

1 Validation of the International IgA risk prediction tool in American Indians and Hispanics.

2 Authors

3 1. Saeed Kamran Shaffi, MD, MS*

4 Staff Nephrologist, Raymond G. Murphy VA Medical Center

5 Associate Professor, Department of Internal Medicine, Division of Nephrology,

6 University of New Mexico

7

8 2. Edger Fischer, M.D., Ph.D.[∞]

9 Professor of Pathology

10 Division of Surgical Pathology and Cytopathology

11 University of New Mexico

12

13 3. Christos Argyropoulos, MD, Ph.D., FASN[^]

14 Associate Professor, Department of Internal Medicine, Division of Nephrology,

15 University of New Mexico

16

17 4. Brent Wagner, MD*

18 Director, Kidney Institute of New Mexico

19 University of New Mexico Health Sciences Center

20 ACOS, Research Service

21 New Mexico VA Health Care System*

22

23 *Raymond G. Murphy New Mexico VA Health Care System

24 1501 San Pedro Dr SE, Albuquerque, NM, 87108

25

26 [∞]Department of Pathology

27 1 University of New Mexico

28 MSC08 4640

29 Albuquerque, NM, 87106

30

31 [^]Division of Nephrology,

32 1 University of New Mexico

33 MSC 04-2785

34 Albuquerque, NM, 87131

35

36 Corresponding author: Saeed Kamran Shaffi,

37 Email: kshaffi@salud.unm.edu

38 Abstract

39 Background: The International Immunoglobulin A nephropathy (IgAN) risk prediction assesses
40 the risk of kidney failure in patients with IgAN. The performance of this risk prediction tool has
41 not been studied in American Indians and Hispanics. We conducted a single-center study to
42 assess the equation performance in this population.

43

44 Methods: We calculated the 5-year risk of developing kidney failure with the IgAN risk
45 prediction equation without race and assessed the equation performance using the metrics of
46 calibration, discrimination, and overall prediction error.

47

48 Results: Thirty-four patients were included, most of whom identified as of Hispanic
49 race/ethnicity (44.1%), or as American Indians (26.5%). At biopsy, the median (IQR) age, serum
50 creatinine, and spot urine protein to creatinine ratio were 38 years (27-45), 2.15 mg/dl (1.51-
51 3.04), and 2.7 g/g (1.5-5.8), respectively. The equation identified patients at high risk of
52 developing kidney failure early with a concordance statistic of 0.79 (95% CI 0.68 – 0.89). The
53 agreement between observed and predicted outcomes at 5 years was marginal, with over-
54 estimation of risk for patients with low observed risk and vice versa. Overall prediction error
55 was suboptimal in this cohort [index of prediction accuracy 0.34 (0.03 – 0.51)].

56

57 Conclusions: The International IgAN risk prediction equation without race accurately identified
58 patients at elevated risk of developing kidney failure. At 5 years, the agreement between the

59 observed and predicted outcomes was sub-optimal, possibly due to advanced kidney disease in
60 this cohort. A diverse development population may improve the risk prediction.
61

- 62 Key words: Immunoglobulin A Risk Prediction – Risk Prediction Model Calibration Assessment – Risk
63 Prediction Model Discrimination Assessment

64 Introduction:

65 IgA nephropathy (IgAN) is one of the most common primary glomerular diseases worldwide, with an
66 incidence of 2.5 per 100,000 in adults. It is associated with end-stage kidney disease (ESKD) in 10-40% of
67 patients by ten years of diagnosis [1]. Our understanding of IgAN pathophysiology has improved due to
68 recent advances, which have led to new therapies [2, 3]. Therefore, it is imperative to identify patients
69 who would rapidly progress to kidney failure and would benefit from early goal-directed treatments.
70 Furthermore, an accurate risk prediction tool - that identifies patients at an elevated risk of developing
71 kidney failure - would assist patients and clinicians to prepare for transplant and renal replacement
72 therapy and instituting measures to decrease cardiovascular risk, which is associated with mortality in
73 patients with chronic kidney disease [4].

74 To address these issues, Barbour et al. developed an International IgA risk prediction tool that uses
75 demographic, clinical, and kidney biopsy variables to assess the risk of kidney failure or >50% decline in
76 the estimated glomerular filtration rate (eGFR) [5]. The authors developed two equations with and
77 without an ethnic variable (Chinese, Japanese, or Other). The equation without a race/ethnicity variable
78 was developed (n = 2781) and validated (n = 1146) in large cohorts comprising Europeans, Chinese, and
79 Japanese. Since the publication of the original report, IgAN risk prediction tool performance has been
80 studied in Chinese [6, 7], South Asian [8], Greek [9], and South American [7] cohorts. Table 2 summarizes
81 the performance of the equation without the race/ethnicity variable in these cohorts. The equation
82 performance has not been studied in the Hispanic and American Indian populations. A New Mexico
83 based kidney biopsy registry data analysis has revealed a higher frequency of IgAN in the American
84 Indians compared to their state population share which may point towards a higher incidence of IgAN in
85 them [10]. Therefore, we conducted this single-center study to assess the performance of the
86 International IgA risk prediction tool without race in a cohort from New Mexico where, Hispanics and
87 American Indians are in the majority.

88

89 Methods:

90 The University of New Mexico Kidney Biopsy registry was established by searching for the current
91 procedural terminology (CPT) codes for kidney biopsies performed at the University of New Mexico
92 (UNM) hospital between 2001 and 2016. The institutional review board at UNM approved the Kidney
93 Biopsy Registry. The principal investigator (SKS) created a HIPPA compliant online Kidney Biopsy registry.
94 We queried the complete biopsy report for the following terms: “Immunoglobulin A nephropathy,” “IgA
95 Nephropathy,” and “IgAN.” We reviewed the biopsies to identify patients with primary IgAN who met
96 the inclusion criteria. The IgA risk prediction tool without the race variable uses demographic and clinical
97 variables at the time of biopsy that includes chronic kidney disease epidemiology (CKD-EPI) equation
98 estimated GFR, mean arterial blood pressure, proteinuria, age, the Oxford kidney biopsy classification
99 scores of mesangial hypercellularity, endothelial proliferation, segmental sclerosis, and tubular atrophy,
100 and medication use of angiotensin-converting enzyme inhibitor or angiotensin receptor blocker, and
101 immunosuppression around the biopsy to predict the risk of kidney failure defined as: 1) end-stage
102 kidney disease requiring dialysis or a kidney transplant or an estimated glomerular filtration rate of < 15
103 ml/min/1.73m² – or, 2) >50% decline in eGFR from the baseline value. For this study, we included
104 patients with primary IgAN and on whom the International IgAN Prediction Tool without race variable
105 predictors were available. Other inclusion criteria included the availability of longitudinal data for at
106 least 12 months. The patients who developed kidney failure or had >50% decline in the eGFR within 12
107 months were also included. We excluded patients who were < 18 years of age, who had kidney failure at
108 the time of biopsy, and on whom the risk prediction variables or longitudinal laboratory data to assess
109 risk were unavailable. The primary outcome was the development of end-stage kidney disease or an
110 eGFR <15 ml/min/1.73m² or a > 50% decline in the estimated GFR from the value at the time of biopsy.

111 We summarized the baseline variables stratified on race and ethnicity and reported the categorical
112 variables as n (%) and continuous variables as median (IQR). For comparisons between groups, we
113 performed the Kruskal-Wallis Rank Sum test or Fisher's Exact test as appropriate. A renal pathologist
114 (EF) reviewed all the biopsies and assigned them a MEST-C score. For survival analysis, patients were
115 censored when they were lost to follow-up or at the end of the study.

116 We calculated the linear predictors and the predicted probability of the primary outcome using the
117 formula without the race variable (S1) as we had few patients with the Chinese or Japanese race in our
118 cohort.

119 We assessed the performance of the equation in our cohort using the metrics of calibration,
120 discrimination, and the overall prediction error [11, 12]. A well-calibrated risk assessment model's
121 predictions should match the observed outcomes. Ideally, a model's predictions should not only match
122 the observed effects at a particular time (e.g., 5 years) - known as calibration in the large - as well as
123 across the entire range of predictions. The population is usually divided into risk categories for
124 calibration assessment. The mean observed risk for each category is obtained from a Kaplan-Meier
125 model and risk predictions from the risk equation. For each category, the observed and predicted risks
126 are plotted to assess agreement. There are a few issues with this approach; first, the risk categories are
127 arbitrary and result in loss of information, and second, this method cannot account for censored
128 patients. To circumvent these issues, we used the pseudo-observations approach for the primary
129 outcome indicator as recommended by the Strengthening Analytical Thinking of Observations Studies
130 (STRATOS) initiative[11, 12]. Observed risk calculations employed weighting techniques with pseudo-
131 observations accounting for the censored patients. We regressed the primary outcome with the linear
132 predictor by performing a Cox-proportional Hazard analysis. We obtained the 5-year estimated and
133 observed risk and plotted them to obtain a calibration plot. The calibration plot should follow a 45-
134 degree line for a well-calibrated equation; risk estimate values above and below the 45-degree line

135 suggest under and over-prediction, respectively. We also obtained numerical summaries of calibration.
136 The ratio of the observed and expected outcomes (O/E) should be 1; a ratio less than 1 signifies
137 overprediction and greater than 1 under-prediction. We obtained the O/E ratio at 5 years. We regressed
138 the primary outcome with the log-transformed predicted risk using a restricted cubic splines model with
139 smoothing to obtain the calibration intercept and slope. An intercept of 0 means that, on average, the
140 observed and the predicted risks are well matched. A slope greater than or less than 1 indicates the
141 homogeneity and heterogeneity of the risk estimates, respectively.

142 Discrimination assesses the ability of a risk model to identify patients who had the outcome of interest
143 earlier than others. We regressed the primary outcome with the linear predictor to assess discrimination
144 and calculated the concordance statistics (C-Statistic). The area under the curve (AUC) was graphically
145 represented by plotting the true positive rate (sensitivity) against the false positive rate (1 – specificity).

146 The index of prediction accuracy (IPA) summarizes both calibration and discrimination and represents
147 the overall prediction error [13]. We used the following formula to calculate IPA:

$$148 \quad PA = 1 - \text{Brier Score (Prediction Model)} / \text{Brier Score (Null Model)}$$

149 Where the Brier score is the mean squared difference between the primary outcome indicator and risk
150 estimates, the prediction model is a Cox-proportional Hazards model regressed with the linear predictor,
151 and the null model had no predictors. A higher IPA score indicates better model performance.

152 We used bootstrapping approach (100 bootstraps with data replacement) for estimating the area under
153 the curve, C-statistics, and IPA to ensure that the standard errors and confidence intervals for these
154 metrics were accurate.

155

156 Results:

157 We identified 74 patients with suspected IgA nephropathy by searching the University of New Mexico
158 Kidney biopsy registry (Fig 1). Out of 74, 34 patients met the eligibility criteria. The median age at the
159 time of the kidney biopsy was 38 years (IQR 27- 45 years) [Table 1]. Most of the patients identified as of
160 the White race (38%). Of the patients who identified as White, the majority were of Hispanic ethnicity
161 (53%). We collated race and ethnicity into a single variable and observed that about 3 out of 7 patients
162 were of Hispanic race/ethnicity (44.1%), 2 out of 7 identified as of American Indian race/non-Hispanic
163 ethnicity (26.5%), and the rest were of other races and ethnicities (29.4%). The other category included
164 Asians (n = 3), non-Hispanic Whites (n=3), African Americans (n =1), and of different races or ethnicities
165 or unknown races and ethnicities (n=3).

166 At the time of biopsy, the median (IQR) serum creatinine and Chronic Kidney Disease Epidemiology
167 equation (CKD-EPI) eGFR were 2.15 mg/dl (1.51 – 3.04) and 34 ml/min/1.73m² (20-51), respectively
168 (Table 1). Median (IQR) proteinuria assessed by a spot urine protein to creatinine ratio around the time
169 of biopsy was 2.7 g/g (1.5-5.8). Mean arterial blood pressure was 99 mm Hg (IQR 89 -104 mm Hg).
170 Thirty-five percent of the patients were on immunosuppression, and 62% were either on an angiotensin-
171 converting enzyme inhibitor or an angiotensin receptor blocker within 12 months of IgAN diagnosis.
172 American Indians showed a non-significant trend towards a higher serum creatinine, mean arterial
173 pressure, and a lower CKD-EPI estimated GFR than the other races. American Indians and Hispanics
174 showed a trend towards higher proteinuria around the time of biopsy than the other races, but it was
175 not statistically significant. On the Oxford classification of IgAN, Hispanics were more likely (100%) to
176 have an S1 score than the other races (p 0.049). Of note, 89% of the American Indians had an S1 score
177 on kidney biopsy.

178

179 Figure 3 shows the plot of the 5-year estimated and observed risk. The observed risk was lower for the
180 estimated risks between 20 and 50%; therefore, the model overestimated risk. The observed risk was
181 higher for the estimated risks above 50%, indicating an underestimation of risk by the model; however,
182 due to a small sample size, the confidence intervals were wide indicating imprecision of our findings.

183 Table 3 summarizes the numerical summaries of model calibration. At five years, the observed and
184 estimated risk ratio (O/E) was less than 1 [0.88 (95% CI (Confidence Interval) 0.51-1.26)], indicating over-
185 prediction; however, the confidence intervals crossed 1. The intercept was close to 0 [0.10 (CI – 0 .37 -
186 0.58)]; since the confidence intervals crossed 0, we cannot make a conclusive statement about the
187 intercept.

188 The risk prediction equation identified patients at a greater risk of developing the primary outcome early
189 with a C-statistic of 0.79 (0.68 - 0.89). Figure 2 shows the receiver operator curve (ROC) with the area
190 under the curve, which was 0.78 (0.58 -0.98).

191 The prediction accuracy index was 0.34 (0.03 -0.51) (table 3), which was marginal due to poor model
192 calibration in our population.

193 In summary, in our cohort, the international IgA risk prediction tool had less than optimal calibration;
194 however, there was uncertainty in our results and excellent discrimination.

195

196 Discussion:

197 We assessed the performance of the International IgA risk prediction tool without race in a cohort from
198 2001 to 2016, which was primarily composed of the Hispanic and American Indian populations.
199 Compared to the equation development cohort, our population had a higher baseline creatinine
200 [median (IQR); 1.0 (0.8-1.4) vs. 2.14 (1.51 - 3.04)] and proteinuria [median (IQR); 1.2 (0.7 -2.2) vs. 2.7
201 (1.5-5.8)] around the time of IgA nephropathy diagnosis [5]. These patients were more likely to have
202 higher segmental sclerosis (S1 77% vs. 88%) and tubular atrophy scores (T1 24.7 vs. 71% and T2 4.6 vs.
203 26%, respectively). Considering these findings, it is reasonable to assume that our cohort had a longer
204 disease duration and severity. Indeed, the equation derivation population was younger when diagnosed
205 with IgAN than the population in our study [Age(years); median (IQR); 35.6 (28.2 -45.4) vs. 38 (27-45)]. In
206 our population, Hispanics were more likely to have segmental sclerosis on kidney biopsy (p 0.049) [Table
207 1]. For these reasons, the probability of achieving the primary outcome was higher in our cohort (68%)
208 than in the equation development cohort (17.7%). All these findings may be explained by healthcare
209 disparities faced by the Hispanic and American Indian residents of New Mexico. Race and ethnicity often
210 provide insight into the socioeconomic and healthcare disparities in a population [14]. A review of the
211 U.S. census 2017 data shows that in New Mexico, the poverty rates were higher for Hispanics (24.1 %)
212 and American Indians (31.9%) than for Whites (17.7%) or Asians (10.5%) [15]. Furthermore, for each
213 dollar earned by the White/Non-Hispanic residents of New Mexico, the American Indians and Hispanics
214 earned approximately one-third less, highlighting the earnings disparities [16].

215 This study confirms the results of a previous study [10] looking at the frequency of IgAN in New Mexico
216 in various races. US census data from 2010 showed that White/non-Hispanic, Hispanic, and American
217 Indian shares of New Mexico's population were 40.6, 46.4, and 8.5%, respectively [17]. The proportion
218 of the Hispanics in our dataset (44%) matched their state population share; however, American Indians'
219 representation in our cohort was more than expected (26%) which may point towards a higher

220 incidence of IgAN in the American Indian. This trend should be systematically evaluated in larger
221 datasets.

222 A good prediction equation should be able to distinguish patients who are at low risk of achieving an
223 outcome of interest from those who are at elevated risk. The international IgAN risk prediction tool did
224 an excellent job identifying patients at a high risk of developing kidney failure early. The C-statistics,
225 receiving operator curve, and the area under it (AUC) were close to what has been reported in the
226 derivation and other validation cohorts (Table 2).

227 A well-calibrated risk equation's predictions should match the observed risk. At 5 years, the equation
228 overpredicted the risk for the patients to whom it assigned a risk of 20-50% and underpredicted the risk
229 in patients to whom it assigned a higher risk (Figure 3); similar findings were also seen in the external
230 validation of this equation in Chinese, Whites, and South Asians (Table 2)[6–8]. However, the equation
231 calibration should be studied in a larger population to achieve more robust results.

232 Our study has weaknesses. First, our sample size is significantly smaller than the other populations in
233 which this equation has been validated; therefore, there is uncertainty in our results, and a more
234 extensive study is needed to assess the equation performance in American Indians and Hispanics. We
235 used the bootstrapping technique for the metrics of C-statistics and IPA to obtain robust estimates and
236 confidence intervals. Furthermore, despite of small sample size, the equation discrimination, and
237 calibration performance match what has been reported in the other cohorts (Table 2). Second, our
238 cohort had more severe disease than the equation development and the other cohorts in which it has
239 been validated apart from the South Asian cohort [8], which could be the reason for the poor equation
240 calibration seen in this study and underscores the need for an IgA risk-prediction equation developed in
241 ethnically and racially diverse patients with advanced kidney disease.

242 Herein we provide the first report of IgA risk prediction tool performance in a population from the U.S.
243 Southwest. We performed a rigorous chart review to ensure the accuracy of our data. We modeled the
244 observed and predicted risks on a continuous scale which precluded the creation of arbitrary risk
245 categories.

246 In summary, we showed that IgAN is diagnosed later in New Mexico, which may point towards
247 economic and health care disparities. The international IgAN risk prediction tool without race can
248 accurately and timely identify patients at a high risk of disease progression. In this cohort, the 5-year
249 observed risk was lower in patients assigned a risk estimate of <50% of developing kidney failure and
250 higher in patients assigned a risk estimate of >50%; these findings indicate over and under-prediction by
251 the IgA International risk prediction tool, respectively. Our study suggests that the international IgAN
252 risk prediction tool can be used to identify patients who are at a high risk of developing kidney failure;
253 however, the calibration of the international IgAN risk prediction equation in our population was less
254 than optimal and underscores the need of an IgAN risk prediction tool developed in a racially diverse
255 population with more severe disease.

256 Disclosures

257 Saeed Kamran Shaffi:

258 Advisory board: Novartis

259

260 Christos Argyropoulos:

261 Consultant Agreements: Bayer, Baxter, Otsuka, Quanta

262 Funding

263 Saeed K Shaffi: The University of New Mexico Kidney biopsy registry was established by funding from the

264 Dialysis Clinic, Inc.

265 Acknowledgments:

266 The authors want to acknowledge the services of University of New Mexico Clinical and Translational

267 Science Center (CTSC) who provided support for entering the data in the University of New Mexico

268 Kidney biopsy registry. We would specifically like to thank Mr. Hugo A Vilchis and Greg Trejo on working

269 on this project.

270 Author Contribution

271 Saeed K Shaffi was responsible for funding acquisition for establishing the kidney biopsy registry, study

272 conceptualization, project administration, data curation, and writing the original draft.

273 Edger Fischer reviewed the kidney biopsies and assigned a MEST-C score.

274 Christos Argyropoulos devised study methodology and contributed to analysis.

275 Brent Wegner advised about project administration, validation and reviewing and editing of the

276 manuscript.

277 Data Sharing Statement:

278 Restrictions to the data apply; as per the University of New Mexico Institutional Review Board, an IRB

279 modification with a data sharing agreement is required to provide data but can be arranged upon

280 request.

281 References

- 282 1. Moran SM, Cattran DC. Recent advances in risk prediction, therapeutics and pathogenesis of IgA
283 nephropathy. *Minerva Med.* 2019; *110*:439–449
- 284 2. Research C for DE and. FDA approves first drug to decrease urine protein in IgA nephropathy, a
285 rare kidney disease. FDA. 2021
- 286 3. Selvaskandan H, Kay Cheung C, Dormer J, et al. Inhibition of the Lectin Pathway of the
287 Complement System as a Novel Approach in the Management of IgA Vasculitis-Associated
288 Nephritis. *Nephron.* 2020; *144*:453–458
- 289 4. Sarnak MJ, Amann K, Bangalore S, et al. Chronic Kidney Disease and Coronary Artery Disease: JACC
290 State-of-the-Art Review. *J Am Coll Cardiol.* 2019; *74*:1823–1838
- 291 5. Barbour SJ, Coppo R, Zhang H, et al. Evaluating a New International Risk-Prediction Tool in IgA
292 Nephropathy. *JAMA Intern Med.* 2019; *179*:942–952
- 293 6. Zhang J, Huang B, Liu Z, et al. External Validation of the International IgA Nephropathy Prediction
294 Tool. *Clin J Am Soc Nephrol CJASN.* 2020; *15*:1112–1120
- 295 7. Zhang Y, Guo L, Wang Z, et al. External Validation of International Risk-Prediction Models of IgA
296 Nephropathy in an Asian-Caucasian Cohort. *Kidney Int Rep.* 2020; *5*:1753–1763
- 297 8. Bagchi S, Upadhyay AD, Barwad A, et al. The International IgA Nephropathy Network Prediction
298 Tool Underestimates Disease Progression in Indian Patients. *Kidney Int Rep.* 2022; *7*:1210–1218
- 299 9. Papasotiriou M, Stangou M, Chlorogiannis D, et al. Validation of the International IgA Nephropathy
300 Prediction Tool in the Greek Registry of IgA Nephropathy. *Front Med.* 2022;9
- 301 10. Smith SM, Tung KSK. Incidence of IgA-related nephritides in American Indians in New Mexico. *Hum*
302 *Pathol.* 1985; *16*:181–184
- 303 11. Heus P, Reitsma JB, Collins GS, et al. Transparent Reporting of Multivariable Prediction Models in
304 Journal and Conference Abstracts: TRIPOD for Abstracts. *Ann Intern Med.* 2020
- 305 12. van Geloven N, Giardiello D, Bonneville EF, et al. Validation of prediction models in the presence of
306 competing risks: a guide through modern methods. *BMJ.* 2022; e069249
- 307 13. Kattan MW, Gerds TA. The index of prediction accuracy: an intuitive measure useful for evaluating
308 risk prediction models. *Diagn Progn Res.* 2018; *2*:7
- 309 14. Flanagin A, Frey T, Christiansen SL, AMA Manual of Style Committee. Updated Guidance on the
310 Reporting of Race and Ethnicity in Medical and Science Journals. *JAMA.* 2021; *326*:621–627
- 311 15. New Mexico Report - 2017. Talk Poverty. [https://talkpoverty.org/state-year-report/new-mexico-](https://talkpoverty.org/state-year-report/new-mexico-2017-report/)
312 2017-report/. Accessed 7 Nov 2022

313 16. Earnings Disparities by Race and Ethnicity | U.S. Department of Labor.
314 <https://www.dol.gov/agencies/ofccp/about/data/earnings/race-and-ethnicity>. Accessed 7 Nov
315 2022

316 17. Promotions C for NM and, Public Information Staff UCB. US Census Bureau 2010 Census.
317 <http://www.census.gov/2010census/data/>. Accessed 22 Apr 2015

318

319

320 Tables with legends

321 **Table 1: Demographic, Clinical, and the Oxford Classification of the patients with IgAN stratified on Race**
 322 **and Ethnicity**

| | Overall, N = 34 [‡] | Hispanic, N = 15 [‡] | American Indian, N = 9 [‡] | Other, N = 10 [‡] | p-value [‡] |
|---|------------------------------|-------------------------------|-------------------------------------|----------------------------|----------------------|
| Age; Median (IQR) | 38 (27 - 45) | 35 (29 - 47) | 40 (26 - 44) | 40 (25 - 50) | >0.9 |
| Sex; n(%) | | | | | >0.9 |
| Female | 18 (53%) | 8 (53%) | 5 (56%) | 5 (50%) | |
| Male | 16 (47%) | 7 (47%) | 4 (44%) | 5 (50%) | |
| Race; n(%) | | | | | <0.001 |
| White | 11 (32%) | 8 (53%) | 0 (0%) | 3 (30%) | |
| American Indian | 9 (26%) | 0 (0%) | 9 (100%) | 0 (0%) | |
| Unknown | 6 (18%) | 4 (27%) | 0 (0%) | 2 (20%) | |
| Asian | 3 (8.8%) | 0 (0%) | 0 (0%) | 3 (30%) | |
| Hispanic | 3 (8.8%) | 3 (20%) | 0 (0%) | 0 (0%) | |
| Other | 1 (2.9%) | 0 (0%) | 0 (0%) | 1 (10%) | |
| Black/African American | 1 (2.9%) | 0 (0%) | 0 (0%) | 1 (10%) | |
| Creatinine (mg/dl); Median (IQR) | 2.15 (1.51 - 3.04) | 2.10 (1.56 - 3.03) | 2.41 (1.80 - 3.99) | 1.70 (1.40 - 2.75) | 0.4 |
| eGFR (ml/min/1.73m ²); Median (IQR) | 34 (20 - 51) | 34 (22 - 48) | 27 (11 - 41) | 40 (31 - 56) | 0.4 |
| Baseline Proteinuria (g/g); Median (IQR) | 2.7 (1.5 - 5.8) | 3.6 (1.6 - 6.1) | 3.6 (2.0 - 4.3) | 1.6 (1.1 - 2.7) | 0.3 |
| Mean Arterial BP (mmHg); Median (IQR) | 99 (89 - 104) | 97 (90 - 107) | 100 (97 - 103) | 95 (85 - 103) | 0.5 |
| Immunosuppression around biopsy; n (%) | 12 (35%) | 6 (40%) | 6 (67%) | 0 (0%) | 0.006 |
| ACEi/ARB around biopsy; n (%) | 21 (62%) | 9 (60%) | 6 (67%) | 6 (60%) | >0.9 |
| Primary Outcome; n (%) | 23 (68%) | 12 (80%) | 6 (67%) | 5 (50%) | 0.3 |
| M; n (%) | | | | | 0.8 |
| M0 | 10 (29%) | 5 (33%) | 3 (33%) | 2 (20%) | |
| M1 | 24 (71%) | 10 (67%) | 6 (67%) | 8 (80%) | |
| E; n (%) | | | | | >0.9 |
| E0 | 20 (59%) | 9 (60%) | 5 (56%) | 6 (60%) | |
| E1 | 14 (41%) | 6 (40%) | 4 (44%) | 4 (40%) | |
| S; n (%) | | | | | 0.049 |
| S0 | 4 (12%) | 0 (0%) | 1 (11%) | 3 (30%) | |
| S1 | 30 (88%) | 15 (100%) | 8 (89%) | 7 (70%) | |
| T; n (%) | | | | | 0.5 |
| T0 | 1 (2.9%) | 1 (6.7%) | 0 (0%) | 0 (0%) | |
| T1 | 24 (71%) | 9 (60%) | 6 (67%) | 9 (90%) | |
| T2 | 9 (26%) | 5 (33%) | 3 (33%) | 1 (10%) | |
| C; n (%) | | | | | 0.8 |
| C0 | 22 (65%) | 10 (67%) | 5 (56%) | 7 (70%) | |
| C1 | 12 (35%) | 5 (33%) | 4 (44%) | 3 (30%) | |

[‡] Median (IQR); n (%)[‡] Kruskal-Wallis rank sum test; Fisher's exact test

324Table 2: Summary of the studies assessing the performance of the international IgAN risk prediction tool in
325various populations

| | N | Population characteristic | Follow-up duration and use | Creatinine; median (IQR) | Proteinuria; median (IQR) | *R2D | [∞] AUC at 5 years ^Ω C-statistics | Calibration slope | Comments |
|---------------------------------|------|---|----------------------------|--------------------------|--|------|--|-------------------|---|
| Barbour et al; 2019: Derivation | 2781 | Predominantly White, Chinese, or Japanese (in that order) | 4.8 (3.0–7.6) | 1.0 (0.8–1.4) | 1.2 (0.7–2.2) | 25.3 | 0.81 | | |
| Barbour et al; 2019: Validation | 1146 | Predominantly Japanese, Chinese or White (in that order) | 5.8 (3.4–8.5) | 1.0 (0.8–1.3) | 1.3 (0.6–2.4) | 35.3 | 0.81 | 1.19 (1.04-1.34) | Excellent discrimination Good calibration |
| J Zhang et al; 2020 | 1373 | Chinese | 2.4 (1.8–3.6) | 1.0 (0.8–1.4) | 1.7 (0.9–3.3) | | 0.862 | 1.30 (1.15–1.45) | Excellent discrimination The model modestly underestimated the risk |
| Y Zhang et al; 2020 | 1275 | Primarily Chinese with a few Argentinian | 4.8 (3.0–7.6) | 1.04 (0.80–1.4) | 1.2 (0.7–2.2) | 37.6 | 0.81 | 0.87 (0.73–1.0) | Excellent discrimination Acceptable calibration for the prediction model without race |
| Bagchi et al; 2022 | 316 | South Asian | | Cr not reported | 2.6 (1.5–4.0) | 49.9 | 0.819 | 1.21 (0.95–1.47) | Excellent discrimination The model underestimated the trajectory of disease progression across all risk groups |
| Papasotiriou et al; 2022 | 264 | Greek | 8.5 (5-10.83) | 1.2 (0.9–1.8) | Not reported: 41% of the patients had a proteinuria of ≤1 gram per day | 35 | 0.70 | 0.40 | Acceptable discrimination |

*R2D R-squared as Deviance, [∞]AUC Area under the curve, ^ΩC-statistics Concordance Statistics

327 *Table 3: Numerical summary of the measures of the International IgAN Risk Prediction Tool without race*
328 *performance in our cohort*

| Discrimination | |
|--|---------------------|
| Metric | Value |
| C-Statistics at 5 years | 0.79 (0.64 – 0.91) |
| Area Under the Curve at 5 years | 0.79 (0.68 – 0.89) |
| Calibration | |
| Observed/Estimated Risk at 5 years (O/E) | 0.88 (0.51 – 1.25) |
| Calibration Intercept | 0.10 (-0.37 – 0.58) |
| Calibration Slope | 3.03 (0.43 – 5.64) |
| Overall Prediction Error | |
| Index of Prediction Accuracy (IPA) | 0.33 (0.02 -0.50) |

329

330 Figure legends

331 Figure 1: Flow diagram showing the patients identified, excluded, and included in the study

332 Figure 2: Receiver operator curve (ROC) at 5-years of the International IgAN risk prediction tool in the
333 study population.

334 Figure 3: Calibration curve (Estimated risks plotted against Observed risk) at 5 years of the International
335 IgAN risk prediction equation in the study population. Values below and above the 45 degrees line
336 signify over and underprediction, respectively. The rug plot shows the distribution of the estimated
337 risks.

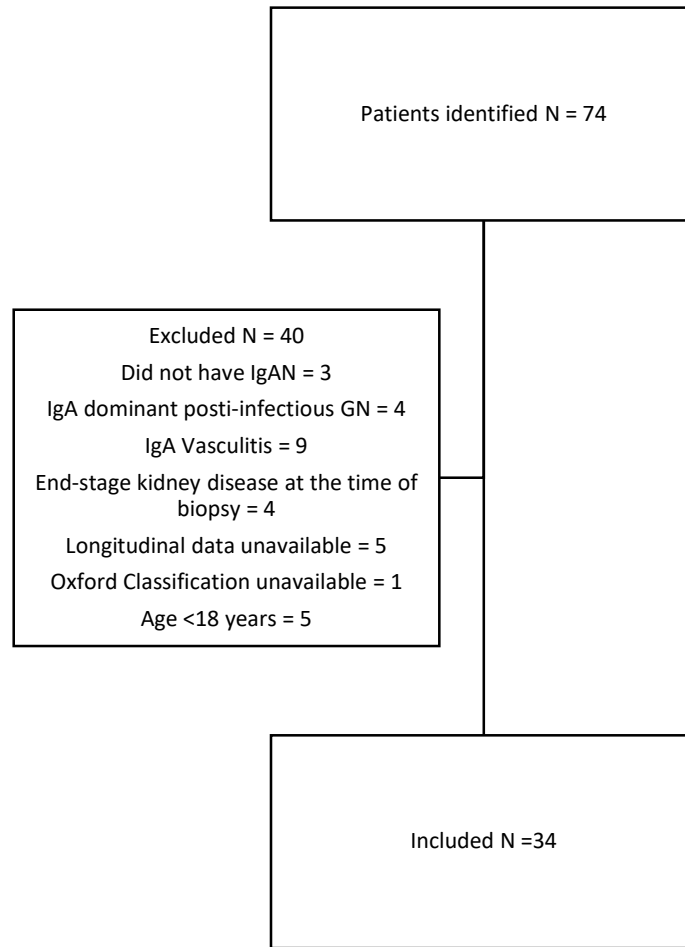


Figure 1: Flow diagram showing the patients identified, excluded, and included in the study

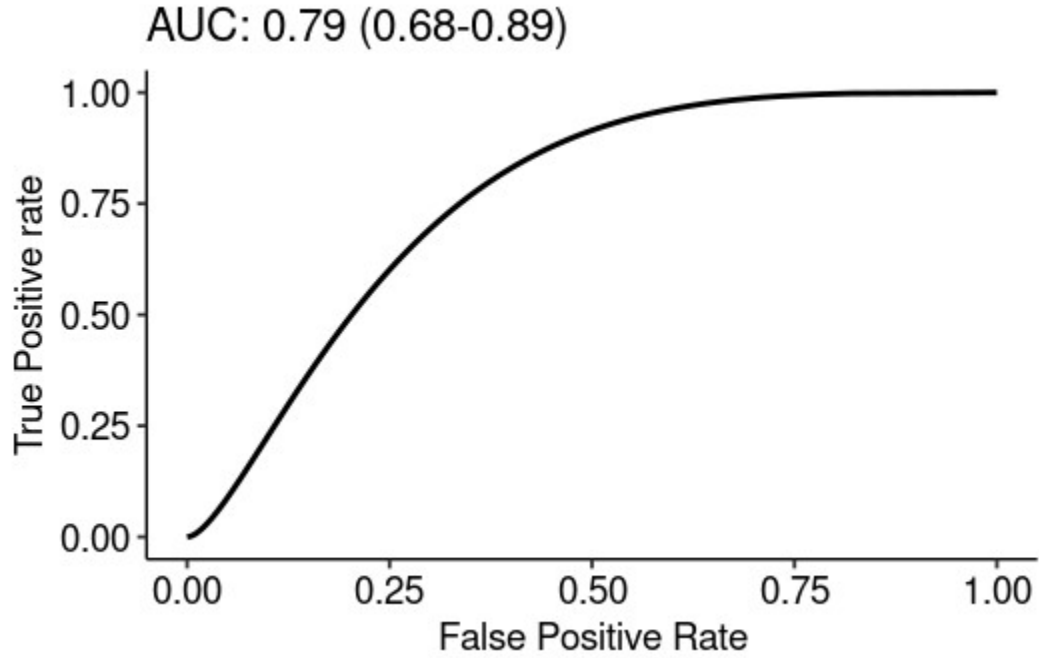


Figure 2: Receiver operator curve (ROC) at 5-years of the International IgAN risk prediction tool in the study population.

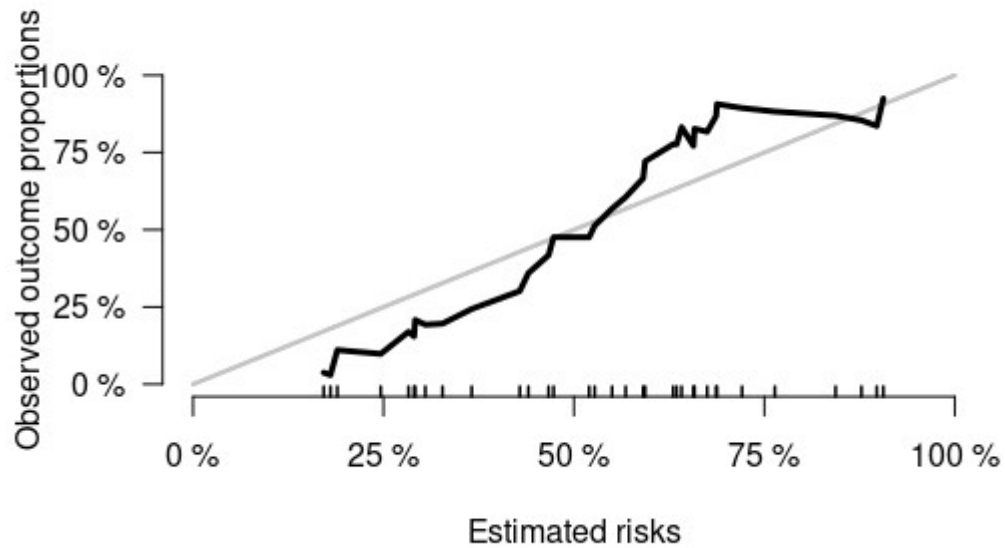


Figure 3: Calibration curve (Estimated risks plotted against Observed risk) at 5 years of the International IgAN risk prediction equation in the study population. Values below and above the 45 degrees line signify over and underprediction, respectively. The rug plot shows the distribution of the estimated risks.