SUPPLEMENTAL TABLES

Supplemental material table of contents

S1a. Baseline characteristics of sickle cell trait sensitivity analysis cohort.

Variable	Reference N = 2,450	SCT N = 1,240	p value
Demographics			
Mean age (SD), years	40 (±14)	40 (±14)	0.63
Age≥65 years	7%	7%	0.72
Female	79%	78%	0.88
Median follow up (IQR), years	9 (5 to 12)	8 (5 to 12)	0.70
Comorbidities			
Hypertension	27%	30%	0.06
Diabetes mellitus	23%	25%	0.15
Cardiovascular disease	14%	14%	0.82
Smoking			<0.01
Never	49%	44%	
Ever	16%	14%	
Missing	35%	41%	
AKI	11.4%	12.3%	0.42
Medications			
ACEi/ARBs	26%	29%	0.16
Aspirin	4%	4%	0.76
Statins	22%	23%	0.61
Hydroxyurea	0.3%	0.3%	0.85
Laboratory values			
Mean eGFR (SD), ml/min/1.73m^2	108 (±27)	104 (±27)	<0.01
eGFR categories			<0.01

≥ 120 ml/min/1.73m²	36%	30%	
90 – 119 ml/min/1.73m²	38%	40%	
60 – 89 ml/min/1.73m²	21%	25%	
30 – 60 ml/min/1.73m²	5%	5%	
Urine ACR			0.48
<30mg/g	3%	3%	
30-299mg/g	1%	1%	
>300mg/g	0%	0%	
Missing	96%	96%	
Mean hemoglobin*(SD), g/dL	12.4 (±1.7)	12.6 (±1.7)	0.04
Mean leukocyte count (SD)*, $x10^5$ cells/mm^3	7.5 (±3.9)	7.2 (±3.0)	0.03
Hemoglobin electrophoresis indications			<0.01
Anemia	47.6%	24.4%	
Perinatal testing	40.8%	38.3%	
Other	4.9%	16.5%	
Unknown	6.6%	20.7%	
Mean fractional hemoglobin S levels (SD)	-	37.0 (±4.3)	
Hemoglobin F > 0.4%	-	24.7%	
Mean fractional hemoglobin A levels (SD)	-	59.4 (±4.4)	
Mean fractional hemoglobin A_2 levels (SD)	-	3.2 (±0.7)	

SCT – sickle cell trait. ACEi/ARBs – angiotensin converting enzyme inhibitors or angiotensin receptor blockers. eGFR – estimated glomerular filtration rate (using Chronic Kidney Disease Epidemiology Collaboration [CKD-EPI] creatinine equation). Urine ACR – urine albumin: creatinine ratio. *Missing data: hemoglobin (5 reference, 8 SCT), leukocyte count (441 reference, 224 SCT), hemoglobin S (21), hemoglobin F (76), hemoglobin A (22) and hemoglobin A_2 (73).

S1b. Baseline characteristics of sickle cell disease sensitivity analysis cohort.

Variable	Reference N = 677	SCD N = 229	p value
Demographics			
Mean age (SD), years	33 (±12)	33 (±12)	0.74
Age≥65 years	0.4%	0.4%	0.99
Female	50%	51%	0.85
Median follow up (IQR), years	8 (4 to 11)	7 (4 to 11)	0.94
Comorbidities			
Hypertension	14%	11%	0.30
Diabetes mellitus	8%	10%	0.58
Cardiovascular disease	9%	28%	<0.01
Smoking			<0.01
Never	47%	40%	
Ever	15%	9%	
Missing	38%	51%	
AKI	8.0%	25.3%	<0.01
Medications			
ACEi/ARBs	15%	21%	0.05
Aspirin	2%	3%	0.34
Statins	12%	12%	0.91
Hydroxyurea	0.3%	47.6%	<0.01
Laboratory values			
Mean eGFR (SD), ml/min/1.73m^2	114 (±25)	128 (±32)	<0.01
eGFR categories			<0.01
≥ 120 ml/min/1.73m^2	42%	66%	
90 – 119 ml/min/1.73m^2	42%	22%	
60 – 89 ml/min/1.73m^2	14%	8%	
30 – 60 ml/min/1.73m^2	2%	4%	
Urine ACR			0.99
<30mg/g	1%	1%	
30-299mg/g	0%	0%	
>300mg/g	0%	0%	
Missing	99%	99%	
Mean hemoglobin* (SD), g/dL	13 (±2)	10 (±2)	<0.01
Mean leukocyte count (SD)*, x10^5 cells/mm^3	7.5 (±4.2)	12.2 (±6.4)	<0.01

SCD – sickle cell disease. ACEi/ARBs – angiotensin converting enzyme inhibitors or angiotensin receptor blockers. eGFR – estimated glomerular filtration rate (using Chronic Kidney Disease Epidemiology Collaboration [CKD-EPI] creatinine equation). Urine ACR – urine albumin: creatinine ratio. *Missing data: hemoglobin (2 reference, 0 SCD), leukocyte count (128 reference, 18 SCD).

S2. Difference in mean eGFR change per year in sickle cell trait and sickle cell disease compared to the reference in the sensitivity analysis cohort.

Exposure	Median eGFR values (IQR)	Unadjusted β, 95% CI ml/min/1.73m² per year	*Adjusted β, 95% CI ml/min/1.73m² per year
SCT Reference	18 (10-34)	0	0
SCT	19 (10-33)	-0.48 (-0.54 to -0.41)	-0.47 (-0.53 to -0.41)
SCD Reference	14 (7-25)	0	0
SCD	48 (22-125)	-1.07 (-1.20 to -0.94)	-1.09 (-1.22 to -0.97)
SCD vs SCT	-	-0.85 (-0.94 to -0.76)	-0.88 (-0.97 to -0.79)

SCT – sickle cell trait. SCD – sickle cell disease. *Adjusted for baseline age, sex, hypertension, diabetes mellitus, history of cardiovascular disease, smoking status, acute kidney injury, ACEi/ARB use, urine albumin: creatinine ratio categories and baseline eGFR. SCT model was adjusted for hemoglobin electrophoresis indication.

S3. Baseline characteristics of cohort with baseline eGFR ≥ 65 ml/min/1.73m²

Variable	Reference N = 8,376	SCT N = 1,141	p value	SCD N = 216	p value
Demographics					
Mean age (SD), years	35 (±11)	38 (±13)	<0.01	32 (±11)	<0.01
Age≥65 years	2%	4%	<0.01	0%	0.03
Female	89%	80%	<0.01	50%	<0.01
Median follow up (IQR), years	8(±4)	8(±4)	0.41	8(±4)	0.31
Comorbidities					
Hypertension	18%	26%	<0.01	9%	<0.01
Diabetes mellitus	14%	21%	<0.01	8%	0.01
Cardiovascular disease	8%	12%	<0.01	26%	<0.01
Smoking			<0.01		<0.01
Never	50%	45%		40%	
Ever	16%	14%		9%	
Missing	34%	41%		51%	
AKI	5.3%	8.4%	<0.01	22.2%	<0.01
Medications					
ACEi/ARBs	18%	24%	<0.01	19%	0.71
Aspirin	20%	24%	<0.01	34%	<0.01
Statins	13%	29%	<0.01	10%	0.13
Hydroxyurea	0.1%	0.3%	0.34	49.5%	<0.01
Laboratory values					
Mean eGFR (SD), ml/min/1.73m²	117 (±24)	108 (±23)	<0.01	133 (±27)	<0.01
eGFR categories			<0.01		<0.01
≥ 120 ml/min/1.73m²	47	32		69	
90 – 119 ml/min/1.73m²	37	43		23	

65 – 89 ml/min/1.73m²	16	24		8	
Urine ACR			<0.01		0.33
<30mg/g	2%	2%		0%	
30-299mg/g	0%	1%		0%	
>300mg/g	0%	0%		0%	
Missing	98%	97%		100%	
Hemoglobin*	12.1 (±1.6)	12.3 (±1.6)	<0.01	9.7 (±2.1)	<0.01
Hemoglobin electrophoresis indications			<0.01		
Anemia	37.8%	22.6%		-	-
Perinatal testing	53.6%	41.5%		-	-
Other	3.3%	16.3%		-	-
Unknown	5.3%	19.6%		-	-

SCT – sickle cell trait. SCD – sickle cell disease. ACEi/ARBs – angiotensin converting enzyme inhibitors or angiotensin receptor blockers. eGFR – estimated glomerular filtration rate (using Chronic Kidney Disease Epidemiology Collaboration [CKD-EPI] creatinine equation). Urine ACR – urine albumin: creatinine ratio. *Missing data: hemoglobin (13 reference, 8 SCT and 0 SCD).

S4. Association of sickle cell trait and sickle cell disease with incident stage 3 chronic kidney disease using sensitivity analysis cohort.

Exposure	Events, n (%)	Unadjusted (HR; 95% CI)	Adjusted* (HR; 95% CI)
SCT Reference	254 (11%)	1	1
SCT	151 (13%)	1.25 (1.02 to 1.53)	1.15 (0.91 to 1.44)
SCD Reference	26 (4%)	1	1
SCD	16 (7%)	2.10 (1.12 to 3.93)	2.51 (1.24 to 5.10)

SCT – sickle cell trait. SCD – sickle cell disease. *Adjusted for baseline age, sex, hypertension, diabetes mellitus, history of cardiovascular disease, smoking status, acute kidney injury, ACEi/ARB use, urine albumin: creatinine ratio categories and baseline eGFR. SCT was also adjusted for hemoglobin electrophoresis indications.

S5. Unadjusted and adjusted coefficients of the difference in mean eGFR change per year in sickle cell disease by severity.

Exposure	Number of eGFR values (IQR)	Unadjusted β, 95% CI ml/min/1.73m² per year	*Adjusted β, 95% CI ml/min/1.73m² per year
Reference	17 (10-29)	0	0
SCD - severe	84 (29-158)	-1.70 (-1.78 to -1.62)	-1.71 (-1.79 to -1.63)
SCD – non severe	31 (17-54)	0.47 (0.32 to 0.63)	0.45 (0.29 to 0.60)

*Adjusted for baseline age, sex, hypertension, diabetes mellitus, history of cardiovascular disease, smoking status, acute kidney injury, ACEi/ARB use, urine albumin: creatinine ratio categories and baseline eGFR.

S6. Adjusted coefficients describing mean annual change in eGFR by baseline eGFR category in

reference, sickle cell trait, and sickle cell disease patients.

Baseline eGFR category	**Reference β (95% CI) ml/min/1.73m² per year**	**SCT β (95% CI) ml/min/1.73m² per year**	**SCD β (95% CI) ml/min/1.73m² per year**
All	-1.21 (-1.23 to -1.19)	-1.64 (-1.69 to -1.59)	-2.48 (-2.56 to -2.40)
≥ 120 ml/min/1.73m²	-1.69 (-1.73 to -1.66)	-1.57 (-1.70 to -1.45)	-2.47 (-2.56 to -2.38)
90-119 ml/min/1.73m²	-0.87 (-0.90 to -0.83)	-1.88 (-1.98 to -1.79)	-2.85 (-3.06 to -2.64)
60-89 ml/min/1.73m²	-1.03 (-1.07 to -0.98)	-1.53 (-1.61 to -1.46)	-1.60 (-2.08 to -1.13)
< 60 ml/min/1.73m²	-1.08 (-1.15 to -1.02)	-1.38 (-1.49 to -1.28)	0.60** (-0.14 to 1.34)

SCT – sickle cell trait. SCD – sickle cell disease. ACEi/ARBs – angiotensin converting enzyme inhibitors or angiotensin receptor blockers. eGFR – estimated glomerular filtration rate (using Chronic Kidney Disease Epidemiology Collaboration [CKD-EPI] creatinine equation). Urine ACR – urine albumin: creatinine ratio. *All coefficients were adjusted for baseline age, sex, hypertension, diabetes mellitus, history of cardiovascular disease, smoking status, acute kidney injury, ACEi/ARB use, and urine albumin: creatinine ratio categories. Reference and SCT were also adjusted for hemoglobin electrophoresis indications. SCD was also adjusted for hydroxyurea.

**Only 9 patients.

APPENDIX

Algorithms for adjudicating covariates using ICD codes

Cardiovascular disease: coronary artery disease

The presence of at least 2 diagnosis codes during follow-up was required. We used the following codes: 410.x, 411.x, I20.0, I21.x, I24.x, I25.1x, I25.7x. These codes were evaluated on a subset of African 696 American CKD patients from the Partners RPDR database with coronary artery disease confirmed on chart review (using results of cardiac catheterizations, stress tests and/or physician documentation of a myocardial infarction).

Sensitivity: 81%, Specificity: 81%, negative predictive value (NPV): 93%, positive predictive value (PPV): 57%

Cardiovascular disease: stroke

The presence of at least 2 diagnosis codes during follow-up was required. We used the following codes: 362.30, 362.31, 362.32, 362.33, 362.34, 433.x, 434.x, 435.x, 436.x, 431.x, G45.x, H34.1x, I63.x, I65.x, I61.x. These codes were evaluated on a subset of 696 African American CKD patients from the Partners RPDR database with stroke confirmed on chart review (using brain imaging radiology reports and/or physician documentation of hemorrhagic or embolic stroke).

Sensitivity: 87%, Specificity: 88%, NPV: 97%, PPV: 59%

Hypertension

The presence of at least 15 diagnosis codes at different times during follow-up was required. We used the following codes: 997.91, 401.0, 401.1, 401.9, 402.00, 402.01, 402.10, 402.91, 402.90, 403.01, 403.00, 404.01, 404.03, 404.91, 404.92, 404.93, 404.11, 404.13, 404.12, 404.10, 402.11, 403.11, 403.10, I10.x, I11.x, I12.x, I13.x, I15.x. These codes were evaluated on a subset of 696

African American CKD patients from the Partners RPDR database with hypertension confirmed on chart review (using multiple physician notes).

Sensitivity: 80%, Specificity: 76%, NPV: 36%, PPV: 96%

Diabetes mellitus

The presence of at least 5 diagnosis codes at different times during follow-up was required. We used the following codes: 250.x, E10.x, E11.x, E12.x, E13.x, E14.x. These codes were evaluated on a subset of 696 African American CKD patients from the Partners RPDR database with diabetes mellitus confirmed on chart review (using multiple physician notes).

Sensitivity: 98%, Specificity: 82%, NPV: 97%, PPV: 86%

Acute Kidney Injury

The presence of at least one diagnosis code was required. We used the following codes: 584.x and N17.x.

Sensitivity: 17%, Specificity: 99.6%, NPV: 92%, PPV: 81.8% (Grams, Morgan E., et al. "Performance and Limitations of Administrative Data in the Identification of AKI." *Clinical Journal of the American Society of Nephrology : CJASN*, vol. 9, no. 4, American Society of Nephrology, Apr. 2014, pp. 682–89).

Sickle cell crisis and acute chest syndrome

The presence of at least one diagnosis code was required. We used the following codes: 282.62,282.64, 282.42, 282.69, 517.3, D57.0*, D57.21*, D57.41*, D57.81*.

Vaso-occlusive crisis codes: Sensitivity: 91.8%, Specificity: 75%, NPV: 94.1%, PPV: 67.7%

Acute chest syndrome codes: Sensitivity: 92.6%, Specificity: 98%, NPV: 99.5%, PPV: 75%

(Ting, Michelle, et al. "Accuracy of Administrative Coding for Sickle Cell Disease." *Blood*, vol. 126, no. 23, 2015).

Sickle Cell Disease Phenotypes and Categorization

Hgb SS – n = 139 (60.4%)

Hgb SC – n = 72 (31.3%)

HGB S-beta thalassemia zero – n = 3 (1.3%)

Hgb S-beta thalassemia trait – n = 14 (6.1%)

Hgb SG – n = 2 (0.9%)

Severe phenotype – Hgb SS and Hgb S-beta thalassemia zero

Non-severe phenotype – Hgb SC, Hgb S-beta thalassemia trait and Hgb SG.

Indications for hemoglobin electrophoresis testing

Anemia	36.7% (3,743)
Routine perinatal/early pregnancy/infertility/reproductive partner testing in sickle cell pregnancy	48.9% (4,992)
Abnormal complete blood count (microcytosis, macrocytosis or high red blood cell count)	2.0% (208)
Other abnormal labs (rhabdmyolysis, unexplained elected liver function tests)	1% (99)
Family history of sickle cell disorders	0.9% (88)
Rule out sickle cell disorders as etiology of acute presentations (hypercoagulable workup, recurrent gallstones, papillary necrosis, bone infarct).	0.9% (91)
Other reasons in an otherwise healthy patient (athlete, pre-employment).	0.2% (24)
Unknown (no documentation of indication and no abnormal labs at time of testing).	7.2% (735)
Sickle cell disease	2.2% (230)