-----	--

				decline in renal function. Predictive power similar to that achieved through clinical monitoring over a 2-year period can be achieved through combination of a patient's MEST-C score with clinical data such as glomerular filtration rate, presence of arterial hypertension and the degree of proteinuria at the time of biopsy. More recently, the international IgAN risk-prediction tool has been developed by Barbour et al. to predict the risk of a 50% decline in estimated glomerular filtration rate (eGFR) or KF within 5-years. Few studies have reported the histological characteristics identified via kidney biopsy in adult patients with IgAN and the impact on disease progression.
Objectives	3	State specific objectives, including any prespecified hypotheses	5	Given the frequently asymptomatic presentation of IgAN and the potential negative implications of uncontrolled disease on long-term prognosis, we sought to characterize disease stage and histological patterns in a modern large real-world cohort of US adult patients with IgAN at the time of kidney biopsy. We further examined potential associations between MEST-C scores and 24-H proteinuria and CKD stage at the time of diagnosis.
Methods				
Study design	4	Present key elements of study design early in the paper	6	This non-interventional, retrospective cohort study comprises patients diagnosed with IgAN in the US between 1 January 2016 to 31 May 2020 using biopsy samples processed by Arkana Laboratories (Little Rock, AR). Arkana Laboratories (Arkana) provides renal pathology, serology, molecular pathology, and neuropathology services from US healthcare institutions across 43 states. The associated dataset contains de-identified demographic and clinical information which are collected at the time of biopsy assessment in addition to histological findings generated from adult patients with biopsy confirmed IgAN.
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6	
Participants	6	(a) <i>Cohort study</i> —Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up <i>Case-control study</i> —Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls <i>Cross-sectional study</i> —Give the eligibility criteria, and the sources and methods of selection of participants	6	To be included in this analysis, patients must have been at least 18 years of age at the time of kidney biopsy, have received a diagnosis of IgAN based upon biopsy results and have no history of prior kidney transplant.

		(b) <i>Cohort study</i> —For matched studies, give matching criteria and number of exposed and unexposed <i>Case-control study</i> —For matched studies, give matching criteria and the number of controls per case	N/A	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	6-7	Clinical characteristics in the de-identified dataset included presence/absence of hypertension, presence/absence of hematuria, serum creatinine, and level of proteinuria-based on urinary protein-creatinine ratio (PCR) or 24h urinary protein excretion rate (24h-PER). To support proteinuria-based analyses, all proteinuria data are presented as 24h-PER (g/d) with PCR values in g/g converted to g/d. eGFR without race/ethnicity modifier was calculated by the research team according to the CKD-EPI Creatinine Equation 2021 using the serum creatinine value closest to the biopsy date. CKD stage (1–5) was assigned using eGFR values: Stage 1 (eGFR ≥ 90); Stage 2 (eGFR 60–89); Stage 3A (eGFR 45–59); Stage 3B (eGFR 30–44); Stage 4 (eGFR 15–29); Stage 5 (eGFR < 15). MEST-C scores were assigned by Arkana for the following histological characteristics: mesangial hypercellularity (M0, no significant; M1, mild to moderate); endocapillary hypercellularity (E0, no proliferation; E1, minimal proliferation); segmental sclerosis (S0, absence of sclerosis or adhesions; S1, presence of sclerosis or adhesions); tubular atrophy (T0, ≤25%; T1, 26–50%; T2, > 50%); crescents (C0, no crescents; C1, present in at least one glomeruli; C2, present in >25% of glomeruli). IgA, IgG and IgM, Cq1 and C3, κ and λ, and fibrinogen levels were assessed using immunofluorescence microscopy.
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	7	Arkana Laboratories reports use of standard renal biopsy processing techniques to include light, immunofluorescence, and electron microscopy. Light microscopy samples were fixed in formalin, embedded in paraffin, serially cut at 3 μm, and stained with hematoxylin and eosin, Jones methenamine silver, Masson trichrome, or periodic acid-Schiff reagent. Tissue for immunofluorescence was embedded in OCT, snap frozen, and sections were cut at 4 μm in a cryostat and stained with fluorescein-tagged polyclonal rabbit anti-human antibodies to IgG, IgA, IgM, C3, C1q, fibrinogen, and κ and λ -light chains

				(Dako, Carpinteria, CA) for 1 h and rinsed; a coverslip was applied using aqueous mounting media. For electron microscopy, 1 mm cubes were removed from the ends of the biopsy sample, dehydrated with graded alcohols, and embedded in Epon/Araldite resin. Sections were cut at 1 µm with an ultramicrotome and stained with toluidine blue and examined with a light microscope for glomerular evaluation. Thin sections were cut at 60 nm and examined in a Jeol JEM 1011 electron microscope (Jeol, Tokyo, Japan) and photomicrographs taken at 4 000, 12 000 and 20 000 x magnification.
Bias	9	Describe any efforts to address potential sources of bias	14	See “Limitations”
Study size	10	Explain how the study size was arrived at	6	See “Study Design” and “Participants”
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	8	Descriptive analyses were conducted; categorical variables were summarized using frequency and percentage, and continuous variables were summarized using mean, standard deviation, median, and interquartile range. Multinomial logistic regression was conducted to examine associations (odds ratio [OR] and 95% confidence interval [CI]) between each MEST-C component (mesangial hypercellularity, endocapillary hypercellularity, segmental sclerosis, tubular atrophy, or crescents) and CKD stage, adjusting for age, sex, and race in each model. Odds of each MEST-C component being present at CKD stages (2, 3A, 3B, 4, and 5) were assessed as compared to CKD stage 1. Logistic regression was used to assess the association between each MEST-C component and proteinuria (24-h PER: ≥ 1 g/d vs < 1 g/d), controlling for age, sex, and race in each model.
Statistical methods	12	(a) Describe all statistical methods, including those used to control for confounding	8	Multicollinearity among predictors and covariates was evaluated using a correlation matrix, and variance inflation factor. All statistical tests were two-sided, with a significance level of P < 0.05. Analyses were performed using SAS Studio 3.81 (SAS Institute, Cary, NC).
		(b) Describe any methods used to examine subgroups and interactions	8	
		(c) Explain how missing data were addressed	8	Missing data was considered as a separate category in the primary objective analysis and described using frequency counts and percentages for both categorical and continuous covariates. Only patients without missing values or ‘unknown’ variables were included in regression models and statistical testing of interest.

		(d) <i>Cohort study</i> —If applicable, explain how loss to follow-up was addressed <i>Case-control study</i> —If applicable, explain how matching of cases and controls was addressed <i>Cross-sectional study</i> —If applicable, describe analytical methods taking account of sampling strategy		N/A
		(e) Describe any sensitivity analyses	11	A similar pattern was evidenced amongst patients with proteinuria values ≥ 3.5 g/d compared with < 3.5 g/d (Supplementary Figure S1)
Results				
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	9	A total of 4 375 patients met the full eligibility criteria for this study (Table 1).
		(b) Give reasons for non-participation at each stage		N/A
		(c) Consider use of a flow diagram		N/A
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	9	The mean age was 47.7 years, and most patients were male (62.7%). Amongst persons with known race (n= 2 998 [68.5% of sample]), 72.4% were Caucasian, 7.5% African American, 7.7% Hispanic and 8.2% Asian. Clinical indication of hypertension (presence/absence) was available for 3 230/4 375 patients (73.8% of sample) with 88.3% (2 852/3 230) exhibiting elevated blood pressure. Hematuria was found to be present in 93.3% (2 761/2 958) of patients with available data. Mean (SD) serum creatinine and eGFR were 3.0 (3.0) mg/dl and 44.2 (32.8) ml/min/1.73m ² respectively, amongst the 3 828 patients with available creatinine values (histograms shown in Figure S1). Amongst this subset, 74.6% presented at CKD stage 3 or higher. Just over half (52.1%) of the cohort had a proteinuria value, measured at mean (SD) of 4.2 (4.0) g/d. Proteinuria values ≥ 1 g/d were observed in 85.0% (1 937/2 280) of patients with available data.
		(b) Indicate number of participants with missing data for each variable of interest		Shown in data tables (number of patients with each variable is shown alongside the percentage of overall patients)
		(c) <i>Cohort study</i> —Summarise follow-up time (eg, average and total amount)		N/A
Outcome data	15*	<i>Cohort study</i> —Report numbers of outcome events or summary measures over time		N/A

		<i>Case-control study</i> —Report numbers in each exposure category, or summary measures of exposure		N/A
		<i>Cross-sectional study</i> —Report numbers of outcome events or summary measures	10	Amongst the 4 375 patient biopsy samples evaluated, almost half (47.3%) showed mild to moderate mesangial hypercellularity (M) and 16.9% exhibited minimal endocapillary (E) hypercellularity (Figure 1). Segmental sclerosis (S) and/or adhesion of tuft to Bowman capsule was present in almost two thirds of biopsies (65.0%). Tubular atrophy (T) of 26–50% of kidney tissue was observed in 29.5% of patients, while 27.9% exhibited tubular atrophy in > 50% of kidney tissue. Crescents (C) were present in at least one glomerulus in 15.1% of patients and 3.4% had crescents in >25% of glomeruli.
Main results	16	(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	10-11	<p>Mesangial hypercellularity was associated with increased odds of being in a more severe CKD stage than Stage 1, with the adjusted odds ratios (95% CI) being 1.43 (1.10, 1.84), 1.61 (1.24, 2.09), 1.76 (1.38, 2.23), 2.37 (1.87, 3.01), and 1.44 (1.14, 1.84) for Stages 2, 3A, 3B, 4, and 5, respectively (Figure 2). Mesangial hypercellularity had the highest odds ratio for Stage 4 versus Stage 1, which indicates that most of patients with mesangial hypercellularity were in Stage 4 when their biopsy was taken. Similarly, CKD stages 3A, 3B, and 4 were each associated with greater odds of segmental sclerosis compared with CKD stages 1 and 2. Odds of tubular atrophy of > 50% of kidney tissue increased significantly with higher CKD stages to a max of 338.51 (146.76, 780.80; P < 0.001) amongst persons diagnosed at CKD stage 5. The odds of observing crescents in > 25% of glomeruli increased with higher CKD stage, reaching statistical significance for stage 3B (2.41 [1.01, 5.78]; P = 0.048) and beyond [stage 4 (3.56 [1.54, 8.21]; P = 0.003) and stage 5 (6.47 [2.87, 14.55]; P < 0.001). While presence of endocapillary hypercellularity showed a similar directional pattern with greatest odds (1.06 [0.79, 1.44]; P = 0.686) identified in patients with CKD stage 4, this finding was not statistically significant.</p> <p>Proteinuria \geq 1 g/d had a statistically significant association with the presence of all MEST-C histological characteristics assessed in this study (Figure 3). Patients with crescents present in at least 25% of their glomeruli had greater odds of having proteinuria \geq 1 g/d at biopsy (4.21 [1.31, 13.50]; P = 0.016), which increased</p>

		further in patients with > 50% tubular atrophy (6.03 [4.08, 8.92]; P < 0.001) . Patients were also at least twice as likely to have proteinuria \geq 1 g/d if they demonstrated endocapillary hypercellularity (2.04 [1.40, 2.96]; P < 0.001), segmental sclerosis (2.27 [1.78, 2.88]; P < 0.001) and tubular atrophy of 26-50% (2.58 [1.94, 3.43]; P < 0.001), and 1.68 times as likely if they presented with mesangial hypercellularity (1.68 [1.33, 2.13]; P < 0.001)
(b) Report category boundaries when continuous variables were categorized	8 (and 6-7 for definition of CKD stages)	Odds of each MEST-C component being present at CKD stages (2, 3A, 3B, 4, and 5) were assessed as compared to CKD stage 1. Patients were categorized into CKD stage using the calculated eGFR values: Stage 1 or Stage 2 (eGFR \geq 60 ml/min/1.73 m ²); Stage 3A (eGFR 45–59 ml/min/1.73 m ²); Stage 3B (eGFR 30–44 ml/min/1.73 m ²); Stage 4 (eGFR (15–29 ml/min/1.73 m ²); Stage 5 (eGFR <15 ml/min/1.73 m ²). Logistic regression was used to assess the association between each MEST-C component and proteinuria (24-h PER: \geq 1 g/d vs < 1 g/d)
(c) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period		N/A

Continued on next page

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	11	A similar pattern was evidenced amongst patients with proteinuria values ≥ 3.5 g/d compared with < 3.5 g/d (Figure S2).
Discussion				
Key results	18	Summarise key results with reference to study objectives	11-14	<p>Among the 3 828 US adults with known serum creatinine values, three quarters (75%) received their diagnosis at CKD stage 3 or later, while 42% had advanced to CKD stage 4 or 5 by the time of diagnosis.</p> <p>Our results demonstrated that a high proportion of patients in our US cohort had substantial kidney damage, with M1, E1, S1, T1/2, and C1/2 scores being assigned to high proportions of biopsies. We show that 47% of patients presented with mesangial hypercellularity, 17% with endocapillary hypercellularity, 65% with segmental sclerosis, 57% with $>25\%$ tubular atrophy, and 19% with crescents.</p> <p>Excluding endocapillary hypercellularity, the remaining MEST-C scores showed a significant association with CKD stage; higher CKD stage, particularly stage 4, resulted in significantly higher odds for the presence of positive MEST-C scores. CKD stage 5 was observed to have a lower OR than stage 4 for mesangial hypercellularity, segmental sclerosis, and mild to moderate tubular atrophy (T1). The lower odds of segmental sclerosis observed at CKD stage 5 may be the result of the greater levels of global glomerulosclerosis expected in patients with advanced disease. A greater proportion of patients in CKD stage 5 were found to exhibit tubular atrophy in $> 50\%$ of tubules, correlating with the expected increase in odds, and a lower proportion exhibited tubular atrophy in 26–50% of tubules (Table S5).</p> <p>Our study provides evidence from a large cohort of patients that the association of MEST-C scores with clinical characteristics at diagnosis is evident for IgAN patients first diagnosed at CKD stages 4 and 5 as well as CKD stages 1–3.</p> <p>Most patients with available data in our cohort had proteinuria ≥ 1 g/d, with a mean value of 4.2 g/d. Analysis of the association between proteinuria and MEST-C in our cohort revealed that the presence of each MEST-C component came with higher odds of</p>

			having proteinuria ≥ 1 g/d, particularly segmental sclerosis and tubular atrophy.
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	14-15 When using secondary data for comparative analyses, several limitations exist. Although this study represents a modern large sample of US adult patients with IgAN, not all US states are included in the database and those that are included may not be equally represented; thereby possibly limiting our ability to generalize results to the overall US IgAN patient population. Patient demographic and clinical information was provided to Arkana on a secondary level by the clinicians who ordered the biopsy. In some cases, the forms accompanying the biopsy tissue were not completed with all requested information such as serum creatinine or urinary protein values. For example, in this study only 52% of patients had available proteinuria data, while 88% had available serum creatinine values. Whether these data points are absent due to a systematic patient-related cause or simply missing at random is beyond the ability of this study to determine. Finally, MEST-C scores were not validated for patients with eGFR <30 ml/min/1.73m ² in the original Oxford classification analysis. Our results in patients with CKD stages 4 and 5 should therefore be interpreted with more caution than results for CKD stages 1–3 when considering potential implications of MEST-C scores at diagnosis in relation to long-term renal outcomes.
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	15 Most US patients with IgAN were diagnosed at CKD stage ≥ 3 and had elevated MEST-C scores for mesangial hypercellularity, segmental sclerosis, and tubular atrophy which suggests significant disease duration at the time of kidney biopsy. Proteinuria ≥ 1 g/d and higher CKD stage at diagnosis are both associated with higher MEST-C scores in diagnostic biopsies and may therefore be early indicators of the need for diagnostic biopsy. Strategies to improve awareness and facilitate earlier detection of IgAN are needed to enable administration of current and future therapies before extensive and irreversible kidney damage has occurred, thereby maximising the potential to slow progression to KF.
Generalisability	21	Discuss the generalisability (external validity) of the study results	14 When using secondary data for comparative analyses, several limitations exist. Although this study represents a modern large sample of US adult patients with IgAN, not all US states are included in the database and those that are included may not be equally

represented; thereby possibly limiting our ability to generalize results to the overall US IgAN patient population.

Other information

Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	16	This work was funded by Traverre Therapeutics, Inc.
---------	----	---	----	---

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at <http://www.plosmedicine.org/>, Annals of Internal Medicine at <http://www.annals.org/>, and Epidemiology at <http://www.epidem.com/>). Information on the STROBE Initiative is available at www.strobe-statement.org.