PREVALENCE AND CHARACTERISATION OF CHRONIC KIDNEY DISEASE IN A RURAL SETTING IN SOUTH AFRICA

June Fabian

Supervisors

Saraladevi Naicker

Jaya Anna George

Alisha Nicole Wade

A thesis submitted to the School of Public Health,

Faculty of Health Sciences,

University of the Witwatersrand, Johannesburg,

in fulfilment of the requirements for the degree of Doctor of Philosophy

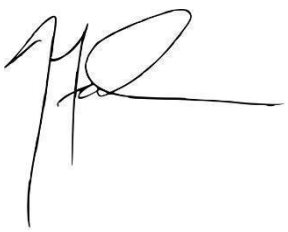
Johannesburg, South Africa, 22 October 2021

Declaration

I, June Fabian, student number 8600952/N, declare this is my original work, submitted for the degree of Doctor of Philosophy in the School of Public Health, Faculty of Health Sciences, University of Witwatersrand. This work has not been submitted before for any degree or examination at this or any other university.

For the duration of my doctoral research, I worked closely with colleagues in Malawi, Uganda, and the London School of Hygiene and Tropical Medicine. Although my primary focus was South Africa, it became clear we stood to gain far more through a synergistic collaboration than pursuing similar projects independently. Originating from a preliminary discussion between Moffat Nyirenda and Stephen Tollman, who jointly facilitated the grouping of partners in the African Research on Kidney Disease (ARK) Consortium, we have worked together over the last six years to bring our individual and collective work to fruition. As the co-principal investigator for the ARK South Africa project, some components of this work are confined to South Africa, and others reflect my original work from South Africa, now incorporated into the broader ARK Consortium. I am first author for each paper presented in this body of work, and for each component of the research and associated papers, my contribution has been detailed, as have been the contributions of others, where appropriate.

This compilation of my original work presents an integrated narrative of the research undertaken, followed by the resultant publications.



June Fabian, 22 October 2021

Dedication

To Sandra Maytham Bailey, for always lifting me on your shoulders, ensuring I keep in sight on the horizon, what I dream of in my heart.

To our beloved men, Christopher and Nicholas, one day I hope my endless obsession with understanding more will enrich your lives.

Format for this doctoral thesis

The format for this doctoral thesis is by publication and follows that prescribed by the School of Public Health, Faculty of Health Sciences, University of the Witwatersrand. Personal reflections from my doctoral research comprise the prologue, titled "Circling back to what was hidden-in-plain-sight". This personal narrative contains carefully selected images taken throughout the duration of my doctoral research. The images are intended to situate the work within the rural Bushbuckridge community – and highlight my personal learning extending beyond the science. After the prologue, the background, study justification, and methods are self-explanatory. Because this thesis is by publication, the results section presents the main themes of the research, drawing together a cohesive framework that contextualises each individual published paper – rather than replicate detail in already published papers.

Letter of contribution: 22 October 2021

June Fabian, student number 8600952/N, made the following contributions to the project and to writing of the papers for the doctoral thesis titled: "Prevalence and characterisation of chronic kidney disease in a rural setting in South Africa".

Contribution from the candidate

We conducted the project in three phases, with the first two at the MRC/Wits-Agincourt Rural Public Health and Health Transitions Research Unit in Bushbuckridge, South Africa. The first phase comprised a population-based study to determine the prevalence of chronic kidney disease (CKD) and associated risk factors. The second phase measured kidney function in a subsample of participants from the first phase, using plasma clearance of iohexol. The final phase comprised analysis and interpretation of our findings from the iohexol studies completed in Uganda, Malawi and South Africa. For this component we formed a working group from among the broader ARK Consortium.

I led the grant application that funded this work; for the first two phases of the study, I wrote the study protocols and standard operating procedures, and obtained initial and subsequent amendment approvals from the University of Witwatersrand Medical Human Research Ethics Committee. I visited the study site for one to two weeks per month for the

duration of the project, depending on what was required. F. Xavier Gómez-Olivé, the Agincourt Research Manager and his team mentored me for the implementation and timeous completion of our project. In Agincourt, I was integrally involved with recruiting, training and managing the ARK study team. The team included a project manager, a fieldwork supervisor, eight fieldworkers, four nurses and four drivers. I worked closely with the Agincourt laboratory manager and two laboratory technicians regarding specimen tracking, processing, storage and shipping. Together with the laboratory team, we set up urine microscopy for schistosomiasis, in close collaboration with John Frean and his staff at the National Institute for Communicable Diseases. I procured stock and ensured quality control for the field work, laboratory and shipping procedures. To represent the ARK Study, I worked with the engagement team in Agincourt to meet community members, healthcare workers at primary care clinics and Tintswalo Hospital, and stakeholders such as the District Department of Health.

For the last phase, a pooled analysis of the iohexol data across the three ARK sites, I have been a core member of the group, meeting weekly for the last 18 months to analyse, interpret and write up our data. Unexpectedly, the last phase of the work has been the most challenging for me, notwithstanding the added impact we collectively had to face from the COVID-19 pandemic.

The final component of the work I have contributed was unplanned. Initially, the statistical analysis plan involved splitting our dataset for a validation component. However, as the analysis evolved, we reviewed our strategy and shifted to using the whole dataset and sourcing an external dataset. The responsibility for sourcing the external dataset fell into my hands, as neither our colleagues in Malawi nor Uganda had access to radionuclide services in their hospitals. Collating the external dataset was enabled by our colleagues in the Department of Nuclear Medicine at Charlotte Maxeke Johannesburg Academic Hospital and the Transplant Unit at Wits Donald Gordon Medical Centre.

Aside from specific project-related contributions, we had regular meetings with the grant fund managers at Wits Health Consortium, and with our funders, GSK, SA-MRC, and the Newton Fund, providing feedback on our progress and findings. I applied for additional funding that we needed for our project, including a Faculty Research Grant from the Faculty

of Health Sciences (University of the Witwatersrand) which assisted with the installing of a dual-energy x-ray absorptiometry (DXA) machine in Agincourt, generously donated by Professor Shane Norris of the Wits Developmental Pathways to Health Research Unit. This enabled the South African component of the ARK study to perform the gold standard assessment of lean muscle mass, critical to our work. This standard became the reference for body composition studies in our ARK Malawi and Uganda sites too.

With Professor Jaya George, Head of Department for Chemical Pathology (University of Witwatersrand), I applied for additional funding from the NHLS for doing cystatin C measurements and establishing the laboratory iohexol method. While not initially conceptualised as part of our work, evaluating cystatin C has critically informed our findings. Unbeknown to us, we became the referral centre for iohexol testing for Uganda and Malawi, who were unable to set up the iohexol method as originally intended.

Professor Jaya George was instrumental in procuring, at no cost, capillary and venous point-of-care devices for creatinine to validate against iohexol measured GFR, as well as evaluating a urine dipstick and point-of-care device on stored urine samples from the iohexol measured GFR study. I was responsible for setting up this point-of-care substudy in the Agincourt Research Clinic. I wrote the standard operating procedures for the clinic and laboratory staff, trained the clinic and laboratory staff of how to use the devices, and established the internal quality control procedures for the point-of-care devices with the laboratory manager.

Through an ongoing collaboration with Professor Cheryl Winkler, National Cancer Institute, USA the *APOL1* genotyping was performed for our population sample at no additional costs other than that of shipping the samples to her laboratory. For this component, I was responsible for finalising the Material Transfer Agreement and the export permit for the DNA samples.

Publications arising from this work

Paper one

Status: under review – Kidney International Journal

Prevalence and associated risk factors for chronic kidney disease in a rural South African setting

*June Fabian, Mwawi Gondwe, Nokthula Mayindi, Shingirai Chipungu, Bongekile Khoza, Alisha N. Wade, F. Xavier Gómez-Olivé, Laurie A. Tomlinson, Michele Ramsay, Stephen Tollman, Cheryl Winkler, [*]Jaya George, [*]Saraladevi Naicker, on behalf of the ARK Consortium. [*Joint senior authors]*

My contribution: I collated the data, performed the data cleaning and statistical analysis, wrote the first draft of the paper, collated comments from co-authors and finalised the manuscript.

Paper two

Methods and reporting of kidney function: a systematic review of studies from Sub-Saharan Africa (95)

June Fabian, Jaya A. George, Harriet R. Etheredge, Manuel van Deventer, Robert Kalyesubula, Alisha N. Wade, Laurie A. Tomlinson, Stephen Tollman, Saraladevi Naicker, on behalf of the African Research in Kidney Disease (ARK) Working Group. Clinical Kidney Journal 2019, vol. 12, no. 6, 778-787 https://doi.org/10.1093/ckj/sfz089

My contribution: I registered the systematic review with PROSPERO; coordinated the systematic review with co-authors; completed the literature search; collated and reviewed abstracts; collated and reviewed full manuscripts; designed the data collection tool for the review; analysed the data; wrote and finalised the manuscript; revised and resubmitted the manuscript based on reviewers' comments.

Paper three

How to estimate glomerular filtration rate in Sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study (85)

[*]Robert Kalyesubula, [*]June Fabian, Wisdom Nakanga, Robert Newton, Billy Ssebunnya, Josephine Prynn, Jaya A. George, Alisha N. Wade, Janet Seeley, Dorothea Nitsch, Christian Holm Hansen, Moffat Nyirenda, Liam Smeeth, Saraladevi Naicker, Amelia C. Crampin, and Laurie A. Tomlinson. [[*]Joint first authors] BMC Nephrology 2020, vol. 21, no. 20, 1-12.

https://doi.org/10.1186/s12882-020-1688-0

My contribution: I co-authored the manuscript with Robert Kalyesubula; wrote the protocol and all standard operating procedures for the South African study site; added the information regarding the methodology of the ARK study for the South African site to the paper.

Paper four

Status: under review – New England Journal of Medicine

Measurement of kidney function in Sub-Saharan Africa

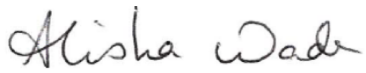
[*]June Fabian, [*]Robert Kalyesubula, Joseph Mkandawire, Christian Holm Hansen, Dorothea Nitsch, Eustasius Musenge, Saraladevi Naicker, Wisdom P. Nakanga, Josephine E. Prynn, Liam Smeeth, Tracy Snyman, Billy Ssebunnya, Stephen Tollman, Amelia C. Crampin, Robert Newton, Jaya A. George, Laurie A. Tomlinson, on behalf of the African Research on Kidney Disease (ARK) Consortium.
[[*]Joint first authors.]

My contribution: aside from my contribution stated in the preface, I cleaned the South African dataset, submitted it for a pooled analysis with collaborating sites in Malawi and Uganda. I was a core member of the working group for the analysis of the pooled iohexol data. I took the lead on writing the manuscript, collated the supplementary appendix, incorporated feedback from co-authors and finalised the manuscript. I presented

preliminary data to the broader ARK Consortium during the course of the analyses for review, and collated data for an external expert review. The primary analyses were led by the statistician, Professor Christian Holm Hansen from the MRC/UVRI & London School of Hygiene and Tropical Medicine Research Unit, Entebbe, Uganda, with additional input from the statistical team at Wits, led by Professors Jonathan Levin and Eustasius Musenge.

Saraladevi Naicker (supervisor)

Jaya Anna George (supervisor)



Alisha Nicole Wade (supervisor)

June Fabian (student)

Abstract

Introduction: In South Africa (SA) and Sub-Saharan Africa (SSA), the burden of chronic kidney disease (CKD) is unknown, associated risk is poorly characterised, and equations to estimate glomerular filtration rate (eGFR) have limited validation. This study aimed to determine the population prevalence of CKD, identify associated risk factors, and measure kidney function using iohexol (mGFR) as the reference for comparing eGFR equations. If inaccurate, a model would be developed to better predict eGFR, therefore more accurately determining CKD burden.

Methods: The research was conducted in the MRC/Wits-Agincourt Unit, Bushbuckridge, Mpumalanga Province. To determine population prevalence of CKD and associated risk, 2 021/2 759 adults, aged 20 to 80 years were recruited (prevalence of CKD negligible in those younger than 20). CKD prevalence was determined by measuring albuminuria (albumin: creatinine ratio (ACR)≥3mg/mmol) and estimating GFR (eGFR<60ml/min/1.73m^2), using the CKD-EPI (creatinine) equation without ethnicity coefficient. CKD was confirmed with repeat screening after three months. Genotyping determined *APOL1* renal risk variants. To measure GFR, a subsample of ~1 000 adults were recruited, stratified by eGFR and sex. Serum creatinine, cystatin C and GFR were measured using the slope-intercept method for iohexol plasma clearance (mGFR). eGFR equations were compared to mGFR, a new eGFR equation was modelled and validated, and multiple imputation modelling trained on mGFR was used to predict CKD. Data was pooled with Malawi and Uganda for analysis.

Results: In SA, the WHO age standardised population prevalence of CKD was 4.0% (95% CI 3.1-4.9). Risk factors comprised high risk *APOL1* genotypes (OR 1.95; 95% CI 1.20-3.17); hypertension (OR 3.34; 95% CI 1.92-5.79); diabetes (OR 3.28 95% CI 1.61-6.70), HIV infection (OR 1.89; 95% CI 1.18-3.03) and hyperuricaemia (OR 1.85; 95% CI 1.13-3.05). Pooled data for mGFR included 2 578 of 3 025 participants. Overall and by country, creatinine-based equations overestimated kidney function compared to mGFR. The greatest bias occurred at low kidney function, where the proportion with mGFR <60ml/min/1.73m^2 was more than double that estimated from creatinine. A new model for estimating GFR did not outperform existing equations, and no eGFR equation achieved the benchmark of estimated GFR within ±30% of mGFR for ≥75% of participants. Imputation modelling

estimated prevalence of kidney disease as two- to three-fold higher than creatinine-based estimates across six SSA countries.

Conclusion: In South Africa, CKD is prevalent and associated with high risk *APOL1* genotypes, HIV infection and cardiometabolic diseases. In Sub-Saharan Africa, estimating GFR using serum creatinine substantially underestimated individual and population-level burdens of CKD, and cystatin C may be a preferable biomarker. Our findings have implications for individual care and public health policy, supporting implementation of screening for early detection and prevention of progression of CKD in those who are at risk.

Acknowledgements

The Agincourt community, who embraced me and my unknowing, and gently taught me what I needed to learn, one step at a time.

Kathleen Kahn, for warmly welcoming me to the PhD program in the School of Public Health and suggesting we apply for the grant that funded this project; my personal turning point for registering my PhD.

Stephen Tollman, for listening and asking hard questions, and for your wisdom and vision.

Kathy and Steve, for supporting me personally, and the ARK project, in countless ways.

My supervisors, Saraladevi Naicker, for sage insights from a lifetime of nephrology experience, always stepping in when needed (and constant reminders of a deadline I had to meet); Jaya George, in typically understated style, for critical laboratory expertise and unwavering commitment to the ARK project, including establishing iohexol assays as a first on the continent; and Alisha Wade, for those invaluable "what's not working" and "how are we going to fix it" meetings, making the unknown terrain navigable.

The innumerable people who made ARK-SA happen in Johannesburg and Bushbuckridge. Beyond the science, this amounted to finding meeting venues, solving stock crises, *those* purchase orders, booking a place to stay and fixing punctures. Dawn Dalby, the goddess of all things, who threw me numerous lifelines just when I needed them.

My colleagues in the ARK Consortium, my brother, Robert Kalyesubula, and our expanding network of collaborators, whose contributions, knowledge and expertise have been an incredible source of learning and enrichment for me.

Our funders and grant managers: Candice Roux from the South African Medical Research Council; The Newton Fund; James van Hasselt, Lindsay Kendall and Maria Davy from GlaxoSmithKline; Faculty of Health Sciences (University of the Witwatersrand); the National Health Laboratory Service; and the International Society of Nephrology. Dorcas Lesolang and Vanessa Du Preez, for your never-ending patience, big mugs of coffee, and management of our grant funding.

Michele Ramsay, aside from much more, for generously sharing data on behalf of the Africa Wits-International Network for the Demographic Evaluation of Populations and their Health Partnership for Genomic Studies (AWI-Gen) in Ghana, Kenya, Burkina Faso and South Africa.

Mboyo-Di-Tamba Vangu, Head of the Department of Nuclear Medicine (Charlotte Maxeke Johannesburg Academic Hospital), Geoffrey Candy (University of the Witwatersrand) and Jean Botha, Head of the Wits Transplant Unit (Wits Donald Gordon Medical Centre, University of the Witwatersrand) for contributing data for the external validation dataset.

Cheryl Winkler from the National Cancer Institute, USA for the *APOL1* genotyping.

Catherine Burns for your effusive energy and your brilliant, incisive mind that always stretches mine.

Sue Tager for supporting me without question, not knowing when the end was in sight, assuring me somehow "it would all be alright in the end".

The amazing Wits Donald Gordon Medical Centre team, for stepping up in ways I could not have imagined.

Heather Maher, one of my pillars, for knowing what I need even when I don't.

Harriet Etheredge, for advice, endless encouragement and the "surviving PhD pearls".

Janet Bartlet and Marike Mapham for editing, formatting and visualising, with endless reworks until we "got it right".

Family and friends, my tribe (eclectic mavericks and misfits come to mind), whose love and care pulled me to the finishing line with spottify songs, whats app's from the tops of mountains, and freshly cut garden flower drop-offs.

My mother, Rona Fabian, who by her very existence defines tenacity. And fortunately made it her life's mission to ensure I received a tertiary education. I am deeply grateful.

Kindness. Generosity. Heaped on me by so many. The culmination of this body of work reflects the immense power of the collective. In return, I will pass on your gifts.

Prologue: Circling back to what was hidden-in-plain-sight

My first experience of Bushbuckridge was in 2000, when I arrived as one of the first group of community service doctors at Tintswalo Hospital. During that year, we never saw a single patient with kidney failure because there were no laboratory reagents to perform creatinine assays. We had intermittent water supply, as did the community. And we performed many procedures considered out of our scope – but we did them knowing that if we did nothing, the patient would most certainly demise.

Fifteen years later I circled back to the MRC/Wits-Agincourt Unit in the heart of the Bushbuckridge community.

This north-eastern part of Mpumalanga was originally the land of the Mapulana people and then included the Xitsonga and Eswatini tribes who came across from the east. The area was declared one of three “bantustans” or ”homelands” during apartheid. The community has a complex and complicated history, made worse by its incorporation into the Limpopo Province in 1994. After a bitter dispute lasting 11 years, the Bushbuckridge Municipality was reintegrated into the Mpumalanga Province. Grounds for the dispute were the community’s fears of being marginalized, poor service delivery and delays in infrastructure development. The community is still plagued by water insecurity – one of many service delivery failures.

Bushbuckridge: *While electrification is reaching most homes, communal taps still service much of the community.*

The influence of European settlers and mission hospitals remains tangible in many villages with names such as Lilliput, Merry Pebble Stream, Ireagh and Kildare. Even the practice of topiary, a style of shaping hedgerows taught by the Scottish missionaries, is a common sight in the community. These ficus bush (*Ficus Benjamina*) sculptures, despite their long and textured history, are proudly displayed in many community gardens, clinics and hospitals.

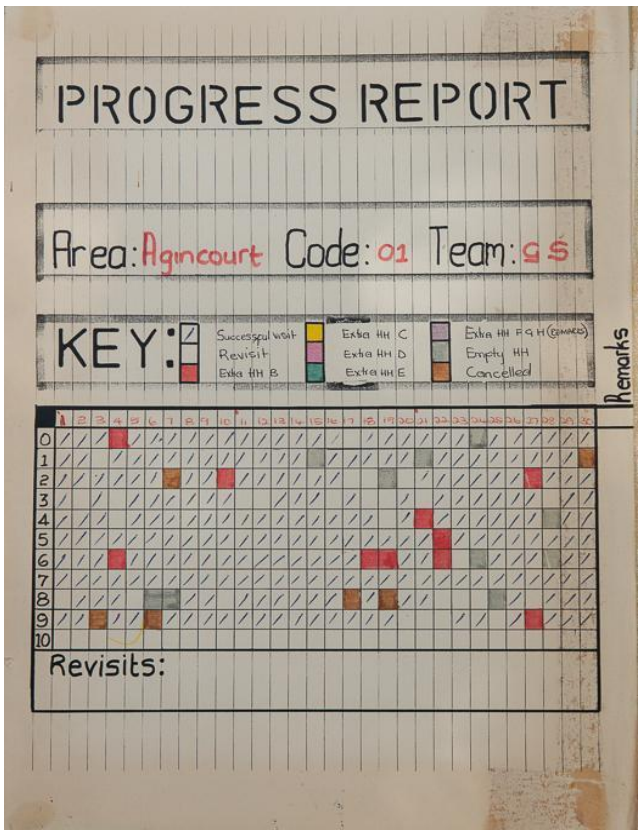
Ficus topiary hedge: *The entrance to the Agincourt Primary Health Care Clinic, with its young topiary hedge of swirls, hearts and loops.*

From the 1800s, missionary hospitals such as Tintswalo, were often important hubs in the community, with food gardens and schools expanding beyond the hospital service. They became places of improvisation and courage, particularly as the impact of apartheid began to take hold, with governmental authorities often threatening their existence. The role of these community-based hospitals continued to be a source of comfort – often doubling up as impromptu places of worship, teaching and education – and provided a budding informal trading post for hawkers of all sorts.

In 1982, the Community Health Department at Wits established a Health Services Development Unit (HSDU) in Bushbuckridge. The aim was to improve local primary health services in the “homeland”, but any initiatives were heavily curtailed by the government.

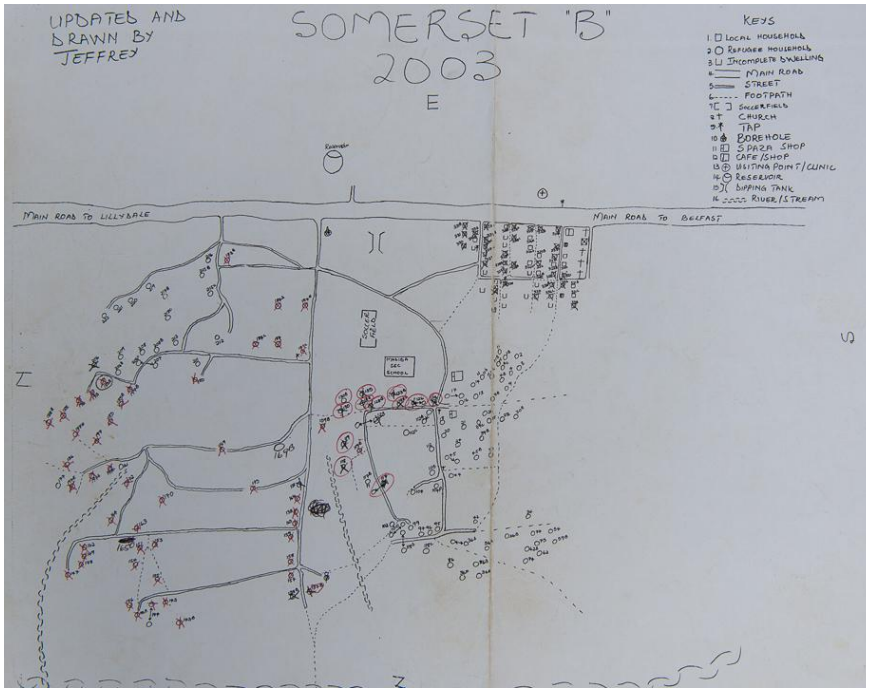
After the dismantling of apartheid in the early 1990s, the Gazankulu Health Services entered a long-awaited partnership with Wits to address health systems development.

In parallel, the Agincourt field site was established and began to contribute important demographic and health surveillance data, detailing vital population health information that had never been collected – a deliberate strategy – prior to 1992. As an example, this early data was among the first to inform the South African government regarding population birth and death rates, infant and maternal mortality. Later, the information collected evolved to include vital statistics about HIV infection and tracked the evolution of the epidemic.



Agincourt Field Site: *Progress Report – an early annual census form held in the Agincourt visual archives.*

As the study site began to expand, maps of each village were carefully charted by the staff, some of whom still work in the study site. Given the persistent absence of infrastructure in the villages – no tarred roads, no street names or house numbers – the Health and Demographic Surveillance System (HDSS) site has GIS location mapping for each participating household.



Top: *A hand drawn map from 2003 of Somerset village, part of the field site.*

Bottom: *Current GIS location technology for mapping the villages in the HDSS site.*

With my clinician's background, I had no appreciation of the depth and scope of the Agincourt Unit's work and the immense power that a population-based longitudinal research platform provides for research. However, accessing the platform doesn't necessarily make for a successful project. While science is critical, the crux lies in the implementation. I realised our success needed much more than the science and would be determined by the

study team and the community, provided I was willing to listen. One of my first lessons was that consent is communal, conducted as a ritual the team called “approaching the household”. This ritual includes entering the potential participant’s yard, after which we would be invited to sit in the shade of a tree and await the head of the household. This could take any amount of time and only thereafter would the discussion about the study begin. Without consent of the head of the household, there would be no participation by the individual.

Approaching the household: *June Fabian and ARK study fieldworker, Surprise Ubisi, await the arrival of the head of the household.*

The Agincourt HDSS platform has a strong commitment to engage with the community ahead of each new project, and to provide feedback to each village for completed studies. On one occasion, an elderly woman refused to participate in the ARK study because she wanted her individual results, not a general outcome summary from the research. We discussed the points raised and the team echoed similar sentiments, which was especially relevant given the almost 30-year relationship between the Agincourt HDSS and the community. While individual feedback is not a prerequisite from the Wits Human Research Ethics Committee, we debated the feasibility and cost of including individual feedback within our timeline. The team were totally committed, willing to work on weekends, and felt the

complexity of the iohexol component of the study required an additional investment from us. We designed participant feedback forms detailing information like height, weight, blood pressure and point-of-care results in Shangaan (Xitsonga) and then went ahead.

Home visit: *The communal home visit including three generations – the elderly head of household on the right, her daughter (participant), and grandchild (foreground).*

Challenges: *Summer temperatures regularly exceeded 38º Celsius, forcing the team to find innovative ways of storing collected samples while ensuring point-of-care devices continued to work in the heat – a practical downside to using imported heat-sensitive devices from Norway.*

After the first phase of field work was completed, we had the daunting task of setting up the Agincourt Research Clinic for iohexol measurement. Despite being commonplace in high-income countries, cohorts of kidney disease patients from which existing samples are drawn for measured GFR studies, do not exist in Sub-Saharan Africa (SSA). Neither do established centres of excellence exist that can administer iohexol and measure GFR. We needed to create both our own cohorts, and then establish the iohexol GFR methods. It was during this phase that not speaking Shangaan (Xitsonga) became incredibly frustrating for myself and the team, who were learning about complex concepts in a second language. Shangaan is rich and expressive and it took the help of Busi, the ARK fieldwork supervisor, to interpret the often-difficult English science parlance into an easily comprehensible Shangaan context. The lesson I learned was to step back. Busi would say, “Junie, just give me 2 hours” and when I returned, we were back on track.

Soon the humble building was transformed into a 7-station iohexol day-clinic. Blue surgical gowns were washed, participants received a meal after a six-hour fast, the iohexol IVI administration room was made sterile, and large wall mounted stopwatches kept the team to a strict schedule for timing of blood draws.

Training: *The training of the nursing staff often required practicing on each other. Here Busi, our fieldworker supervisor, becomes the participant.*

For the iohexol GFR cannulation procedure, participants were asked to fast from the previous night. They were collected by the study drivers each morning and screened on arrival. The iohexol procedure required interval sampling for at least 4 hours after administration of the IVI push, meaning a minimum 4 hours' stay in the research laboratory. During their time with us, participants underwent numerous additional tests, some of which included kidney ultrasound, carotid intima-medial thickness measurement and ECG measurement. This meant participants would also need a meal and taken back to their homes. All in all, a big ask.

Logistical challenges: *With no running water in the main iohexol station, the team improvised with water dispensers and collection buckets.*

Wash day: *Patient gowns washed and ready each day.*

Meal time: *A field worker offers a clean bowl of water for a participant to wash their hands before receiving a meal*

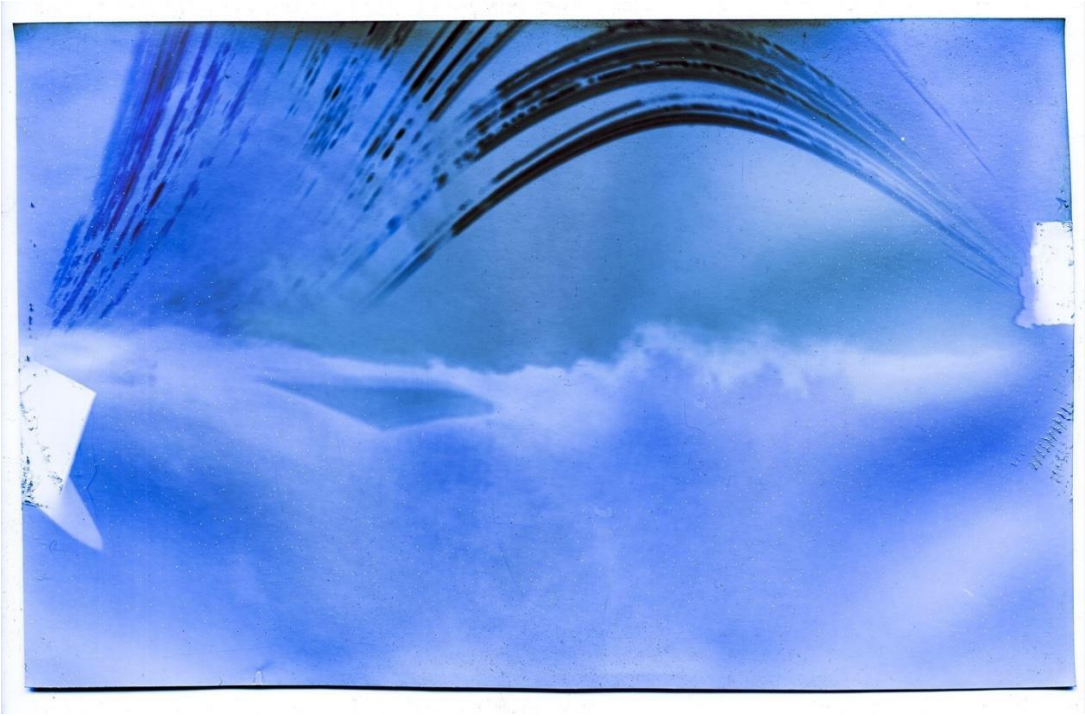
The existing infrastructure in the Agincourt Unit's research lab enabled the complicated machinations of the project to fall into place. For example, processing and storing multiple iohexol samples onsite before batched shipping to Johannesburg.

Research Lab: *Agincourt's lab facilitated the storage and shipment of the samples to Johannesburg.*

In closing, the blue solargraph image below was created by mounting a small pinhole camera on the roof of the study house at Wits Rural Facility, where I often stayed. This camera was exposed for the duration of the project. The arc, burnt by the sun onto the light-sensitive paper for each day of the study, created a permanent record of our work. It symbolises a personal journey, as I circled back to Bushbuckridge.

My early exposure was entirely focused on hospital-based individualist care. Now, with less hubris, more maturity and resilience, my eyes were opened to the previously unseen: school children waiting for their lift on the corner, where to find the best popcorn in the village, groundnut and spinach meals with the team, always going home with a boot full of avocados and some homemade mango atchar. As I transition to a broader population-health lens, I see that anchoring our work within rural communities in South Africa and on the continent – where voices are mostly silent – is fundamental to being reflexive and responsive to their needs, enables relevant context-specific research, and gives profound meaning to the science.

I am deeply grateful to everyone at the Agincourt HDSS site, and the ARK team for your wisdom, energy, time and commitment. And to the Bushbuckridge community for your trust and participation.



All ARK participants and staff gave written, informed consent to be photographed and for their images to be used in publications and research arising from the study.

Table of contents – Prevalence and Characterisation of Chronic Kidney Disease in a Rural Setting in South Africa

List of Abbreviations

ACR: albumin creatinine ratio (mg/mmol)

APOL1: apolipoprotein L1

ARK Consortium: African Research on Kidney Disease Consortium

ART: anti-retroviral therapy

AWI-Gen: Africa Wits-International Network for the Demographic Evaluation of Populations and their Health Partnership for Genomic Studies

BMI: body mass index

BP: blood pressure

BSA: body surface area

CAP: College of American Pathologists

CI: confidence interval

CKD: chronic kidney disease

CKD-EPI equation: Chronic Kidney Disease Epidemiology Collaboration equation

CKD ACR staging based urine ACR (mg/mmol):

A1: <3.0

A2: 3-30

A3: >30

CKD GFR staging based on eGFR (ml/min/1.73m^2):

G1: ≥90

G2: 60-89

G3a: 45-59

G3b: 30-44

G4: 15-29

G5: <15

CLS: Central Laboratory Services

DNA: deoxyribonucleic acid

DTPA: diethylenetriamine penta-acetic acid

DXA: dual-energy x-ray absorptiometry

EDTA: ethylenediamine tetra-acetic acid

eGFR: estimated glomerular filtration rate

ESKD: end stage kidney disease

FAS equation: Full Age Spectrum equation

GFR: glomerular filtration rate

GSK: GlaxoSmithKline

HDSS: Health and socio-Demographic Surveillance System

HIV: human immunodeficiency virus

HPLC: High Pressure Liquid Chromatography

IDMS: isotope dilution mass spectrometry

IQR: interquartile range

ISN: International Society of Nephrology

KDIGO: Kidney Disease: Improving Global Outcomes

KRT: kidney replacement therapy

LC-MS/MS: Liquid Chromatography- Mass Spectrometry

LMIC: low and middle income country

MDRD equation: Modification of Diet in Renal Disease equation

mGFR: measured glomerular filtration rate

MRC: Medical Research Council

NCD: non-communicable disease

NHANES: National Health and Nutrition Examination Survey

NHLS: South African National Health Laboratory Service

NICD: National Institute for Communicable Diseases

NKDEP: Laboratory Working Group of the National Kidney Disease Education Program

OR: odds ratio

POC: point-of-care

RMSE: root mean square error

SA: South Africa

SBIMB: Sydney Brenner Institute for Molecular Biosciences

SD: standard deviation

STEPS: stepwise approach to surveillance

SSA: Sub-Saharan Africa

TB: tuberculosis

UK: United Kingdom

USA: United States of America

WHO: World Health Organization

List of Figures

List of Tables

Background

In 2017, there were 1.2 million deaths from chronic kidney disease (CKD) reported worldwide, and 697 million cases of all-stage CKD were recorded – resulting in a global prevalence of 9%. Compared to 1990, the global all-age prevalence of CKD increased by 29% and most of the CKD associated disability was a consequence of diabetic nephropathy. Most of the burden of CKD was concentrated in the three lowest quintiles of socio-demographic Index (SDI). In southern Sub-Saharan Africa (SSA) the burden of CKD was much higher than expected, with poor outcomes are likely related to the absence of screening for early detection in those at risk, and limited access to resource-intense kidney replacement therapy (KRT) for those severely affected.(1)

In high-income settings, most CKD is caused by obesity-induced type 2 diabetes and secondary hypertension, and most associated deaths are cardiovascular in origin.(1). Follow-up from population-based longitudinal cohort studies revealed an inexorable decline to kidney failure was not the most common trajectory for those with CKD. Rather, those with less severe kidney disease died prematurely with relatively preserved kidney function, mostly from a cardiovascular event.(2) These findings resulted in the classification of CKD as an independent risk factor for premature ischaemic cardiac events. Initially, adverse outcomes were examined in relation to glomerular filtration rate (GFR) as the measure of kidney function but an additional association between albuminuria and adverse outcomes began to emerge.(3) Subsequently, similar cohort studies showed that persistent albuminuria was an independent risk factor for cardiovascular and all-cause mortality.(4) This led to changes in the Kidney Disease: Improving Global Outcomes Clinical Practice Guideline for the Evaluation and Management of CKD (KDIGO CKD guideline) that included albuminuria as a criterion for evaluating CKD.(5) Insights from such valuable data informed public health initiatives promoting screening, early detection and CKD management, and facilitated access to kidney replacement therapy (KRT) in the form of chronic dialysis or transplantation.(6)

In low and middle income (LMIC) settings, including South Africa (SA) and SSA, the true prevalence of CKD remains unknown because of the absence of large population-based studies, and prevalence rates from smaller studies in at-risk groups can inflate estimates.

Additionally, although associated risk factors for CKD are poorly characterised, there are many potential infectious and non-communicable risks for CKD and a high propensity for adverse outcomes, given the limited access to appropriate care. CKD risk is thought to be higher, in part, because of multiple events throughout the life course that potentially impact functioning nephron mass. Nephron mass can be adversely affected by maternal hypertension, diabetes and obesity, foetal intrauterine growth retardation or premature delivery, and childhood malnutrition with repeated diarrhoeal infections in early life.(7) Added to these factors are hereditable susceptibility and epigenetic interactions – a case in point being the association between inheriting high risk alleles for apolipoprotein 1 (*APOL1*) that increase risk for kidney disease, and the potent upregulation of *APOL1* induced by interferon.(8, 9) This is relevant in African populations where the mutations that emerged for *APOL1* confer protection against trypanosomiasis (one copy of a risk allele), and yet two copies of the risk allele confers risk for CKD. The persistent burden of infectious diseases and their role in upregulating interferon in those with high risk *APOL1* genotypes remains unknown in African populations. Context specific, less traditional factors associated with CKD, such as the contribution of endemic and other infections (malaria, schistosomiasis, TB, HIV); over the counter or herbal medicine consumption; and agricultural pesticide or heavy metal toxin exposure (lead) from water pollution, remain unknown. While well-characterised longitudinal population cohorts in SSA are emerging, that in time will address outcomes for those with CKD, current knowledge gaps impair implementation of appropriate screening, prevention and management strategies.(10) The consequences are dire, with most who develop severe kidney disease suffering from premature disability or death.(11)

(i) Measuring kidney function – measured GFR (mGFR)

Currently, measuring GFR is the most accurate method to assess kidney function. The ideal way to measure GFR is to administer an exogenous filtration marker that is safe, not protein-bound and freely filtered by the glomerulus; it is neither reabsorbed nor secreted by the kidney tubules, not metabolised or excreted by the gastrointestinal tract or hepatobiliary system, and easily measurable in collected urine or blood. The ideal “gold standard” marker is inulin, but it is expensive, not readily available, with no standard reference to calibrate

laboratory accuracy.(12) Also, measuring GFR using inulin requires constant intravenous infusion and urinary catheterisation for the duration of the procedure, which may be as long as eight hours, depending on kidney function. Alternative reference markers that are easier to use and more cost-effective are either radiolabelled with a nuclear isotope, such as tri-iodinated iothalamate (^{125}I-iothalamate), chromium-51 ethylenediamine tetra-acetic acid (^{51}Cr-EDTA) or technetium-99 diethylenetriamine-penta-acetic acid (^{99}Tc-DTPA); or non-radiolabelled, such as iohexol or iothalamate, both radiocontrast media. After administration of a bolus dose of the molecule (normally intravenously), the concentration in urine or blood is measured at intervals and plotted as a concentration (y-axis) versus time (x-axis) curve. Clearance is calculated using the area under the curve and incorporated into various mathematical models that allow for GFR calculation.(12) Most commonly, one compartment modelling is applied, based on the slow elimination phase of the tracer, using the Brochner Mortensen equation,(13) which corrects for the rapid elimination phase. This phase reflects distribution from the intravascular to extracellular space, immediately after administration of the tracer. After two hours, the rate of elimination from plasma is considered a reflection of glomerular function.(14)

Because exogenous reference markers differ from one another, regional use depends on preference and availability. Iothalamate, which was used in the United States to develop the Modification of Diet in Renal Disease (MDRD) and Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations, is secreted by renal tubules, lending to slight overestimation of GFR. In contrast, iohexol which was used in Scandinavia for the Full Age Spectrum (FAS) and Revised Lund-Malmö Study equations, is not secreted by renal tubules, has an excellent safety profile, is temperature stable, does not need to be processed in real time and is affordable.(15) The laboratory assays for iohexol are sensitive, allowing small doses to be administered without compromising accuracy. Also, it is the only filtration marker with an external quality assurance program to reduce interlaboratory bias (Equalis AB, Sweden).(16, 17) These various characteristics make iohexol appealing for studies of kidney function in resource-limited settings, allowing for pooled samples to be transported to centres of excellence for processing.(18)

Using alternative reference markers in preference to inulin is considered a pragmatic balance between acceptable levels of accuracy and feasibility in clinical and research

settings. However, to interpret measured GFR research studies it is important to understand intra-individual and analytic bias related to the marker, the performance of the marker compared to inulin, and the performance of markers relative to each another. Bias can be introduced with (i) the time-frame used for plasma sampling with relatively longer sampling intervals required for those with impaired kidney function; (ii) single versus multi-sample techniques, although multi-sample techniques are recommended for research studies; (iii) laboratory methods to measure the marker, for example high pressure liquid chromatography (HPLC) can yield different results to liquid chromatography-mass spectrometry methods (LC-MS/MS); and (iv) using non-standardised protocols.(12) It is relevant that while iohexol has become the preferred reference marker for measured GFR studies there are still no consensus guidelines which would serve to minimise methodological and analytic bias. By way of example, the British Nuclear Medicine Society Guidelines were developed for centres using radionuclides to measure GFR because an audit revealed large between-centre discrepancies.(14) The correlation between reference markers and urinary inulin clearance (as the gold standard) has been investigated: urinary clearance of iothalamate, urinary and plasma clearance of 51Cr-EDTA, and plasma clearance of iohexol are considered sufficiently accurate for measuring GFR; urinary clearance of DTPA and urinary clearance of iohexol have insufficient accuracy; while 24-hour creatinine clearance and plasma clearance of DTPA are the least accurate.(12, 15) Relating specifically to iohexol plasma clearance, studies report 86-98% correlation with urinary inulin excretion.(12) When comparing reference markers with one another, plasma clearances of 51Cr-EDTA and iohexol have excellent agreement. Plasma clearance of iothalamate consistently overestimates that of iohexol by ±10% (measured GFR is consistently higher) because of tubular secretion of iothalamate – this effect may be accentuated when using LC-MS/MS compared to HPLC. (12)

In summary, when compared to urinary clearance of inulin, the plasma clearance of iohexol is sufficiently accurate for measured GFR studies. When comparing commonly used reference markers, iohexol plasma clearance is highly concordant with 51Cr-EDTA plasma clearance but overestimated by iothalamate plasma clearance – the degree of which can be influenced by the method of laboratory measurement. Intra-individual variation associated

with iohexol plasma clearance studies ranges between 4 and 10%, potentially contributing additional biological bias.

(ii) Predicting kidney function – estimated GFR (eGFR)

A more pragmatic approach to measuring GFR would be to predict GFR using a validated estimating equation (eGFR). Estimating equations are derived by combining endogenous (rather than exogenous) biomarkers, such as creatinine or cystatin C, with variables known to influence their concentrations. Serum cystatin C is influenced by age, sex, smoking, obesity, inflammation, corticosteroid administration and thyroid disease, but is not secreted by tubules, or dependent on dietary factors and muscle mass.(19) On the contrary, individual factors influencing serum creatinine include age and sex (because of muscle mass), dietary protein ingestion and kidney function (tubular secretion of creatinine increases proportionately with declining kidney function). Aside from intraindividual biological variation, creatinine measures are subject to seasonal changes, particularly in hot climates.(20) Analytic bias also influences serum creatinine or cystatin C concentrations. In the case of creatinine, bias is inversely related to creatinine concentration: as creatinine concentrations decrease, the analytic bias increases. This has implications for population screening where most individuals are expected to have normal creatinine.(21)

To reduce analytic bias, standard reference methods and materials have been developed for creatinine and cystatin C, against which laboratories calibrate their assays. For creatinine, the enzymatic method is more accurate than the Jaffe method,(21) however, the relative expense of enzymatic assays has precluded widespread uptake in SSA. Since the Jaffe method is subject to interference from non-creatinine chromagens, resulting in over-reading true creatinine, laboratories correct for this effect using a compensated Jaffe method. The disadvantage of this might be artificially low creatinine levels, especially problematic in children. A recent study based in a local clinic in Johannesburg showed that delays in laboratory processing of ≥6 hours artificially increased creatinine concentrations using the Jaffe method, but not the enzymatic method.(22) It is reasonable to anticipate even longer delays in underserviced areas, such as those in rural SA.

All currently available eGFR equations have been developed and validated for use in high-income, predominantly Caucasian populations. In the USA, when modelling the MDRD equation, it was observed that for a given mGFR, participants with African ancestry had higher serum creatinine levels than their Caucasian counterparts.(23) On this basis, an ethnicity coefficient was developed to adjust eGFR by a factor of 1.18 (+18%) when calculating eGFR in Americans with African ancestry (the term "ethnicity coefficient" was derived by the authors of the study). The authors hypothesised that the higher creatinine levels they observed were due to greater muscle mass in those of African ancestry compared to other population groups, and cited three studies in support of this hypothesis.(23) However, none of the studies to which they referred were designed to test their hypothesis and have limited scientific validity.(24-26) A similar approach was taken when developing the CKD-EPI (creatinine) equation, resulting in another ethnicity coefficient, adjusting eGFR by 1.159(~+16%).(27)

To date, there is no scientific justification for racialised coefficients in either of the eGFR equations, and the reported higher serum creatinine for a given mGFR in those of African ancestry, remains unexplained.(28) A recent study using DXA scanning to assess muscle mass differences in patients with end stage kidney disease (ESKD) concluded differences in creatinine ascribed to African ancestry could not be explained by differences in muscle mass alone.(29) Other potential explanations, such as differences in renal tubular handling of creatinine, have not been definitively proven.(30) Use of racialised coefficients may inflate true GFR with potentially significant consequences. One of these is delaying or missing the diagnosis of early CKD, thus hindering opportunities for intervention or nephrology referral for investigation and treatment.(31)

In 2002 the National Kidney Foundation recommended the MDRD equation for estimating GFR in clinical practice.(32) Subsequently, in 2012, the KDIGO CKD guideline recommended the CKD-EPI (creatinine) equation.(5) For the nephrology community outside the USA, the question was how to interpret the significance of ethnicity coefficients in non-Caucasian populations. Though the guidelines advise that eGFR equations should be validated for specific populations, this may not be feasible, particularly in resource-limited settings. Studies from China,(33, 34) South Korea,(35) Thailand,(36) Taiwan,(37) and Japan,(38, 39) revealed discrepancies in the results, fuelling an ongoing debate in Asia about which

coefficients are accurate, or reflect methodological differences in the measurement of GFR. In China, for example, a coefficient of 1.233 was proposed for the MDRD equation,(33) contrasting with the coefficient of 0.763 from Japan,(39) translating to a potential difference in eGFR of almost 40%.(40)

Contrary to validation studies in Asia, those from West, East and Southern Africa consistently demonstrate ethnicity coefficients in African participants overestimate mGFR.(41-47) Consequently, in countries such as SA, the National Health Laboratory Service (NHLS), which provides services to approximately 80% of the population, omits ethnicity coefficients when estimating GFR; however, many SSA countries continue to use these.

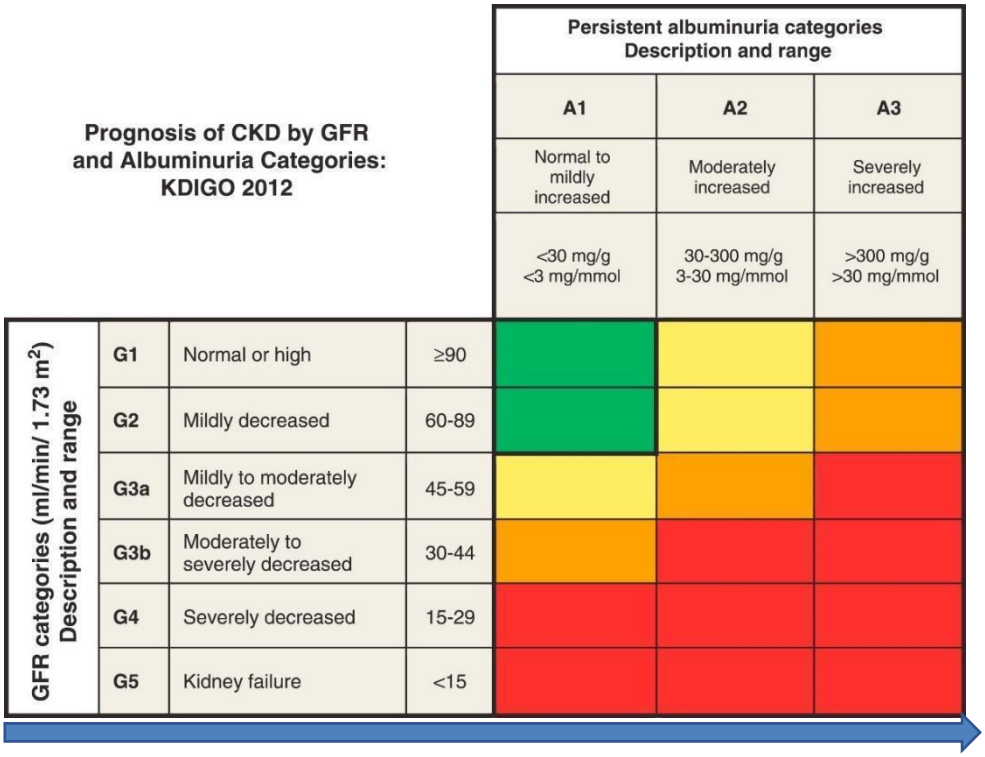
(iii) Defining chronic kidney disease (CKD)

Globally, most nephrologists accept the KDIGO Clinical Practice Guideline for the Evaluation and Management of CKD as their reference for clinical care.(5) The diagnosis of CKD requires either (i) the presence of eGFR<60ml/min/1.73m^2 using the CKD-EPI (creatinine) equation, or (ii) markers of kidney damage, including any of the following: albuminuria (preferably measured using an ACR≥3.0mg/mmol), abnormal urinary sediment, tubular disorders, evidence of kidney damage on imaging or histology, or a history of kidney transplant (of all these, albuminuria is the most common criterion). The key to confirming chronicity is demonstrating that these abnormalities persist, with a minimum period of three months before repeat testing. If CKD is confirmed, further classification of CKD by stage of GFR and stage of ACR allows for stratification into risk for adverse events (Figure 1). However, the risk stratification and associated threshold values for GFR and ACR stages were derived from large general population cohorts in high-income countries, without any representation of African populations.(4, 48) Many high-income populations are older than those in SSA and have potentially different risk profiles, and so applying the same thresholds for risk is likely inappropriate. Some have advocated that the definition for low GFR be adjusted to accommodate the young and elderly, as current thresholds under- and overdiagnose CKD respectively, in these groups. More so, the current GFR staging does not accommodate senescent decline kidney function.(49) Albuminuria may also need to be stratified by GFR stage, with the lower threshold for ACR increased.(3, 49) Despite these challenges and

controversies, the KDIGO CKD guideline has created a framework for clinicians, with standardised points of reference, and have highlighted critical gaps for future research.

To confirm CKD in clinical practice, it is feasible to repeat individual screening. On the other hand, this is much harder to achieve at population level, with large cross-sectional epidemiological studies usually only performing a single screen. Inconsistencies among studies, with some only using estimated GFR (calculated using different equations) and others assessing albuminuria using different methods, have been a key challenge to understanding the true global burden of CKD.(50, 51)

Figure 1: KDIGO CKD classification and prognosis



Blue: increasing risk associated with increasing severity of albuminuria (vertical); or increasing risk associated with worsening decline in GFR (horizontal); Green: low risk (if no other markers of kidney disease, no CKD); Yellow: moderately increased risk; Orange: high risk; Red, very high risk. Adapted from the KDIGO CKD guideline(5)

(iv) Prevalent CKD and associated risk in South Africa (SA) and Sub-Saharan Africa (SSA)

Although the true prevalence of CKD in SA and SSA is unknown, it is predicted to be high due to risks associated with prevalent infectious and non-communicable diseases (NCD).(10) A

recent systematic review reported higher CKD prevalence in SSA compared to North Africa, with the highest prevalence in West Africa.(52) Studies either represented high-risk groups (usually hospital-based) or were population-based, making comparisons difficult. Also, CKD prevalence varied substantially depending on which eGFR equation was used and whether chronicity was confirmed. In high-risk groups, the most common risk factors associated with CKD were HIV infection, hypertension and diabetes. Reported prevalence may be influenced by survivor bias, since most with severe kidney disease demise in the absence of appropriate care.

In SA, isolated non-population-based studies have identified HIV infection, hypertension and diabetes as the dominant risk factors for CKD.(53, 54) In those presenting with ESKD, the South African Renal Registry reported the cause as unknown in one third of cases on dialysis, followed by hypertension-associated kidney disease, diabetes mellitus and primary glomerular disease, in decreasing frequency.(55) To date, the only population-based study to evaluate kidney dysfunction in SA included a mixed-ancestry cohort from the Western Cape in which age, sex and hypertension were identified as significant risk factors.(56) Subsequently, *APOL1* high-risk genotypes were independently associated with systolic hypertension and reduced GFR in the same cohort, although frequencies of the high-risk genotypes were lower (1.01%) than those observed in the US-resident African diaspora (10-15%).(57) A strong association between high-risk *APOL1* genotypes and HIV-associated nephropathy has also been reported from Johannesburg.(58)

Although “conventional” risk factors for CKD, namely HIV infection, hypertension and diabetes have frequently been identified, the cause of kidney failure is unknown in one third of South Africans. A recent study from rural East Africa demonstrated only one third of participants had conventional risk factors associated with CKD.(59) The study also highlighted the presence of unexplained leukocyte-esterase negative leukocyturia and haematuria as potential indicators in those with undiagnosed causes of CKD. In a recent population-based study from Malawi, there were no conventional risk factors in 20% of participants with low GFR, and risk for kidney disease was associated with living in a rural rather than urban area, and working in agriculture or fishing.(60) This data suggests that additional causes of CKD in SSA might be tubular, interstitial or glomerular in origin, and environmental or infectious exposures remain unidentified.

Study justification and objectives

From a clinical perspective, managing kidney disease requires functional primary health care and referral structures, with flexibility for multilevel and horizontal management approaches.(61) The South African health system is relatively inflexible, with a vertical, siloed approach. In the absence of point-of-care (POC) diagnostics, screening for kidney disease is cumbersome, requiring at least one urine protein/albumin test and a serum sample for creatinine (eGFR), and retention in care for follow-up. POC assays could potentially improve screening but there is a misperception that this approach is less accurate than existing pathology laboratory testing,(62, 63) and its relatively high cost may be prohibitive. Often, primary kidney disease requires a kidney biopsy for definitive diagnosis, while secondary kidney disease occurs consequent to systemic illnesses which often coexist (such as HIV and diabetes), all of which need patient-centred management and appropriately trained health care workers. In the absence of screening, most present with late-stage disease, at high risk of severe morbidity or death. The cardiovascular risk associated with kidney disease remains under appreciated and is often obscured by comorbidities.(64)

Unfortunately, the decision by the World Health Organization Independent High-Level Commission on NCD to prioritise cardiovascular disease and stroke, chronic respiratory disease, cancer and diabetes, later adding mental health, has also contributed to a lack of urgency among policy makers to prioritise kidney health.(65) Treating late-stage kidney disease is costly for health systems which leads health policy makers to conclude that KRT is too expensive to treat in resource-limited settings. Consequently, these costs are passed on to vulnerable households, with an affected member with late-stage kidney disease often resulting in catastrophic out-of-pocket expenditure.

Therefore, public health policy on the management of kidney disease in SA needs to be prioritised. To achieve this, further data on the accurate measurement of kidney function and consequently, the accurate prediction of the burden of kidney disease is necessary to ensure improved care for those at risk for, or who already have, kidney disease. Currently, there are no population-based CKD studies from SA located within rural communities. This omission is particularly relevant given many rural communities are at various stages of epidemiological transition and are thus vulnerable to an emerging burden of

cardiometabolic diseases, combined with existing infectious, environmental and socio-demographic stressors, all of which contribute to potential CKD risk. The paucity of data from rural communities in SA reflects the more pervasive absence of accurate data from the global South regarding CKD prevalence and associated risk that is context-specific – including infectious and non-communicable causes, compared to well characterised CKD prevalence and associated risk profiles in the North.

Therefore, the specific objectives of this doctoral work were to:

(i) Determine the population prevalence of CKD in rural South Africans;

(ii) Identify risk factors associated with CKD in rural South Africans;

(iii) Measure kidney function (mGFR) using intravenous iohexol in rural South Africans;

(iv) Evaluate performance of various eGFR equations compared to iohexol mGFR by country, and overall;

(v) Develop a regional GFR estimating equation that better predicts iohexol mGFR than existing GFR estimating equations; and

(vi) Accurately predict the prevalence of impaired kidney function in large population datasets from West, East and Southern Africa, with the improved GFR estimating equation.

For the first two objectives, similar studies were independently conducted in Malawi and Uganda, as part of the ARK Consortium. For the third objective of the study, study protocols were prospectively harmonised across the three ARK partner sites (Malawi, SA and Uganda). For the last three objectives, data from all three partner countries were pooled for a combined analysis.

Regarding the fulfilment of requirements for objectives three to five, many countries in Sub-Saharan Africa (SSA) do not have the appropriate facilities or infrastructure for measured GFR studies (including Uganda and Malawi), nor are there well characterised kidney disease cohorts from which to source appropriate datasets to validate GFR estimating equations. As

such, our methods for conducting the measured GFR studies in our study needed to be different from high-income settings. We debated how to do this within the ARK Consortium and agreed the most scientifically robust approach would be to preserve the pooled dataset for the development of the ARK equation, and source a diverse external validation dataset in South Africa (where we were able to access an external validation dataset). By preserving the pooled dataset, we would allow for between-country comparisons which we considered a critical component of our analysis, given the population diversity that characterises SSA. While it is accepted that an internal validation will perform better as it is derived from the same dataset in which the equation was developed, the performance of a new GFR estimating equation is best tested in a (different) external validation dataset.

By prior agreement in the ARK Consortium, data from all three sites were pooled for analysis and the iohexol measured GFR modelling of the ARK equation was led by the chief statistician in Uganda, as the statistical knowledge required was beyond the scope of this PhD.

To predict the regional prevalence of impaired kidney function, population datasets were derived from the ARK Consortium,(60, 66) and the Africa Wits-International Network for the Demographic Evaluation of Populations and their Health Partnership for Genomic Studies (AWI-Gen) in Ghana, Kenya, Burkina Faso and SA.(67, 68) The term “impaired kidney function” is used as the population studies were single-screen studies that did not repeat measures to confirm CKD.

Methods

1. Determining CKD prevalence and associated risk in Rural South Africans

This component of the study encompasses objectives (i) and (ii).

Study setting and sampling strategy

This cross-sectional study was conducted in the MRC/Wits-Agincourt Unit, Bushbuckridge, a rural subdistrict of the Mpumalanga Province, in the north-eastern part of SA. The Agincourt Health and socio-Demographic Surveillance System (HDSS) site (90 000 people) was established in 1992. Annual updates of vital information on residents have facilitated a longitudinal platform supporting observational and interventional work along the life course involving individual and population health, migration and urbanisation.(69)

A minimum sample size of 1 800 was required to provide at least 80% power to determine CKD prevalence >5%, provided the true prevalence was ≥6.5%. To ensure the study sample was population-representative for adults aged 20 to 80 years, proportional allocation of participants by sex and age-strata was performed based on the Agincourt HDSS population census round conducted in 2014. The sample size was increased proportionately to 2 759 individuals to accommodate a predicted non-participation rate of 25%.

Participant recruitment and study procedures

Ethics approval was obtained from the Faculty of Health Sciences Medical Human Research Ethics Committee, University of the Witwatersrand (certificate number M170583). From November 2017 to September 2018, a team of locally trained fieldworkers and nurses performed home visits in 31 research villages in the Agincourt HDSS.(70) For each participant recruited to the study, fieldworkers obtained written, informed consent in the participant's first language (primarily Xitsonga), administered a CKD risk questionnaire, performed anthropometry (including height, weight, hip and waist circumference measurements) and measured blood pressure.

For all anthropometric measures, participants were asked to remove heavy clothing and footwear. Weight was measured to the nearest 0.1kg using a digital flat scale (Seca, Germany) and weighing scales were calibrated weekly. Standing height was measured to the

nearest 0.1cm with a portable stadiometer (Seca, Germany). Height and weight were used to calculate body mass index (BMI) as weight(kg)/height(m)^2. Waist and hip circumference were measured to the nearest 0.1cm with a circumferential tape measure (Seca, Germany). Waist circumference was measured at the midpoint between the lowest palpable rib and the iliac crest in the mid-axillary plane. Hip circumference was measured at the most protruding part of the buttocks with the participant standing. Waist and hip circumferences were measured three times, with the average of each metric used for analyses.

Blood pressure (BP) was measured using automated upper arm devices (Omron M6W, Intellisense BP785 large cuff, Japan). Participants were asked to sit comfortably with legs uncrossed for five minutes before readings were taken. Three measurements were taken using an appropriate-sized cuff on the left arm, at two-minute intervals. Of three BP readings, the first was discarded and an average of the second and third readings used for analyses.

Nurses performed capillary POC random cholesterol and random glucose testing (Cardiochek® PA analyzer, PTS Panels® test strips, PTS Diagnostics, USA), and POC haemoglobin testing (HemoCue® Hb 201+ analyzer and test strips, Norway). Their respective suppliers calibrated all POC and blood pressure monitoring devices before commencing the study.

If a participant knew their HIV status as positive, this was recorded, and they were asked about adherence to antiretroviral therapy (ART).(71) If participants did not know their HIV status or had previously tested negative, nurses offered voluntary POC screening and testing (Alere™ HIV Combo, Abbott, USA), according to South African Department of Health guidelines.(72) A positive test result was confirmed with a second test (Uni-Gold™ Recombigen® HIV-1/2, Trinity BioTech, USA). Nurses collected blood samples and a freshly voided urine sample for laboratory testing. A spot urine pregnancy test was performed for premenopausal women (Abon™ One Step Pregnancy Test, Pharmaland, UAE). All samples collected in the field were stored in isothermally padded bags, prepacked with ice-bricks, to maintain temperatures between 2 and 6°C and temperatures in each bag were monitored (Easylog, Lascar Electronics, UK). Each day, after completion of fieldwork, the team delivered

all samples to the Agincourt Research Laboratory for onsite processing and storage at -80°C, according to standardised protocols.

Laboratory procedures

For DNA, buffy coats from centrifuged EDTA specimens were lifted and stored in the Agincourt Research Laboratory and transported to the Sydney Brenner Institute for Molecular Biosciences (SBIMB), Johannesburg, for DNA extraction and biobanking. A 20µl aliquot of DNA was transported to the Frederick National Laboratory at the National Cancer Institute, USA, for *APOL1* genotyping as described in David et al.(73) DNA was genotyped using TaqMan assays (ThermoFisher Scientific, USA). *APOL1* G1 allele comprises two missense variants, rs73885319 (G1g) and rs60910145 (G1m); however, the presence of only the G1g variant is sufficient to define the G1 risk allele. The *APOL1* G2 allele consists of a six bp in-frame deletion, rs717185313. High risk genotypes were defined as those containing two high risk alleles (G1/G1, G1/G2 or G2/G2); low risk genotypes were defined as those containing zero or one risk allele. The number of *APOL1* risk alleles (either G1 or G2 allele) carried by each subject was then coded as 0 for the G0/G0 genotype, 1 for the G0/G1 or G0/G2 genotype, or 2 for the G1/G1, G1/G2 or G2G2 genotypes. The *APOL1* genotypes were further coded as "high renal risk (HR) genotype" (if the subject carried any combination of 2 risk alleles) or "low renal risk (LR) genotype" (if the subject has 0 or 1 risk alleles).(73)

For serum and urine testing, all specimens were transported at -80°C to the Central Laboratory Services (CLS) in Johannesburg. Serum creatinine was measured by an IDMS-traceable modified Jaffe method (Cobas© 6000/c501, Roche Diagnostics.GmbH, Mannheim). Uric acid was measured by enzymatic colorimetric testing on a Roche/Hitachi Cobas© c analyzer and hepatitis B surface antigen titers, defined as positive or negative based on a prespecified laboratory cut-off value (ARCHITECT, Abbott, Chicago, USA). Urine albumin was measured using a colorimetric method on the Cobas© 6000/c501 analyzer; urine creatinine by the modified Jaffe method on the same analyzer. The CLS laboratory adhered to standard daily internal quality control procedures and complied with the requirements of the external quality control programme through the College of American Pathologists (CAP), with approved certification for urine and serum chemistry for the study period.

Study definitions

The CKD risk factor questionnaire was designed using the WHO Stepwise approach to surveillance (STEPS) as a framework.(74) For individual participants, the highest level of education and a household assets score were received from the Agincourt HDSS.(75) Increased waist circumference was defined as >=80cm for women and >=94cm for men. Increased waist to hip ratio was defined as >0.85 for women and >1.0 for men.(76) BMI (kg/m^2) was used to classify: underweight (<18.5); normal weight (18.5-24.99); overweight (25.0-29.99); obese (≥30.0): obese class I (30.0-34.99); obese class II (35.0-39.99); obese class III (>=40.00).(77) Participants were classified according to the 7th Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure as: normotensive (systolic SBP<120mmHg and DBP<80mmHg); prehypertensive (SBP≥120mmHg and <140mmHg or DBP≥80mmHg and <90mmHg); or hypertensive (SBP≥140mmHg or DBP≥90mmHg).(78) Diabetes was defined as a non-fasting glucose >=11.1mmol/L,(79) hypercholesterolaemia as a non-fasting total cholesterol >5.0mmol/L,(80) hyperuricaemia as uric acid >0.43mmol/L in men and >0.36mmol/L in women,(81) and anaemia as haemoglobin (Hb) levels <11.0g/dL in pregnant women; <12.0 g/dL in non-pregnant women and <13.0 g/dL in men.(82) For apolipoprotein L1 (*APOL1*) genotyping, analyses were performed using high versus low renal risk stratification, where individuals with zero or one risk allele were defined as low renal risk, and individuals with two risk alleles defined as high renal risk.(73)

Characterising kidney disease

As per the KDIGO CKD guideline,(5) CKD was defined using estimated glomerular filtration rate (eGFR<60ml/min/1.73m^2), calculated using the CKD-EPI (creatinine) equation without ethnicity coefficient,(27) and albuminuria (albumin:creatinine ratio (ACR) >=3.0ml/mmol). Based on our initial screen, all participants with an eGFR<60ml/min/1.73m^2 or urine ACR>=3.0mg/mmol were revisited by the study team for repeat sampling, where possible, after a minimum period of 3 months, to fulfil the temporal requirements for the definition of CKD. Once confirmed, CKD was further classified(5) by eGFR stage (ml/min/1.73m^2), (G1: eGFR >=90; G2: eGFR 60-89; G3a: eGFR 45-59; G3b: eGFR 30-44; G4: eGFR 15-29; G5: eGFR <15), and ACR stage (mg/mmol) (A1: <3; A2: 3-30; A3: >30) to stratify severity and risk for

adverse outcomes. While this classification may not be appropriate in South African and SSA populations, it was used in the absence of an alternative. All data was age-standardised using the revised WHO World Standard Population Distribution.(83) Results from the first screen were also presented to facilitate comparison with other epidemiological studies that only have a single screening result – which is by far the majority.

Data analysis

Continuous variables were represented as mean (± standard deviation) if normally distributed, and median (interquartile range) if non-normally distributed. Categorical variables were expressed as frequencies (percentage). To identify factors independently associated with CKD, it was defined as a dichotomous variable which included those in whom albuminuria (ACR>=3mg/mmol) and/or eGFR<60ml/min/1.73m^2 were confirmed with follow-up. Logistic regression was used to estimate odds ratios (OR), with corresponding 95% confidence intervals (95% Cis) in a parsimonious, forward stepwise approach to develop a multivariable model adjusting for age, sex, phenotypic, genotypic (*APOL1* high vs. low renal risk), infectious and metabolic comorbidities. Using the same approach, we determined whether the same risk factor profile was associated with markers of impaired kidney function at first screening (as opposed to confirmed at follow-up as CKD), using ACR>=3.0mg/mmol and eGFR<60ml/min/1.73m^2 as dichotomous variables. We also performed bivariate analyses using ACR and eGFR, as continuous and dichotomous variables for the screening and follow-up visits, to determine any association with *APOL1* high versus low renal risk genotypes. Missing data was reported as missing and not imputed. Statistical analyses were performed using STATA 15 SE (Stata Corp, Texas, USA).

Participant referral for care

Participants with abnormal tests detected by the study team were referred to a local primary health care clinic for confirmatory testing and further management. Referrals were facilitated through an active, ongoing engagement programme with the communities and the Agincourt HDSS, health care workers in local clinics, regional hospitals and the district Department of Health.(84)

2. Measuring kidney function using intravenous iohexol in a sample of rural South Africans

This component of the methods section pertains to the measurement of GFR using iohexol in the MRC/Wits-Agincourt Unit, Bushbuckridge, SA (objective (iii)).

Study setting and sampling strategy

The study was conducted in the MRC/Wits-Agincourt Unit in Bushbuckridge, a rural subdistrict of the Mpumalanga province, in the north-eastern part of SA. From the population-based CKD prevalence study, a sub-sample of potential adult participants was identified for the iohexol mGFR study. The estimated sample size (N=730) was calculated for 90% power to detect whether eGFR was within 5% of mGFR (difference between eGFR and mGFR) at an eGFR of 60ml/min/BSA, assuming a standard deviation of 25ml/min and two-sided α=0.05. Sample size was stratified by sex and eGFR stage to overrepresent lower levels of kidney function and increased to 1 000 to accommodate a predicted 25% non-participation rate, and technical and screening failures.(85)

Study procedures

Potential participants, 18 years and older, were screened for eligibility between November 2018 and July 2019. Screening included an assessment for contraindications to intravenous iohexol administration, namely pregnancy or breastfeeding in women, uncontrolled epilepsy, severe, uncontrolled hypertension, or any acute pyrexial illness. If eligible, a trained fieldworker obtained written, informed consent in the participant's first language. On the day of iohexol measurement participants were asked to fast and were transported to the Agincourt Research Laboratory for the measured GFR procedure. In the clinic, a repeat screen was performed measuring blood pressure and temperature. In addition, height and weight were recorded, and trained nurses established two intravenous (IV) access points in the antecubital fossa (in most cases) of the dominant and contralateral arm. From the dominant arm, a baseline serum sample for creatinine and cystatin C was drawn and the cannula flushed with 10ml normal saline. A single bolus of 5ml Omnipaque^{TM} (350mg iodine/ml) was administered intravenously over 30 seconds, followed by a 10ml normal saline flush and the cannula was then removed. Time zero (T0) was the start time of

administration of the intravenous bolus of iohexol. In the contralateral arm, at 5, 120, 180, and 240 minutes from T0, 1ml of aspirated venous blood was discarded, followed by the collection of 4mls of venous blood in heparinized tubes. The exact time of sample collection was recorded for each sampling time. Using a calibrated laboratory scale, each syringe was weighed pre- and post-administration of iohexol in grams, to two decimal points. These weights were used to calculate the dose of iohexol administered.

To calculate GFR adjusted for body surface area (BSA)(mL/min/1.73m^2) from plasma clearance of iohexol, the slope-intercept method for three time points was used (using the exact time of measurement of the samples intended to be taken at 120, 180 and 240 minutes) in the second (slow) exponential phase of iohexol elimination.(14) The following steps were followed to calculate a measured GFR in accordance with the British Nuclear Medicine Society Guidelines (14):

1. The extrapolated y-intercept was the calculated concentration of iohexol at time zero, and the coefficient (k) was the slope of the semilog plot of plasma iohexol concentration against time using the formula: [$P(t)=P_0 \exp(-kt)$].

2. Adjustment for BSA was performed using the Haycock formula:
[$BSA = 0.024265 * Wt^{0.5378} * Ht^{0.3964}$].

3. The Brochner-Mortensen (BM) correction for adults was applied to account for iohexol plasma clearance in the first (rapid) exponential phase using the formula:
[$BM = (0.9908 * GFR) - (0.001218 * GFR^2)$].(13, 86)

Laboratory methods and testing

Iohexol

Iohexol plasma samples were processed in the Agincourt laboratory on the same day of collection and stored at -80°C, batched and transported to the NHLS in Johannesburg. The NHLS is accredited to the ISO 15189 standard and participates in the Equalis External Quality Assurance Programme for iohexol (Uppsala, Sweden).(16) Using a published method, all iohexol samples were analysed using ultra high performance liquid chromatography-tandem mass spectrometry ((SCIEX (Redwood, CA, USA) 5500 QTRAP (LC-MS/MS)).(87) Certified reference materials for iohexol (CRM: USP H0J211) and ioversol (CRM: USP 34510F) were

purchased from Industrial Analytical (Kyalami, SA) for the calibration curve and internal standard, respectively. Both compounds were lyophilised and weighed to make a stock standard of 10g/L in deionised water for iohexol, and methanol for ioversol. Working standards were then prepared with further dilutions to create a three-point calibration curve of 50mg/L, 500mg/L and 1 500mg/L for iohexol, by spiking the stock standard into drug-free serum. The working solution for ioversol was prepared by diluting the stock standard in methanol to a final concentration of 20mg/L. Internal quality control (IQC) samples were prepared by spiking the certified reference material for iohexol (standard) to create final concentrations of 100 and 1 000mg/L for respective low and high IQC. Coefficients of variation for IQC with the iohexol standard at 100 and 1000mg/L were 4.1% and 4.2%, respectively. Equalis samples were included in every run as an additional quality check. Iohexol samples were run on a C18 column purchased from SCIEX (Redwood, CA, USA). Mass spectrometry was carried out using electrospray ionisation in positive mode (ESI+) and multiple reaction monitoring was used for identification of iohexol and ioversol with transitions of 821.7>803.7m/z and 807.9>589m/z, respectively. Study sample iohexol concentrations (mg/L) were within the calculated z-score of <2.0.

Creatinine and cystatin C

The CLS standardised their creatinine assay using an IDMS assay traceable to a standard reference material for creatinine and used the compensated Jaffe method. Cystatin C samples were processed by the NHLS in Johannesburg using the Tina-quant® Cystatin C Gen.2 test kits and all samples were analysed as a single batch after completion of the study to reduce inter-assay variability. The assay is an immunoturbidimetric assay standardised to the ERM-DA471/IFCC reference material (Roche Diagnostics, USA).

Quality control procedures

Daily IQC was performed as per standard laboratory practice and inter-assay coefficients of variation for each laboratory biomarker were calculated based on analyses of commercial controls. For external quality assurance, standard serum was distributed to the participating laboratory for testing of common analytes. For creatinine, the laboratory in SA complied with the requirements of the CAP Quality Assurance Program. For cystatin C, the laboratory

complied with the requirements of the Equalis External Quality Assurance Program (Uppsala, Sweden).

Recalibration and split sample testing study for creatinine and cystatin C

Systematic bias may arise between study site laboratories because of different pre-analytic and analytic methods for biomarker measurement. As far as possible, bias can be eliminated in the pre-analytic stages through standardising study protocols prior to initiation of the study. Bias during the analytic stage can be controlled by a recalibration study, where measurements from one study site are recalibrated to the measurements in a reference study site. Recalibration allows for correction of inter-site laboratory differences in time or space, which relate to the types of assay, the manufacturer, and the analytic platform.(88) Despite the inherent quality assurance of manufacturer assays and the availability of standardised reference materials to aid calibration, evidence suggests some assays do not meet optimal bias limits and calibration differences persist, which explains the need for rigorous internal and external quality control procedures in each study laboratory. All these potential sources of analytic bias may impact research data. For epidemiological studies involving population-level data, these small systematic differences may result in a shift of the distribution of a biomarker, potentially biasing estimates of mean values and the prevalence of dichotomously defined variables; in this case, the presence or absence of kidney disease.(88) In this recalibration and split sample testing study, inter-laboratory bias was assessed for serum creatinine and cystatin C biomarkers measured in the ARK study site partner laboratories. If significant systematic differences were observed, the recalibration equation to correct for the analytic bias was determined. To assess inter-site creatinine and cystatin C laboratory measurement variability, randomly selected split serum samples from Uganda (N=50) and Malawi (N=20) were transported to reference laboratories in SA for repeat testing.

Recalibration of creatinine

For Uganda and Malawi respectively, split sample results were compared with the reference laboratory in SA using scatter and differential plots. Outliers were flagged for review, defined as creatinine values more than three standard deviations from the mean difference between paired values.(88) A single outlier for creatinine measurement from Uganda was excluded

after confirming it was transcriptional error during data entry. Agreement analysis was performed using Bland-Altman plots and Lin's concordance correlation coefficient.

Recalibration of cystatin C

A similar approach to that for creatinine was used for recalibration of cystatin C. Cystatin C measurements from Malawi and Uganda were performed in the Uganda laboratory, so for this component of the study, split samples from Uganda were compared to repeat samples in the SA reference laboratory. Scatter and differential plots were examined, no outliers were identified, and the agreement analysis proceeded. A cumulative sum (Cusum) test was performed to assess the linearity and Passing-Bablok regression analysis was used to determine the calibration function for the relationship between the paired cystatin C values, assuming the regression equation Y=A(intercept)+B(slope)*X. Statistical calculations were performed in Stata/SE, version 16.1.

Inter-site calibration for serum creatinine

Correlation was good between original Malawi and Uganda samples and repeat assays in the reference laboratory in SA (Table 1, Figures 2 and 3). With SA as reference, the bias was -2.70µmol/L (95% CI -6.70 – 1.59) and +9.29µmol/L (95% CI 7.25 – 11.33 µmol/L) for Malawi and Uganda creatinine samples, respectively. Since laboratories in Malawi and SA used the modified Jaffe method and Uganda used the enzymatic method, the systematic bias was ascribed to these methodological differences across the study sites. Ideally, KDIGO recommends the enzymatic method in preference to the Jaffe method, as the former is less biased and less prone to interference.(5) On this basis, creatinine values for Malawi and SA were recalibrated by a constant factor of +9.29µmol/L, to align with the Uganda laboratory.

Inter-site calibration for serum cystatin C

When comparing Ugandan cystatin C values to the repeat values obtained in SA, the correlation was good, with a bias of 0.10mg/dL (95% CI 0.36 – 0.17). With SA as the reference laboratory for the split sample, Passing-Bablok regression analysis was performed to determine the constants for the recalibration equation. (Table 1, Figures 4 and 5). The resulting regression coefficients (slope and intercept) were used to recalibrate the Uganda

and Malawi cystatin C measurements for comparability to the SA reference laboratory values.

Table 1: Agreement and recalibration coefficients for creatinine and cystatin C

Study Site	Analyte	No of excluded outliers	R^2 Lin's Concordance Correlation Coefficient	Intercept	Slope
Malawi	creatinine(µmol/L)	0	0.977	n/a	n/a
Uganda	creatinine(µmol/L)	1	0.967	n/a	n/a
Uganda	cystatin C (mg/dl)	0	0.942	0.178 (95% CI 0.090 – 0.349)	0.922 (95% CI 0.776 – 1)

Figure 2: Bland-Altman plot comparing agreement between serum creatinine measurements for Malawi (MW) and South Africa (SA)*

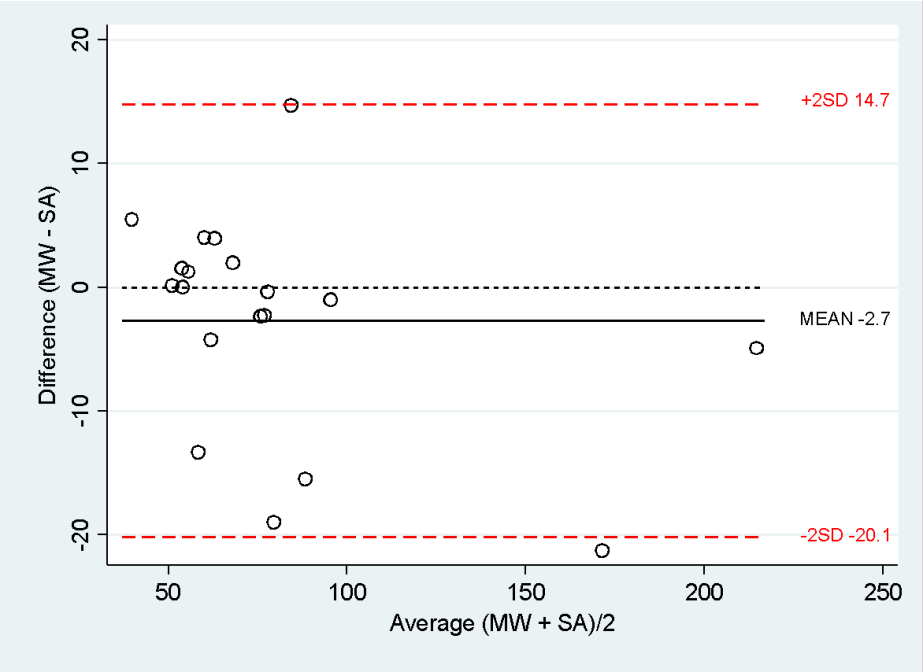
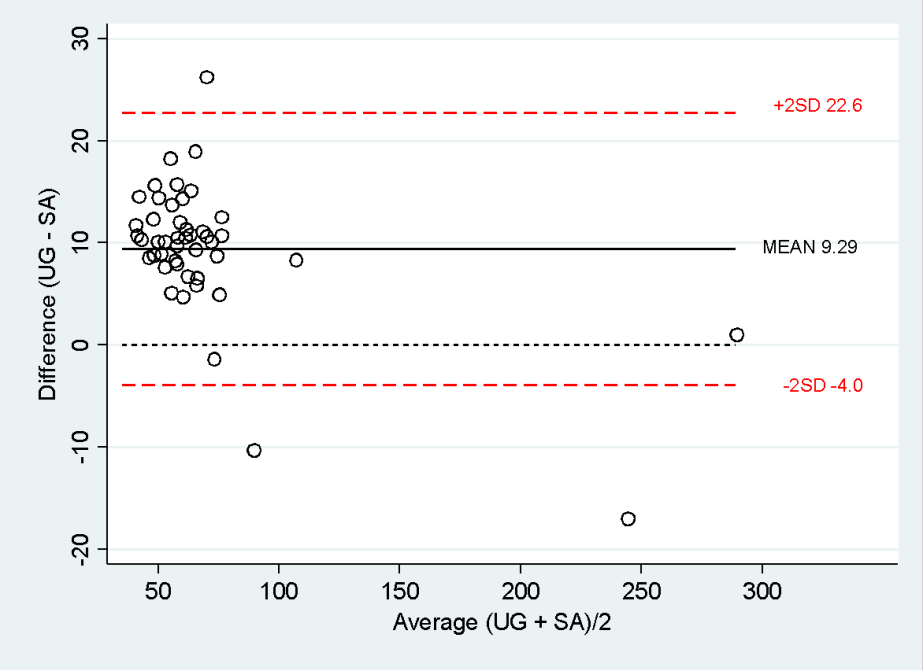
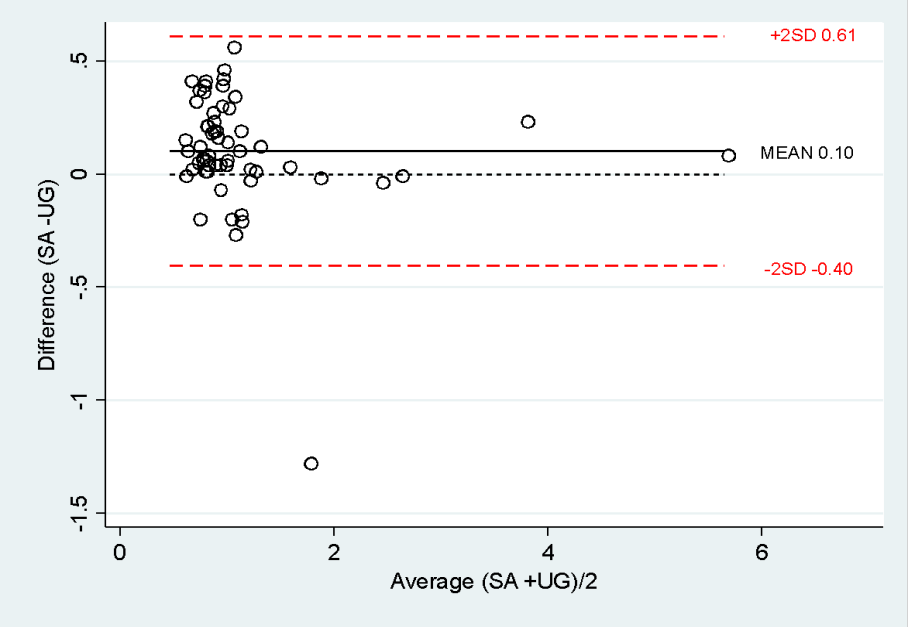


Figure 3: Bland-Altman plot comparing agreement between serum creatinine measurements for Uganda (UG) and South Africa (SA)*



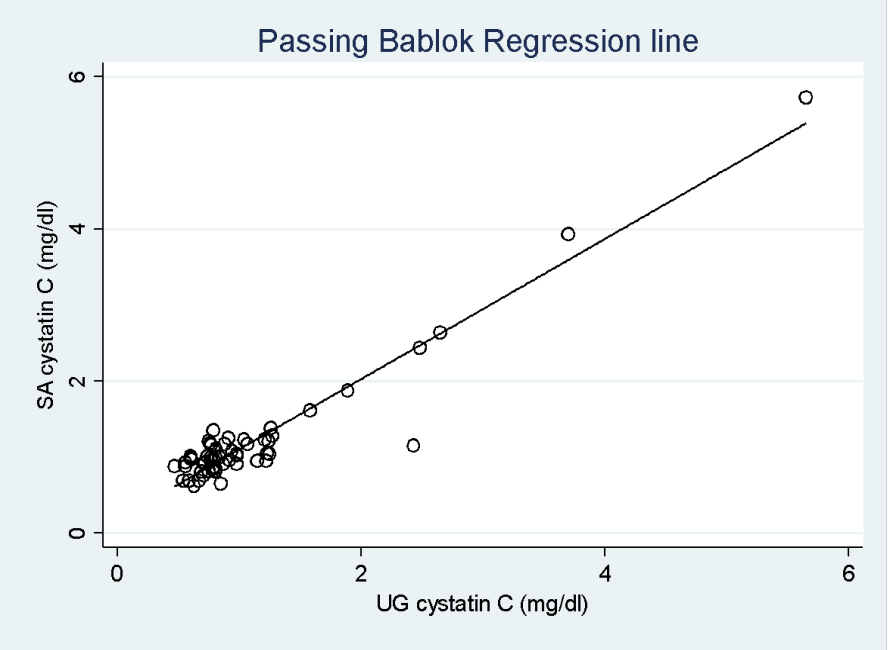
*Dashed red lines represent the 95% confidence intervals for creatinine in the calibration samples.

Figure 4: Bland-Altman plot comparing agreement between serum cystatin C measurements in Uganda (UG) and South Africa (SA)*



*Dashed red lines represent the 95% confidence intervals for cystatin C in the calibration samples.

Figure 5: Passing-Bablok regression line for recalibration of cystatin C in South Africa (SA) and Uganda (UG)



GFR estimating equations

For all eGFR equations, standard conventional units (mg/dL) for serum creatinine values were rounded to the nearest 100th of a whole number. Serum creatinine values expressed as standard international (SI) units (µmol/L) were rounded to the nearest whole number. The formula used to convert serum creatinine from conventional to SI units = [serum creatinine (conventional units) x 88.4]. The units for serum cystatin C are mg/L, and all equations were adjusted for BSA with units for GFR as mL/min/1.73m^2. The performance of the following serum creatinine and serum cystatin C-based GFR estimating equations was evaluated:

Cockroft Gault equation adjusted for BSA

GFR = [140-age (years) x weight (kg) x (0.85 if female) x 1.73m^2)] / [sCr x BSA$^{(86)}$ (m^2)]

In its original form, the Cockroft Gault equation did not adjust for body surface area (BSA).[89] However, this adjustment is required when comparing performance to other eGFR equations (all of which are adjusted for BSA). For our analyses we adjusted Cockroft Gault for BSA using the Haycock formula.[86]

4-variable Modification of Diet in Renal Disease (4-vMDRD) equation[90]*

GFR = 175 x sCr$^{-1.154}$ x age$^{-0.203}$(years) (x 0.742 if female) (x 1.1212 if African American)

*re-expressed for IDMS assays traceable to a standard reference material for creatinine

2009 Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) creatinine equation[27]

GFR = 141 × min(sCr/κ, 1)$^{\alpha}$ × max(sCr/κ, 1)$^{-1.209}$ × 0.993^{age} (× 1.018 if female) (× 1.159 if African American)

κ is 0.7mg/dl (62 µmol/L) for females and 0.9mg/dl (80 µmol/L) for males

α is -0.329 for females and -0.411 for males

min indicates the minimum of sCr/κ or 1

max indicates the maximum of sCr/κ or 1

Equations expressed for specified sex and serum creatinine level[5]

Female ≤0.7 mg/dl (≤62 mmol/L) = 144 x (sCr/62)$^{-0.329}$ x 0.993^{age} [x 1.159 if African American]

Female >0.7 mg/dl (>62 mmol/L) = 144 x (sCr/62)$^{-1.209}$ x 0.993^{age} [x 1.159 if African American]

Male ≤0.9 mg/dl (≤80 mmol/L) = 141 x (sCr/80)$^{-0.411}$ x 0.993^{age} [x 1.159 if African American]

Male >0.9 mg/dl (>80 mmol/L) = 141 x (sCr/80)$^{-1.209}$ x 0.993^{age} [x 1.159 if African American]

2012 CKD-EPI cystatin C equation[91]

GFR = 133 x $\min(scysC/0.8, 1)^{-0.499}$ x $\max(scysC/0.8, 1)^{-1.328}$ x 0.996^{age} (x 0.932 if female)

min indicates the minimum of scysC/0.8 or 1

max indicates the maximum of scysC/0.8 or 1

Equations expressed for serum cystatin C level[5]

scysC ≤0.8: 133 x $(scysC/0.8)^{-0.499}$ x 0.996^{age} [x 0.932 if female]

scysC >0.8: 133 x $(scysC/0.8)^{-1.328}$ x 0.996^{age} [x 0.932 if female]

2012 CKD-EPI creatinine-cystatin C equation[91]

GFR = 135 x $\min(sCr/\kappa, 1)^{\alpha}$ x $\max(SCr/\kappa, 1)^{-0.601}$ x $\min(scysC/0.8, 1)^{-0.375}$ x $\max(scysC/0.8, 1)^{-0.711}$ x 0.995^{age} (x 0.969 if female) (x 1.08 if African American)

κ is 0.7mg/dl (62 µmol/L) for females and 0.9mg/dl (80 µmol/L) for males

α is -0.248 for females and -0.207 for males

min indicates the minimum of sCr/κ or 1

max indicates the maximum of sCr/κ or 1

min(scysC/0.8, 1) indicates the minimum of scysC/0.8 or 1

max(scysC/0.8,1) indicates the maximum of scysC/0.8 or 1

Equations expressed for specified sex, serum creatinine, and serum cystatin C level[5, 92]

Female sCr≤62 cysC≤0.8 = 130 x $(sCr/62)^{-0.248}$ x $(scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]

Female sCr≤62 cysC>0.8 = 130 x $(sCr/62)^{-0.248}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]

Female sCr>62 cysC≤0.8 = 130 x $(sCr/62)^{-0.601}$ x $(scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]

Female sCr>62 cysC>0.8 = 130 x $(sCr/62)^{-0.601}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]

Male sCr≤80 cysC≤0.8 = 135 x $(sCr/80)^{-0.207}$ x $(/scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]

Male sCr≤80 cysC>0.8 = 135 x $(sCr/80)^{-0.207}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]

Male sCr≤80 cysC≤0.8 = 135 x $(sCr/80)^{-0.601}$ x $(scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]

Male sCr≤80 cysC>0.8 = 135 x $(sCr/80)^{-0.601}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]

Revised Lund-Malmö Study equation[15]

GFR = eX–0.0158×Age+0.438×ln(Age)

Female sCr<150: X = 2.50+0.0121 × (150–sCr)

Female sCr≥150: X = 2.50–0.926 × ln(sCr/150)

Male sCr<180: X = 2.56+0.00968 × (180–sCr)

Male sCr≥180: X = 2.56–0.926 × ln(sCr/180)

Full Age Spectrum (FAS) creatinine equation[93]

$$GFR = \frac{107.3}{(sCr/Q)} \quad \text{(if ≤40 years)}$$

$$GFR = \frac{107.3}{(sCr/Q)} \times 0.988^{(Age\ -40)} \quad \text{(if >40 years)}$$

[sCr/Q (female) = 0.70mg/dL or 61.88µmol/L
sCr/Q (male) = 0.90mg/dL or 79.56µmol/L.[94]]

Determining population reference creatinine for the FAS (creatinine) equation

One of the variables needed to estimate GFR using the FAS (creatinine) equation is a population-specific reference serum creatinine measurement.(93, 94) The population prevalence data from the baseline ARK studies was used to establish country-specific population reference creatinine ranges. To do this, all serum creatinine measures ≥30µmol/L from apparently healthy adults were included (healthy adults were defined as those without HIV infection, hypertension, diabetes and obesity (BMI >30kg/m^2). All participants recruited for the iohexol mGFR study were excluded to ensure independent datasets. The median creatinine for men and women, by study site, was used to calculate the Q-value for the FAS equation.

Table 2: Population reference serum creatinine measurements for the FAS equation, by country and sex

Country	Sex	Sample size (N)	Serum Creatinine (median, µmol/L)
Malawi	Female	1693	67
	Male	1542	84
South Africa	Female	269	64
	Male	337	81
Uganda	Female	2111	59
	Male	1585	72

Data management and statistical analysis

Anonymized datasets from each country were pooled and participants who met all of the following criteria were included in further analyses: complete recordings of age, sex, height and weight; exact times for administering intravenous iohexol (T0) and subsequent sampling; iohexol plasma concentrations at each timepoint, demonstrating monotonic decline; pre- and post-administration syringe weights; and serum creatinine concentration ≥30µmol/L.

Distribution of data overall and by country was examined, and agreement between mGFR and creatinine and cystatin C-based eGFR equations was evaluated using Bland-Altman plots. For MDRD and CKD-EPI equations, performance was evaluated with and without ethnicity coefficients. Performance for each eGFR equation was compared to the reference mGFR using bias, precision and accuracy, and the proportion of participants correctly classified as having CKD by mGFR stage, was compared.

The parameters for performance expressed as GFR (mL/min/1.73m^2) were:
absolute bias – median difference of each individual's eGFR and their respective mGFR (eGFR-mGFR);
relative bias – median percentage difference of (eGFR-mGFR), expressed as a percent;
precision – as interquartile range (IQR) of (eGFR-mGFR), and as root mean square error (RMSE), representing standard deviation of (eGFR-mGFR);

accuracy – median of the absolute difference of (eGFR-mGFR), expressed as a percent of mGFR for estimates within 10% (P_{10}) and within 30% (P_{30}) of the mGFR.

To evaluate whether our findings were influenced by systematic error, sensitivity analyses were performed with three restricted datasets: (i) r (correlation coefficient) >0.985 for the slope-intercept measured GFR derivation; (ii) normal sex-specific calculated volumes of distribution as 11-17L for women, and 13-20L for men; and (iii) a double restricted dataset combining (i) and (ii).(14)

Development of a new creatinine-based GFR estimating equation

To develop a new eGFR model using the whole iohexol GFR sample, candidate predictors were kept to a minimum, and creatinine was chosen over cystatin C to ensure clinical utility in SSA. Candidate predictors were examined in models regressing log mGFR on log creatinine, with covariates of age and sex, with and without body mass index (BMI), mirroring the functional form of the CKD-EPI (creatinine) and Lund-Malmö eGFR equations. Regression coefficients were allowed to vary by sex and the model included a spline knot, with the position chosen using lowess plots. Likelihood ratio tests and adjusted R^2 were used to decide between candidate models and assessed model performance as previously defined, compared to existing creatinine-based eGFR equations. For external validation, performance of the new eGFR model was compared to existing eGFR equations and to mGFR from radionuclide studies of 651 individuals (South Africa) whose creatinine measurements, age and sex were available.

Determining true population prevalence of reduced kidney function in SSA

Since the predictive power of all eGFR equations was limited, including the newly modelled eGFR equation, a multiple imputation model was developed, trained on the mGFR sample to predict individual GFR based on creatinine, age and sex. Predictions were based on 100 imputed datasets. The imputation model builds a predictive distribution around each unobserved GFR, with variance and mean determined through a linear regression on the chosen predictor variables. This approach was used in six distinct large population-representative datasets from (i) the baseline prevalence studies in the ARK Consortium comprising adults between 20 and 80 years from Malawi, Uganda, and SA

(N= 12 454)(60, 66) and (ii) the AWI-Gen in Ghana, Kenya, Burkina Faso and SA, which recruited adults between 40 and 60 years (N= 11,701).(67, 68)

Funding

The SA study was jointly funded by the SA-MRC, MRC UK (via the Newton Fund), and GSK Africa Non-Communicable Disease Open Lab (Project Number: 074). The Uganda and Malawi studies were funded by a project grant from the GSK Africa Non-Communicable Disease Open Lab (Project number: 8111). Additional sources of funding for the SA project were obtained from the International Society for Nephrology (ISN) Clinical Research Program: 15-2-015_Validation of eGFR equations in South Africans, the National Health Laboratory Services (NHLS), Faculty of Health Sciences Research Incentive Grant; Grant number: 00128384342035121105000000000000000004550; University of the Witwatersrand; FHS (Wits).

Results

Determining the prevalence of CKD and associated risk factors in Bushbuckridge, Mpumalanga, South Africa (SA)

The first two objectives of this doctoral work are addressed in paper one (the full text can be found in the appendix).

Paper one (submitted to the Kidney International Journal, undergoing review)

Prevalence and associated risk factors for chronic kidney disease in a rural South African setting

June Fabian MD, Mwawi Gondwe, Nokthula Mayindi, Shingirai Chipungu, Bongekile Khoza, Alisha N. Wade PhD, F. Xavier Gómez-Olivé PhD, Laurie A. Tomlinson PhD, Michele Ramsay, Stephen Tollman PhD, Cheryl Winkler PhD, [*]Jaya George PhD,[*]Saraladevi Naicker PhD, on behalf of the ARK Consortium. [[*]Joint senior authors]

Summary: Determining the prevalence of CKD in rural Agincourt

We conducted the study between November 2017 and September 2018. We completed recruitment of a population-based sample of 2 021 of 2 759 participants between 20 and 80 years (a 73% response rate). Overall, our study population was relatively young, with mean age of 38 (SD ±14) years; nine percent had no formal education (predominantly women); and 30% were in the lowest quartile of the household assets score. With the initial screening for impaired kidney function, we had complete data on 2 004 of 2 021 participants. The WHO age- standardised prevalence of low GFR (<60ml/min/1.73m^2) was 1.6% (95% CI 1.1-2.2), albuminuria was 8.2% (95% CI 7.1-9.5), and low GFR and/or albuminuria was 9.1% (95% CI 7.9-10.4). Follow-up confirmed eGFR<60ml/min/1.73m^2 in 41% (12/29) of participants, and albuminuria in 53% (118/221) of participants. The overall WHO age-standardised CKD prevalence was 4.0% (95% CI 3.1-4.9).

Summary: Determining risk factors associated with CKD in rural Agincourt

Regarding CKD risk, population frequency of the high risk *APOL1* genotype was 11%. One quarter of men, and more than one third of women were overweight, with disproportionate obesity in women, approaching 40%. Many more men smoked and had higher rates of elevated blood pressure and hyperuricaemia than women. However, the prevalence of HIV

infection, hypercholesterolemia and diabetes mellitus, was higher in women. For those fulfilling the temporal definition of CKD, multivariable modelling identified the following as independently associated risk factors for CKD: a high risk *APOL1* genotype (OR 1.95; 95% CI 1.20-3.17); hypertension (OR 3.34; 95% CI 1.92-5.79); diabetes mellitus (OR 3.28; 95% CI 1.61-6.70); HIV infection (OR 1.89; 95% CI 1.18-3.03); and hyperuricaemia (OR 1.85; 95% CI 1.13-3.05).

Measuring kidney function in Sub-Saharan Africa: Individual and population-level implications

The last four objectives of this doctoral research are addressed in this section, grouped as a core theme synthesizing related aspects of measuring and estimating kidney function, how this impacts CKD diagnosis, and the consequent implications for individual and population level health. Three papers included in this section represent various aspects of this work (the full texts can be found in the appendix).

Paper two

Methods and reporting of kidney function: a systematic review of studies from Sub-Saharan Africa (95)

June Fabian, Jaya A. George, Harriet R. Etheredge, Manuel van Deventer, Robert Kalyesubula, Alisha N. Wade, Laurie A. Tomlinson, Stephen Tollman, Saraladevi Naicker on behalf of the African Research on Kidney Disease (ARK) Working Group. Clinical Kidney Journal 2019, vol. 12, no. 6, 778-787 *https://doi.org/10.1093/ckj/sfz089*

Paper three

How to estimate glomerular filtration rate in Sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study (85)

[*]Robert Kalyesubula, [*]June Fabian, Wisdom Nakanga, Robert Newton, Billy Ssebunnya, Josephine Prynn, Jaya George, Alisha N. Wade, Janet Seeley, Dorothea Nitsch, Christian Hansen, Moffat Nyirenda, Liam Smeeth, Saraladevi Naicker, Amelia C. Crampin, and Laurie A. Tomlinson.

[*] Joint first authors.

BMC Nephrology 2020, vol. 21, no. 20, 1-12. https://doi.org/10.1186/s12882-020-1688-0

Paper four (submitted to the New England Journal of Medicine, undergoing review)

Measurement of kidney function in Sub-Saharan Africa

[*]June Fabian, [*]Robert Kalyesubula, Joseph Mkandawire, Christian Holm Hansen, Dorothea Nitsch, Eustasius Musenge, Wisdom P Nakanga, Josephine E Prynn, Gavin Dreyer, Tracy Snyman, Billy Ssebunnya, Michele Ramsay, Liam Smeeth, Stephen Tollman, Saraladevi Naicker, Amelia Crampin, Robert Newton, Jaya A George, Laurie Tomlinson on behalf of the African Research on Kidney Disease (ARK) Consortium.

[*]Joint first authors.

Thematic overview: assessing kidney function in Sub-Saharan Africa

In paper two, we appraised the methods for measuring and reporting kidney function in SSA and identified critical knowledge gaps for future work. Paper three is a natural extension of paper two, as we introduced research partners from Malawi and Uganda, and detailed our methods for the proposed work, and meticulously designed and harmonised across the three partner countries. Collectively, throughout the study design, we tried to anticipate any shortfalls and limitations from prior work. Harmonising the protocols ensured study procedures were standardised, and our combined platform became a good support for each partner, to maintain the integrity of the study. Our strategy was to allow for analyses by country, for between country comparisons, and for a pooled analysis of all three countries. Given the biodiversity of African populations, we needed to explore whether findings were population specific, or potentially applicable to the region. Paper four is a culmination of 18 months of ongoing analyses, rigour and scientific interpretation of our findings, by the smaller working group. The manuscript is understandably complex, as it synthesizes critical findings that impact individual and population kidney health in SSA. The supplementary appendix to paper four is an in-depth collation of our analyses. We shared these with our readers to facilitate open scientific scrutiny by our peers, and to disseminate essential information that might facilitate replication of our work in other countries, particularly in SSA.

Analytic bias

The overarching hypothesis underpinning this work was that available eGFR equations overestimate true kidney function in African populations. To explore our hypothesis, we needed to understand contributing factors, their mechanisms, and potential interactions with one another. The first consideration was whether laboratories in SSA measure creatinine correctly. For creatinine, there are many sources of analytic bias that affect eGFR estimates and these might be accentuated in resource-limited settings. For example, delays in centrifugation and laboratory processing of blood samples after collection can affect the measured creatinine concentration (particularly when using the Jaffe method). In underserviced areas, especially rural, such delays are common. Furthermore, many laboratories use the Jaffe method because it is much cheaper than the enzymatic assay. The Laboratory Working Group of the National Kidney Disease Education Program (NKDEP) published guidelines in 2006 to standardise laboratory creatinine measurement, with considerable global uptake, but more so in well-resourced settings.(21) Ideally, all laboratories should use an IDMS-traceable assay and an enzymatic method for creatinine measurement.

We explored analytic bias in paper two where the primary focus was to determine laboratory implementation of the NKDEP guidelines in SSA. We conducted a systematic review of all published research from SSA between 2008 and 2018, in which kidney function was measured. In summary, more than two thirds of studies did not report their laboratory methods for creatinine measurement; more than 80% failed to report whether their creatinine method was standardised (IDMS-traceable); and over 90% of laboratories still used the Jaffe method. The MDRD equation was most frequently used, followed by the CKD-EPI creatinine equation, and only half of the studies stated whether they applied ethnicity coefficients. Only six studies measured GFR and compared performance of eGFR equations. It was clear that potential for bias in the measurement of creatinine and estimates of kidney function in SSA was substantial. Furthermore, limitations in sample size and study design compromised the potential scientific contribution from many published studies. We were also aware of the paucity of population-representative mGFR studies.

It follows that reducing as many sources as possible of analytic bias was critical when designing the iohexol mGFR study, a core component of our methods paper (paper three). As the collective ARK Consortium, we held two meetings (one hosted by SA and the other by Malawi) where we discussed at length the measures we would implement to minimise analytic bias. We also held regular virtual meetings throughout the project to troubleshoot as many issues as possible, in real time. All partner laboratories confirmed they used IDMS-traceable creatinine assays; all participated in an accredited external quality assurance program and were willing to share internal quality control data (if needed); and had external quality assurance certification. To address potential sources of inter-laboratory bias, we agreed to conduct a split sample recalibration study, essentially repeating measures from one partner laboratory in another, and calibrating results to accommodate any difference.

Iohexol GFR

For measuring GFR using iohexol plasma excretion, there were no established centres of excellence in any partner country that offered such testing as part of clinical service. In SA, GFR could be measured using radionuclide markers at selected specialist referral centres with nuclear medicine facilities. In contrast, in many high-income settings measuring GFR using iohexol is a routine clinical service, with a body of expertise for training the study staff. Furthermore, despite widespread use of iohexol in high-income settings, there are no consensus guidelines on how to use it to measure GFR, so protocols are centre-specific and not standardised. In the absence of mGFR guidelines for iohexol, we used the British Nuclear Medicine Society Guidelines(14) as our reference.

We consulted widely with experts in the UK, agreed upon a standard protocol, and initiated the study in Uganda. We accepted there would be learnings at whichever of the study sites started first. As an example, there were no laboratory scales at the Uganda site to measure pre- and post-administration syringe weights, for calculating the dose of iohexol (this data was discarded in the final analysis). Once the Uganda team were established, the Malawi team travelled there to learn the required methods; likewise, the South African team trained in Malawi. The iohexol laboratory assay was established in SA, with the intention to train the Ugandan team after the purchasing of a mass spectrometer (unfortunately the equipment

was never purchased and the laboratory manager fell ill). Consequently, all the iohexol sample processing was centralised to the laboratory in SA.

Throughout the project we were acutely aware of minimising potential sources of systematic bias in the iohexol phase of the study. While this was an integral part of the paper three (the methods paper), it was also integral to paper four, titled “Measurement of kidney function in Sub-Saharan Africa” (submitted). We performed extensive sensitivity analyses by country, and overall, detailed in the supplementary appendix of paper 4.

Laboratory testing for creatinine and cystatin C

For some clarity regarding the laboratory testing of creatinine and cystatin C, the creatinine assays were completed at the start of the study (N=2578). Subsequently, after we had procured additional funding for cystatin C testing, we performed all the cystatin C the assays as a single batch in the SA and Uganda laboratories (for improved assay standardisation using kits from the same batch) at the end of the study. Unfortunately, as the Uganda team were completing their assays the COVID-19 lockdown was implemented and remaining cystatin C assays were not completed (N=2433).

Discussion

Synopsis of this doctoral work

The purpose of undertaking this doctoral work was to address critical knowledge gaps regarding kidney disease prevalence and associated risk factors, and the accurate assessment of kidney function in Sub-Saharan African populations. While its focus was SSA and its regional contribution is substantial, the body of knowledge accrued contributes vital information to global nephrology. Presenting this work structured by discrete study objectives is necessary and comprises the remainder of this discussion. However, it is also somewhat artificial and may fragment the intention, findings and implications of this work. This synopsis provides an overall synthesis of the intention and the findings of this doctoral work, situates it in the global South, and contextualises it within the existing corpus of knowledge which is dominated by the global North.

As a nephrologist trained in South Africa, I became aware that simple answers to questions like “What causes kidney disease?” and “How much kidney disease is in South Africa?” were unanswered. Furthermore, work from my colleagues showed that GFR estimating equations appeared to overestimate true kidney function in South Africa.(45-47) Aside from clinical impression, a rigorous evaluation of the status quo was needed to inform future research questions – and this evaluation comprised the systematic review (referred to as “Paper two”). The systematic review identified that many laboratories in SSA did not following best practice guidelines for measurement of creatinine; most studies used the MDRD (and not the CKD-EPI) equation to estimate GFR; race-based coefficients were used widely throughout SSA without having been validated for use; the performance of GFR estimating equations and the impact of race-based coefficients had only been evaluated in a handful of small studies; and criteria for diagnosing CKD were haphazardly applied. Therefore, it was impossible to accurately assess kidney function at individual level, nor assess the burden of kidney disease at population level. As an outcome of the systematic review, we devised a checklist for reporting kidney function that could serve as a guide for future studies and used it to guide our subsequent study. We also knew our next research questions.

While the primary aim was to address these knowledge gaps in South Africa, through the ARK collaboration we could address similar gaps in Malawi and Uganda. Fortuitously,

Uganda and Malawi applied for separate grant funding to South Africa and both funding applications were successful and synchronous – ensuring we could pursue our research agenda. To agree upon a way forward, the first scientific advisory meeting for the ARK partner countries was held in Johannesburg in May 2016. As there were no kidney disease cohorts available in any of the ARK partner countries from which samples could be sourced for a measured GFR study (as is the case with countries in the north), and all partner countries had access to longitudinal population cohorts, we agreed to start with establishing the population prevalence of kidney disease in each partner country. The large population prevalence studies included an assessment of kidney function, a profile for associated risk for kidney disease, and would serve as the basis from which to derive the iohexol measured GFR study samples – enriched with participants who had decreased kidney function.

The second scientific advisory meeting for the ARK partner countries was held in Malawi in June 2018. The intention of this meeting was to ensure harmonisation of protocols across the study sites for: sampling; administration of iohexol and plasma sampling and storage; laboratory assays and their internal and external quality control measures; centralisation of the iohexol testing at the NHLS reference laboratory in Johannesburg; and the intersite laboratory calibration study for biomarkers analysed in each partner country (detailed in the *Thematic overview: assessing kidney function in Sub-Saharan Africa*).

The third ARK scientific advisory meeting was held in Uganda in November 2019 to review the preliminary results for the iohexol measured GFR study, the intersite laboratory calibration study results, and plan the way forward for the final analysis. At this, and subsequent meetings between January 2020 and March 2021, senior scientists in the ARK collaboration evaluated results from the analyses for: the pooled dataset; between country comparisons; sensitivity analyses to evaluate systematic bias; sourcing and performance of the external validation dataset; cystatin C-based GFR estimating equations; and performance of the modelling for the ARK equation. Although unexpected, in the absence of an improved creatinine-based ARK GFR estimating equation, we were forced to revise our statistical methodology to predict, more accurately, the prevalence of kidney disease in SSA. Through our collaboration with other studies and networks in SSA, we were able to source large population-based datasets from the AWI-Gen study that would complement the three partner countries in the ARK collaboration. Our study findings overall were replicated within

each country, sensitivity analyses did not identify sources of systematic error, and our conclusions can be summarised as follows: context-specific non-GFR determinants of creatinine and the Jaffe method for laboratory measurement make it a poor biomarker for the evaluation of kidney function; race-based coefficients for the MDRD and CKD-EPI equations substantially overestimate GFR in African populations; cystatin C performed better than creatinine – particularly in the prediction of GFR (CKD equivalent: Stage 3); and the population prevalence of kidney disease in SSA is underestimated using current GFR estimating equations.

Based on the findings of this work, recommendations that could be implemented with relative ease would be for laboratories to use the enzymatic method for measuring creatinine, for race-based coefficients to be omitted from GFR estimating equations, and cystatin C to be used as a confirmatory test for individuals in which kidney disease is suspected – particularly relevant in SSA given that many countries do not have access to confirmatory measured GFR testing. These recommendations should be formulated as regional guidelines for the clinical management of kidney disease as an alternative to the current KDIGO guidelines which are largely informed by data from the USA.

Future work will need to evaluate region-specific explanations for the poor performance of creatinine by including the impacts on functioning nephron mass of adverse developmental and life course events, investigating factors that influence non-GFR determinants of creatinine (including muscle mass), and identifying mechanisms for tubular secretion of creatinine and genetic mutations that might be specific to African populations. As alternatives to creatinine, modelling an ARK cystatin C-based GFR estimating equation for use in SSA and exploring other biomarkers that could be used either alone or in combination with cystatin C (such as beta-2 microglobulin and beta trace protein) must be evaluated to improve the accuracy of GFR estimation in sub-Saharan African populations.

Prevalence and associated risk for chronic kidney disease in a rural South African setting

The following original contributions to the existing body of knowledge on CKD in SA have been made in this doctoral research:

- This was the only study from SSA to have complied with the KDIGO requirements for the confirmation of CKD, with follow-up screening after a minimum of 3 months, for eGFR and albuminuria. In so doing, we demonstrated that single-screen epidemiological studies are likely to overestimate CKD prevalence.
- This was the first population-based study, from a rural community in SA, in which genotypic and phenotypic risk factors for CKD were well characterised.
- This was the first study in a Xitsonga-ethnic group to determine the population prevalence of high risk *APOL1* genotypes for kidney disease. Furthermore, it showed that high risk *APOL1* genotypes were independently associated with CKD, namely persistent albuminuria, not low GFR.
- Finally, this was the first study to demonstrate a significant association between CKD and hyperuricaemia in a rural South African population. This potentially signifies substantial cardiometabolic risk in a community undergoing rapid epidemiological transition, and likely resembles that of other SSA rural populations.

While this is the first study of its kind to be conducted in rural South Africans, there are no comparable data for urban populations of black ethnicity in SA. It is pertinent to consider whether these results can be generalised, as characterisation of risk for rural communities is likely to differ from urban communities, thus affecting CKD prevalence. However, the widely held assumption that urban communities are at higher risk for CKD is potentially flawed because the risk in rural communities is largely unknown. Potential risk in rural communities comprises untreated endemic infections such as genitourinary schistosomiasis, other genitourinary or systemic infections such as tuberculosis, infectious or immune-mediated glomerulonephritides, high rates of HIV infection, exposure to environmental pollutants (predominantly agricultural) and toxins (heavy metals from acid-wash of in the coal mining belt); and herbal toxin ingestion. The contribution to CKD from repeated bouts of acute kidney injury from intermittent infectious diseases, such as malaria, also require elucidation. In addition, the impact of NCDs that confer increased risk for CKD is high in rural

communities – for example hypertension and diabetes. While we did characterise APOL1 high risk allele frequencies, further research is required into the factors that might comprise the "second hit" needed for APOL1-associated nephropathies, and for the gene-environment interactions as more communities in traditionally rural areas undergo rapid epidemiological transition.

Aside from our study, there is only one other African study (Morocco) that repeated eGFR and urine albumin measurements to confirm CKD.(96) They showed similar findings, with at least one third of participants with dipstick positive proteinuria or eGFR<60ml/min/1.73m^2 not confirmed at follow-up.(96) In our data, transient albuminuria occurred most frequently in young women, and for both sexes it declined with age, while the corollary was true for persistent albuminuria. Potential causes of transient albuminuria in the rural Agincourt community require further investigation but may relate to undiagnosed urinary schistosomiasis or other urinary tract infections, and acute febrile illnesses such as malaria (endemic). Albuminuria that might regress in response to an undiagnosed or untreated medical condition is also a potential cause – as all participants with elevated blood pressure, hyperglycaemia, or HIV (not on treatment) were referred for care to their nearest facility. It is noteworthy that most of the cases of hyperglycaemia we detected on screening occurred in women who were unaware of a prior diagnosis of diabetes so starting treatment prior to rescreening in the study might have impacted repeat screening results. Our results confirm that rescreening is essential to confirm CKD, and that potentially treatable causes of albuminuria need to be considered by attending clinicians – particularly untreated or undiagnosed infections or medical conditions. These findings must inform future studies and become a template for promulgating region-specific CKD screening guidelines.

Recent large single-screen population-based studies from ARK study sites in Malawi,(60) Uganda(66) and the AWI-Gen(68) reported rates of reduced eGFR (eGFR <60ml/min/1.73m^2) similar to our study, and substantially lower than predicted for the region.(52) Similar methodologies in the studies, namely IDMS-traceable creatinine assays, eGFR using the CKD-EPI (creatinine) equation without the ethnicity coefficient, and age-standardisation, allowed for regional comparison. For comparison, the National Health and Nutrition Examination Survey (NHANES) Unites States population survey (1999-2004) reported 8.05% of the population over 30 years with eGFR <60ml/min/1.73m^2(97) and in the Health Survey for

England, 7% of the population over 35 years had eGFR <60ml/min/ 1.73m^2.(98) A potential explanation might be that creatinine-based eGFR equations are inappropriate in SSA and might have overestimated GFR in our and other studies, with resultant under reporting of lower eGFR. Alternative considerations are whether the eGFR cut-off (eGFR<60ml/min /1.73m^2) is appropriate for SSA populations. With reason, the utility of applying this threshold across all age groups is the source of intense debate.(49, 99) In the context of aging populations with a senescent decline in kidney function, especially in the absence of albuminuria, a decline in eGFR to 45ml/min/1.73m^2 might be completely normal. At the risk of over diagnosing CKD in the elderly, the opposite is true in relatively young populations, such as those in our study. Given younger people with low GFR are more likely to have a potentially treatable glomerular condition, it might be reasonable to consider an eGFR cut-off <75ml/min/1.73m^2, in the presence of albuminuria, as reason to suspect CKD. A threshold of eGFR<75ml/min/1.73m^2 has been proposed for those <40 years.(49) Another potential explanation for the relatively small population percentage with low GFR might be due to a survivor effect, as those with severe progressive chronic kidney disease demise in the absence of appropriate care.(11)

In this work, we confirmed findings from prior work in SA and SSA that have likewise demonstrated significant risk for CKD with HIV infection, hypertension and diabetes.(52-54, 56, 60, 66, 68, 100) While the population frequency of *APOL1* high risk genotypes in Agincourt was substantially higher than in an admixed population from the Western Cape,(57) it was similar to frequencies observed in the African diaspora population, now resident in the USA, in which the high risk genotype confers a lifetime risk of developing kidney disease of approximately 20%.(8, 101) *APOL1*-associated kidney disease requires a "second hit" that activates interferon pathways, as occurs with viral infection or autoimmune disease. *APOL1*, an innate immune factor, is potently upregulated by interferons.(9) Potential "second hits" include the high prevalence of HIV infection,(102) as HIV-associated nephropathy has been strongly associated with the high risk *APOL1* genotype in other parts of SA.(58) The association between the high risk *APOL1* genotype and CKD was driven by persistent albuminuria and not reduced eGFR, an observation recently reported from a study of HIV-infected adults from Nigeria.(103) While ours was a baseline cross-sectional study, longitudinal studies from relatively young cohorts in the USA, with mostly

preserved kidney function, showed that high risk *APOL1* genotypes predispose to an earlier age of onset of incident albuminuria coincident with a decline in eGFR.(104) A subsequent study found that *APOL1* high risk genotypes are associated with a greater risk of proteinuria, and that the onset of proteinuria was followed by a rapid decline in eGFR in persons with CKD attributed to hypertension.(105)

The association we observed between hyperuricaemia and CKD has previously been reported in a hospital-based sample from Kwa-Zulu Natal(106) and this study confirms the association in a rural African population.(107) Studies from the 1970s describe higher rates of hyperuricaemia in urban than rural South Africans,(108, 109) yet we observed similar concentrations in men from Agincourt. Paradoxically, women in Agincourt have lower rates of hyperuricaemia, despite relatively higher obesity and diabetes (also noted in studies from the North West Province).(110) Whether hyperuricaemia is a predictor of cardiovascular, metabolic or renal outcomes, and whether urate-lowering therapy (both affordable and readily available) modifies these outcomes, has not been proven and remains to be determined in this context.(111, 112)

Measuring kidney function in Sub-Saharan Africa: Individual and population-level implications

The following original contributions to the existing body of knowledge on CKD in SA have been made through this doctoral research:

- In three Sub-Saharan African countries, comprised of diverse urban and rural populations, creatinine-based eGFR equations substantially overestimate kidney function compared to GFR measured with iohexol (mGFR).
- The overestimation worsened at lower levels of mGFR and was further exacerbated by the inclusion of ethnicity coefficients.
- Cystatin C-based equations performed better than creatinine-based equations but none of the eGFR equations achieved an accuracy (P_{30}), considered appropriate for individual clinical decision-making.(5)
- At population level, using creatinine-based eGFR equations substantially underestimated the prevalence of impaired kidney function.

- A new creatinine-based equation to better estimate GFR in SSA was not possible due to wide age- and sex-independent variability in the relationship between creatinine and mGFR, even with weight or BMI adjustment.

Our finding that creatinine-based eGFR equations overestimate well-preserved kidney function (stages G1-2), but underestimate declining kidney function (stages G3-5), has been replicated in studies from Scandinavia and West Africa.(42, 92) However, it differs from those in Japan and the USA.(38, 90, 91) Inaccuracies of GFR estimation from creatinine-based equations have been described in studies from Asia,(40) and ascribed to lower muscle mass. There are reasons why the relationship between creatinine and GFR may be more variable in SSA, compared to US or European populations. Our study and others reported lower BMI, with likely lower muscle mass and therefore lower baseline creatinine, in continental Africans when compared to the African diaspora.(41, 44, 60, 66) Perinatal and early childhood factors, resulting in undernutrition and growth stunting, predisposes to low, lean muscle mass and short stature in adulthood (even in the presence of adult obesity), and are common in SSA.(69, 113-115) Wasting from chronic infection or inflammation, such as tuberculosis and HIV infection, low dietary protein ingestion, and undiagnosed liver disease also impact muscle mass and creatinine generation.(116-119) Finally, variants in genes affecting pathways of creatinine production and tubular secretion may impact the association between serum creatinine and true GFR in SSA.(120, 121)

The implications of our findings need to be considered at an individual and population level. Overall, we do not recommend using ethnicity coefficients for calculating eGFR with the MDRD or CKD-EPI equations. For an individual, overestimating GFR could potentially result in delayed diagnosis of CKD, missed opportunities to address modifiable risk factors to slow the progression of kidney disease, and possible harm from unadjusted doses of renally cleared drugs. At population level, underestimation of the burden of CKD in SSA limits prioritisation of country-specific and regional strategies to minimise CKD progression and manage ESKD. Limited access to specialist renal care in SSA compounds the individual risk for progression, premature disability, and death. While cystatin C identified a greater proportion of individuals with CKD – and might be a valuable confirmatory test – assays are more expensive than creatinine and not widely available in SSA.

While not in the scope of this work, establishing a framework of clinical guidelines informed by our findings for the region would be an appropriate next step. For CKD screening using GFR, there is irrefutable data to support the transition of diagnostic laboratories away from Jaffe methods to enzymatic assays, reinforcing the NKDEP guidelines.(21); this should be implemented in all our national laboratories. In those individuals with impaired kidney function, especially CKD stage 3, a confirmatory test using cystatin C would likely improve the sensitivity of confirmatory screening.

Another contention would be to review whether the dichotomous eGFR <60ml/min "cut-off", as the diagnostic criterion for CKD, is relevant in our relatively young populations (it may be more appropriate to flag an eGFR <75ml/min/1.73m^2 in those younger than 40 years of age, for rescreening).

For individuals with albuminuria, confirmatory screening would be mandatory, given the high prevalence of transient albuminuria. Those with persistent albuminuria would require risk factor management, given the potential association with high risk *APOL1* genotypes. Having said that, we cannot prognosticate outcomes associated with GFR or ACR in SSA populations without longitudinal cohort studies that have quantified associated risk.

Another consideration is whether we should be using creatinine at all to estimate GFR in our setting. Failure to model a creatinine-based eGFR equation that better predicts kidney function in our work is testament to the poor performance of creatinine as a biomarker. While this work has focused on minimising analytic bias, which is relevant, the biological causes for variation in creatinine measurement requires attention. Biological variation in serum creatinine has many contributing factors, as discussed. While there is some assumption that continental Africans have higher muscle mass than other population groups, based on findings in those of African heritage, now resident in high income countries, our findings contradict this.

Another critical aspect of managing CKD lies in the diagnostics. We have witnessed the evolution of HIV management, from expensive assays in hospital-initiated treatment models, to point-of-care, inexpensive, real-time testing at primary care level, using standardised treatment protocols. While we know the true burden of CKD is underestimated, CKD screening and management needs to be restructured for scalability in SSA. To achieve this,

accurate and affordable point-of-care technologies for CKD are urgently needed. A sub-study of this work in Agincourt evaluated capillary and venous point-of-care creatinine testing and point-of-care urine dipstick and ACR testing. Our results showed that POC creatinine devices are as accurate, if not more so than pathology laboratory testing (particularly when diagnostic laboratories use the Jaffe method). In addition, POC urine testing for albuminuria has high specificity, useful for screening algorithms.(62, 63) Community-based CKD screening and prevention programmes in Thailand(122) have been implemented and could inform similar initiatives in SSA.

The use of ethnicity coefficients in US-based MDRD and CKD-EPI equations has recently come under the spotlight and sparked intense debate, mostly confined to the USA, regarding racialised aspects of medicine.(123-125) With our history of Apartheid in SA, the issues raised resonate deeply and mandate critical reflection. As a starting point, why are research findings from the African diaspora assumed to hold true for all African populations. Also, why are many aspects of nephrology practice, such as the KDIGO CKD guideline, which is very US-centric, assumed to be appropriate in African populations. As an example, the Revised Lund-Malmo Study equation performed the best in all three of the ARK countries, and the CKD-EPI creatinine equation, particularly with inclusion of the ethnicity coefficient, was one of the worst. This work will contribute, in part, to informing critical decision-making regarding the appropriateness of eGFR equations in some African populations, while acknowledging that additional work remains to be done.

Limitations

For the CKD prevalence and associated risk study in SA, our cross-sectional study design precluded longitudinal follow-up and restricted interpretation of our findings. This relates specifically to understanding the impact of CKD on cardiovascular and all-cause mortality in our setting. Additionally, the association with hyperuricaemia and high risk *APOL1* genotypes requires longitudinal follow-up to ascertain the risk for CKD progression. Numerous life course and environmental exposures potentially impact kidney function in our rural communities, and these remain unquantified. While we evaluated traditional risk factors for CKD, we did not explore the impact of endemic and other infections (aside from HIV),

exposure to traditional and over-the-counter medicines, and environmental toxins (agricultural and heavy metal).

For the iohexol measured GFR study, an expert team in the UK reviewed our iohexol data and one of their observations was that our volumes of distribution for iohexol appeared larger that those they routinely observed in the UK. While their experience was with radionuclide tracers and not iohexol, we discovered there were no population references for "normal" volumes of distribution in SSA and were unable to cross reference our findings. Consequently, the volume of distribution cut-offs applied to our dataset were derived from European populations in much colder climes and might not have been appropriate.

We only measured creatinine and mGFR at a single time point in the iohexol mGFR study (paper 4), thereby not fulfilling the temporal requirements for the definition of CKD. Despite this, our approach was in keeping with other large cohort studies of GFR.(51) In addition, participants were well and therefore, aside from seasonal changes in eGFR,(20) we would not have anticipated substantial variation in kidney function. Our sampling frame was population-based, with high mean kidney function compared to CKD cohorts used to develop GFR equations.(23, 91,126) Measurement of low creatinine values is more prone to biological and analytic variations than high levels.(127) It is possible that variability in direct GFR measurement inherent to a study largely conducted in rural SSA, despite rigorous quality control procedures, may have contributed to the reduced performance of GFR estimating equations. The diversity of African populations means that there is likely to be greater variability in the relationship between creatinine and GFR than we have measured. This could affect the accuracy of the imputed GFR estimates in the AWI-Gen datasets, but in internal validation, the model estimates prevalence of impaired kidney function close to that measured in the ARK datasets, and with differences substantially less than the variation we see between creatinine-based and imputed estimates. Also, pertaining to the biodiversity of African populations, while we replicated our findings in Malawi, SA and Uganda, we would be reluctant to assume our findings are generalisable to other SSA populations.

Future Work

Using our baseline findings, we have the opportunity to contextualise future work that is responsive to the communities in which research is conducted. The Agincourt HDSS site, alongside the other ARK and AWI-Gen sites in West, East and Southern Africa, are perfectly positioned, with access to longitudinal population cohorts that are able to address longstanding research questions. Most importantly, these would address adverse outcomes (cardiovascular, renal and all-cause mortality), for those with any stage of kidney impairment. Such information would inform population health policy and clinical practice.

During the iohexol mGFR study in Agincourt, we ensured that all participants had a body composition DXA scan which included lean muscle mass measurement. We plan to use this data to investigate the relationship between muscle mass, creatinine and mGFR, and examine correlates with lean muscle mass estimated with portable bioimpedance devices in all three partner countries. Recently, genome-wide association studies of kidney function indicators from our colleagues in Uganda identified a new mutation involved in renal tubular handling of creatinine. While it remains to be clarified whether this is a loss or gain of function mutation, it may very well be another potential explanation for low serum creatinine, in the case of enhanced tubular excretion, independent of muscle mass.(121) A smaller sub-study in Uganda was conducted to evaluate 24-hour urinary creatinine excretion and these data will be analysed to understand tubular creatinine secretion. Similar studies will be replicated in South Africa for future work.

As an extension of the iohexol mGFR work, we will determine the accuracy of single point iohexol GFR compared to multiple point (we used three points), and correlate dried blood spot capillary iohexol GFR to plasma excretion (this component of the study was performed in Uganda). Lastly, using the iohexol GFR samples, we will be exploring the accuracy of additional biomarkers, such beta trace protein and β2-microglobulin, as alternative biomarkers for estimating GFR. The role of cystatin C as an alternative biomarker for estimating GFR will be explored through acquiring funding for population-based cystatin C testing among the three sites.

Conclusion

In a relatively young, rural South African community undergoing rapid epidemiological transition, CKD is prevalent and associated with hereditable risk and combined burdens of infectious and cardiometabolic disease. However, several context-specific risk factors for CKD along the life course remain to be explored in further work. In Malawi, SA and Uganda, creatinine-based GFR estimating equations overestimate kidney function, compared to iohexol and cystatin C measures, suggesting the true burden of kidney disease is markedly underestimated in Sub-Saharan Africa. This has substantial implications for individual health care and public health interventions to address the challenge of kidney disease in resource-limited settings.

References

1. Bikbov B, Purcell CA, Levey AS, Smith M, Abdoli A, Abebe M, et al. Global, regional, and national burden of chronic kidney disease, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2020;395(10225):709-33.
2. Go AS, Chertow GM, Fan D, McCulloch CE, Hsu C. Chronic kidney disease and the risks of death, cardiovascular events, and hospitalization. N Engl J Med. 2004;351(13):1296-305.
3. Gansevoort RT, de Jong PE. The case for using albuminuria in staging chronic kidney disease. J Am Soc Nephrol. 2009;20(3):465-8.
4. Matsushita K, van der Velde M, Astor BC, Woodward M, Levey AS, de Jong PE, et al. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. Lancet. 2010;375(9731):2073-81.
5. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int Suppl. 2013;3:1-150.
6. Forbes A, Gallagher H. Chronic kidney disease in adults: assessment and management. Clin Med (Lond). 2020;20(2):128.
7. Luyckx VA, Bertram JF, Brenner BM, Fall C, Hoy WE, Ozanne SE, et al. Effect of fetal and child health on kidney development and long-term risk of hypertension and kidney disease. Lancet. 2013;382(9888):273-83.
8. Yusuf AA, Govender MA, Brandenburg JT, Winkler CA. Kidney disease and APOL1. Hum Mol Genet. 2021.
9. Nichols B, Jog P, Lee JH, Blackler D, Wilmot M, D'agati V, et al. Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1. Kidney Int. 2015;87(2):332-42.
10. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, et al. Chronic kidney disease: global dimension and perspectives. Lancet. 2013;382(9888):260-72.
11. Ashuntantang G, Osafo C, Olowu WA, Arogundade F, Niang A, Porter J, et al. Outcomes in adults and children with end-stage kidney disease requiring dialysis in sub- Saharan Africa: a systematic review. Lancet Glob Health. 2017;5(4):e408-e17.
12. Delanaye P, Ebert N, Melsom T, Gaspari F, Mariat C, Cavalier E, et al. Iohexol plasma clearance for measuring glomerular filtration rate in clinical practice and research: a review. Part 1: how to measure glomerular filtration rate with iohexol? Clin Kidney J. 2016;9(5):682- 99.

13. Bröchner-Mortensen J. A simple method for the determination of glomerular filtration rate. Scand J Clin Lab Invest. 1972;30(3):271-4.

14. Fleming JS, Zivanovic MA, Blake GM, Burniston M, Cosgriff PS. Guidelines for the measurement of glomerular filtration rate using plasma sampling. Nucl Med Commun. 2004;25(8):759-69.

15. Soveri I, Berg UB, Björk J, Elinder C-G, Grubb A, Mejare I, et al. Measuring GFR: a systematic review. Am J Kidney Dis. 2014;64(3):411-24.

16. Nordin G, Ekvall S, Kristoffersson C, Jonsson A-S, Bäck S-E, Rollborn N, et al. Accuracy of determination of the glomerular filtration marker iohexol by European laboratories as monitored by external quality assessment. Clin Chem Lab Med. 2019;57(7):1006-11.

17. Nilsson-Ehle P. Iohexol clearance for the determination of glomerular filtration rate: 15 years' experience in clinical practice. EJIFCC. 2001;13(2):48.

18. Delanaye P, Melsom T, Ebert N, Bäck S-E, Mariat C, Cavalier E, et al. Iohexol plasma clearance for measuring glomerular filtration rate in clinical practice and research: a review. Part 2: Why to measure glomerular filtration rate with iohexol? Clin Kidney J. 2016;9(5):700- 4.

19. Benoit SW, Ciccia EA, Devarajan P. Cystatin C as a biomarker of chronic kidney disease: latest developments. Expert Rev Mol Diagn. 2020:1-8.

20. Ephraim RKD, Asamoah CA, Abaka-Yawson A, Kwadzokpui PK, Adusei S. Climate change causes changes in biochemical markers of kidney disease. BMC Nephrol. 2020;21(1):1-8.

21. Boutten A, Bargnoux A-S, Carlier M-C, Delanaye P, Rozet E, Delatour V, et al. Enzymatic but not compensated Jaffe methods reach the desirable specifications of NKDEP at normal levels of creatinine. Results of the French multicentric evaluation. Clin Chim Acta. 2013;419:132-5.

22. Chikomba CE, Padoa CJ, Tanyanyiwa D. Evaluation of the impact of delayed centrifugation on the diagnostic performance of serum creatinine as a baseline measure of renal function before antiretroviral treatment. South Afr J HIV Med. 2020;21(1):1-15.

23. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Ann Intern Med. 1999;130(6):461-70.

24. Cohn S, Abesamis C, Zanzi I, Aloia J, Yasumura S, Ellis K. Body elemental composition: comparison between black and white adults. Am J Physiol Endocrinol Metab. 1977;232(4):E419.

25. Harsha DW, Frerichs RR, Berenson GS. Densitometry and anthropometry of black and white children. Hum Biol. 1978:261-80.

26. Worrall J, Phongsathorn V, Hooper R, PAICE EW. Racial variation in serum creatine kinase unrelated to lean body mass. Br J Rheumatol. 1990;29(5):371-3.

27. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.

28. Peralta CA, Risch N, Lin F, Shlipak MG, Reiner A, Ziv E, et al. The association of African ancestry and elevated creatinine in the Coronary Artery Risk Development in Young Adults (CARDIA) Study. Am J Nephrol. 2010;31(3):202-8.

29. Hsu J, Johansen KL, Hsu CY, Kaysen GA, Chertow GM. Higher serum creatinine concentrations in black patients with chronic kidney disease: beyond nutritional status and body composition. Clin J Am Soc Nephrol. 2008;3(4):992-7.

30. Jacobs Jr DR, Murtaugh MA, Steffes M, Yu X, Roseman J, Goetz FC. Gender-and race-specific determination of albumin excretion rate using albumin-to-creatinine ratio in single, untimed urine specimens: the Coronary Artery Risk Development in Young Adults Study. Am J Epidemiol. 2002;155(12):1114-9.

31. Delanaye P, Mariat C, Maillard N, Krzesinski J-M, Cavalier E. Are the creatinine-based equations accurate to estimate glomerular filtration rate in African American populations? Cl J Am Soc Nephrol. 2011;6(4):906-12.

32. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis. 2002;39(2 Suppl 1):S1-266.

33. Ma Y-C, Zuo L, Chen J-H, Luo Q, Yu X-Q, Li Y, et al. Modified glomerular filtration rate estimating equation for Chinese patients with chronic kidney disease. J Am Soc Nephrol. 2006;17(10):2937-44.

34. Kong X, Ma Y, Chen J, Luo Q, Yu X, Li Y, et al. Evaluation of the Chronic Kidney Disease Epidemiology Collaboration equation for estimating glomerular filtration rate in the Chinese population. Nephrol Dial Transplant. 2013;28(3):641-51.

35. Jeong TD, Lee W, Yun YM, Chun S, Song J, Min WK. Development and validation of the Korean version of CKD-EPI equation to estimate glomerular filtration rate. Clin Biochem. 2016;49(9):713-9.

36. Praditpornsilpa K, Townamchai N, Chaiwatanarat T, Tiranathanagul K, Katawatin P, Susantitaphong P, et al. The need for robust validation for MDRD-based glomerular filtration rate estimation in various CKD populations. Nephrol Dial Transplant. 2011;26(9):2780-5.

37. Chen L-I, Guh JY, Wu KD, Chen YM, Kuo MC, Hwang SJ, et al. Modification of diet in renal disease (MDRD) study and CKD epidemiology collaboration (CKD-EPI) equations for Taiwanese adults. PLoS One. 2014;9(6).

38. Horio M, Imai E, Yasuda Y, Watanabe T, Matsuo S. Modification of the CKD epidemiology collaboration (CKD-EPI) equation for Japanese: accuracy and use for population estimates. Am J Kidney Dis. 2010;56(1):32-8.

39. Imai E, Horio M, Nitta K, Yamagata K, Iseki K, Hara S, et al. Estimation of glomerular filtration rate by the MDRD study equation modified for Japanese patients with chronic kidney disease. Clin Exp Nephrol. 2007;11(1):41-50.

40. Teo BW, Zhang L, Guh JY, Tang SCW, Jha V, Kang DH, et al. Glomerular Filtration Rates in Asians. Adv Chronic Kidney Dis. 2018;25(1):41-8.

41. Eastwood JB, Kerry SM, Plange-Rhule J, Micah FB, Antwi S, Boa FG, et al. Assessment of GFR by four methods in adults in Ashanti, Ghana: the need for an eGFR equation for lean African populations. Nephrol Dial Transplant. 2010;25(7):2178-87.

42. Bukabau JB, Sumaili EK, Cavalier E, Pottel H, Kifakiou B, Nkodila A, et al. Performance of glomerular filtration rate estimation equations in Congolese healthy adults: the inopportunity of the ethnic correction. PLoS One. 2018;13(3):e0193384.

43. Bukabau JB, Yayo E, Gnionsahé A, Monnet D, Pottel H, Cavalier E, et al. Performance of creatinine-or cystatin C–based equations to estimate glomerular filtration rate in sub-Saharan African populations. Kidney Int. 2019;95(5):1181-9.

44. Wyatt CM, Schwartz GJ, Ong'or WO, Abuya J, Abraham AG, Mboku C, et al. Estimating kidney function in HIV-infected adults in Kenya: comparison to a direct measure of glomerular filtration rate by iohexol clearance. PLoS One. 2013;8(8):e69601.

45. van Deventer HE, George JA, Paiker JE, Becker PJ, Katz IJ. Estimating glomerular filtration rate in black South Africans by use of the modification of diet in renal disease and Cockcroft-Gault equations. Clin Chem. 2008;54(7):1197-202.

46. van Deventer HE, Paiker JE, Katz IJ, George JA. A comparison of cystatin C-and creatinine-based prediction equations for the estimation of glomerular filtration rate in black South Africans. Nephrol Dial Transplant. 2010:gfq621.

47. Seape T, Gounden V, van Deventer HE, Candy GP, George JA. Cystatin C-and creatinine-based equations in the assessment of renal function in HIV-positive patients prior to commencing highly active antiretroviral therapy. Ann Clin Biochem. 2016;53(1):58-66.

48. Winearls CG, Glassock RJ. Dissecting and refining the staging of chronic kidney disease. Kidney Int. 2009;75(10):1009-14.

49. Glassock R, Delanaye P, El Nahas M. An age-calibrated classification of chronic kidney disease. JAMA. 2015;314(6):559-60.

50. Coresh J. Update on the burden of CKD. J Am Soc Nephrol. 2017;28(4):1020-2.

51. Brück K, Stel VS, Gambaro G, Hallan S, Völzke H, Ärnlöv J, et al. CKD prevalence varies across the European general population. J Am Soc Nephrol. 2016;27(7):2135-47.

52. Kaze AD, Ilori T, Jaar BG, Echouffo-Tcheugui JB. Burden of chronic kidney disease on the African continent: a systematic review and meta-analysis. BMC Nephrol. 2018;19(1):1-11.

53. Madala N, Thusi G, Sibisi N, Mazibuko B, Assounga A. CKD and CVD Risk Factors in the poorly-resourced northern Kwazulu-Natal, South Africa. Nephrol Dial Transpant. 2012:27:380.

54. Adeniyi AB, Laurence CE, Volmink JA, Davids MR. Prevalence of chronic kidney disease and association with cardiovascular risk factors among teachers in Cape Town, South Africa. Clin Kidney J. 2017:sfw138.

55. Davids MR, Jardine T, Marais N, Zunza M, Jacobs JC, Sebastian S. South African Renal Registry Annual Report 2017. Afr J Nephrol. 2019;22(1):60-71.

56. Matsha TE, Yako YY, Rensburg MA, Hassan MS, Kengne AP, Erasmus RT. Chronic kidney diseases in mixed ancestry south African populations: prevalence, determinants and concordance between kidney function estimators. BMC Nephrol. 2013;14(1):1.

57. Matsha TE, Kengne AP, Masconi KL, Yako YY, Erasmus RT. APOL1 genetic variants, chronic kidney diseases and hypertension in mixed ancestry South Africans. BMC Genet. 2015;16(1):69.

58. Kasembeli AN, Duarte R, Ramsay M, Mosiane P, Dickens C, Dix-Peek T, et al. APOL1 risk variants are strongly associated with HIV-associated nephropathy in black South Africans. J Am Soc Nephrol. 2015;26(11):2882-90.

59. Muiru AN, Charlebois ED, Balzer LB, Kwarisiima D, Elly A, Black D, et al. The epidemiology of chronic kidney disease (CKD) in rural East Africa: A population-based study. PLoS One. 2020;15(3):e0229649.

60. Nakanga WP, Prynn JE, Banda L, Kalyesubula R, Tomlinson LA, Nyirenda M, et al. Prevalence of impaired renal function among rural and urban populations: findings of a cross-sectional study in Malawi. Wellcome Open Res. 2019;4:92.

61. Luyckx VA, Al-Aly Z, Bello AK, Bellorin-Font E, Carlini RG, Fabian J, et al. Sustainable Development Goals relevant to kidney health: an update on progress. Nat Rev Nephrol. 2020.

62. Currin SD, Gondwe MS, Mayindi NB, Chipungu S, Khoza BL, Tollman S, et al. Diagnostic accuracy of semiquantitative point of care urine albumin to creatinine ratio and urine dipstick analysis in a primary care resource limited setting in South Africa. BMC Nephrol. 2021;22(1):1-9.

63. Currin S, Gondwe M, Mayindi N, Chipungu S, Khoza B, Khambule L, et al. Evaluating chronic kidney disease in rural South Africa: comparing estimated glomerular filtration rate using point-of-care creatinine to iohexol measured GFR. Clin Chem Lab Med. 2021.

64. Xie Y, Bowe B, Mokdad AH, Xian H, Yan Y, Li T, et al. Analysis of the Global Burden of Disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. Kidney Int. 2018;94(3):567-81.

65. Nishtar S, Niinistö S, Sirisena M, Vázquez T, Skvortsova V, Rubinstein A, et al. Time to deliver: report of the WHO Independent High-Level Commission on NCDs. Lancet. 2018;392(10143):245-52.

66. Kalyesubula R, Hau JP, Asiki G, Ssebunya B, Kusemererwa S, Seeley J, et al. Impaired renal function in a rural Ugandan population cohort. Wellcome Open Res. 2018;3:149.

67. Ramsay M, Crowther N, Tambo E, Agongo G, Baloyi V, Dikotope S, et al. H3Africa AWI-Gen Collaborative Centre: a resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. Glob Health Epidemiol Genom. 2016;1.

68. George JA, Brandenburg J-T, Fabian J, Crowther NJ, Agongo G, Alberts M, et al. Kidney damage and associated risk factors in rural and urban sub-Saharan Africa (AWI-Gen): a cross-sectional population study. Lancet Glob Health. 2019;7(12):e1632-e43.

69. Kahn K, Collinson MA, Gómez-Olivé FX, Mokoena O, Twine R, Mee P, et al. Profile: Agincourt health and socio-demographic surveillance system. Int J Epidemiol. 2012;41(4):988-1001.

70. MRC/Wits-Agincourt Unit, Agincourt Health and Demographic Site, cartographer. HDSS Study Area – Research Villages. South Africa; 2020.

71. Rohr JK, Xavier Gómez-Olivé F, Rosenberg M, Manne-Goehler J, Geldsetzer P, Wagner RG, et al. Performance of self-reported HIV status in determining true HIV status among older adults in rural South Africa: a validation study. J Int AIDS Soc. 2017;20(1):21691.

72. National Department of Health. National consolidated guidelines for the prevention of mother-to-child transmission of HIV (PMTCT) and the management of HIV in children, adolescents and adults. Pretoria, South Africa: National Department of Health; 2015.

73. David VA, Binns-Roemer EA, Winkler CA. Taqman Assay for Genotyping CKD- Associated APOL1 SNP rs60910145: A Cautionary Note. Kidney Int Rep. 2019;4(1):184-5.

74. World Health Organization. WHO STEPS surveillance manual: the WHO STEPwise approach to chronic disease factor surveillance. Geneva: World Health Organization. 2005.

75. Kabudula CW, Houle B, Collinson MA, Kahn K, Tollman S, Clark S. Assessing changes in household socioeconomic status in rural South Africa, 2001–2013: a distributional analysis using household asset indicators. Soc Indic Res. 2017;133(3):1047-73.

76. Alberti K, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood

institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. Circulation. 2009;120(16):1640-5.

77. World Health Organization. Managing the Global Epidemic of Obesity: Report of the World Health Organization Consultation of Obesity. Geneva: World Health Organization. 1997.

78. Chobanian AV, Bakris GL, Black HR, Cushman WC, Green LA, Izzo JL, Jr., et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. JAMA. 2003;289(19):2560-72.

79. American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2018. Diabetes Care. 2018;41(Supplement 1):S13-S27.

80. Klug E, Raal F, Marais A, Smuts C, Schamroth C, Jankelow D, et al. South African dyslipidaemia guideline consensus statement: 2018 update A joint statement from the South African Heart Association (SA Heart) and the Lipid and Atherosclerosis Society of Southern Africa (LASSA). S Afr Med J. 2018;108(11b):973-1000.

81. Tietz AH. clinical guide to laboratory tests. Missouri: WB Saunders Company. 2006;1138.

82. World Health Organization. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Geneva: World Health Organization. 2011.

83. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. Geneva: World Health Organization. 2001;9(10).

84. Twine R. Public engagement throughout the research process in a health and socio-demographic surveillance system study area in rural South Africa [dissertation]. Johannesburg: University of Witwatersrand; 2019.

85. Kalyesubula R, Fabian J, Nakanga W, Newton R, Ssebunnya B, Prynn J, et al. How to estimate glomerular filtration rate in sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study. BMC Nephrol. 2020;21(1):1-12.

86. Haycock GB, Schwartz GJ, Wisotsky DH. Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. J Pediatr. 1978;93(1):62-6.

87. Annesley TM, Clayton LT. Ultraperformance liquid chromatography–tandem mass spectrometry assay for iohexol in human serum. Clin Chem. 2009;55(6):1196-202.

88. Parrinello CM, Grams ME, Couper D, Ballantyne CM, Hoogeveen RC, Eckfeldt JH, et al. Recalibration of blood analytes over 25 years in the atherosclerosis risk in communities study: impact of recalibration on chronic kidney disease prevalence and incidence. Clin Chem. 2015;61(7):938-47.

89. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron. 1976;16(1):31-41.

90. Levey AS, Coresh J, Greene T, Stevens LA, Zhang YL, Hendriksen S, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Ann Intern Med. 2006;145(4):247-54.

91. Inker LA, Schmid CH, Tighiouart H, Eckfeldt JH, Feldman HI, Greene T, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. N Engl J Med. 2012;367(1):20-9.

92. Nyman U, Grubb A, Larsson A, Hansson L-O, Flodin M, Nordin G, et al. The revised Lund-Malmö GFR estimating equation outperforms MDRD and CKD-EPI across GFR, age and BMI intervals in a large Swedish population. Clin Chem Lab Med. 2014;52(6):815-24.

93. Pottel H, Hoste L, Dubourg L, Ebert N, Schaeffner E, Eriksen BO, et al. An estimated glomerular filtration rate equation for the full age spectrum. Nephrol Dial Transplant. 2016;31(5):798-806.

94. Pottel H, Vrydags N, Mahieu B, Vandewynckele E, Croes K, Martens F. Establishing age/sex related serum creatinine reference intervals from hospital laboratory data based on different statistical methods. Clin Chim Acta. 2008;396(1-2):49-55.

95. Fabian J, George JA, Etheredge HR, van Deventer M, Kalyesubula R, Wade AN, et al. Methods and reporting of kidney function: a systematic review of studies from sub-Saharan Africa. Clin Kidney J. 2019;12(6):778-87.

96. Gharbi MB, Elseviers M, Zamd M, Alaoui AB, Benahadi N, Trabelssi EH, et al. Chronic kidney disease, hypertension, diabetes, and obesity in the adult population of Morocco: how to avoid “over”-and “under”-diagnosis of CKD. Kidney Int. 2016;89(6):1363-71.

97. Castro AF, Coresh J. CKD surveillance using laboratory data from the population- based National Health and Nutrition Examination Survey (NHANES). Am J Kidney Dis. 2009;53(3):S46-S55.

98. Fat LN, Mindell J, Roderick P. Health Survey for England 2016 Kidney and liver disease. Health and Social Care Information Centre, London. 2017.

99. Glassock RJ, Delanaye P, Rule AD. Should the definition of CKD be changed to include age-adapted GFR criteria? YES. Kidney Int. 2020;97(1):34-7.

100. Madala ND, Thusi GP, Assounga AG, Naicker S. Characteristics of South African patients presenting with kidney disease in rural KwaZulu-Natal: a cross sectional study. BMC Nephrol. 2014;15(1):61.

101. Limou S, Nelson GW, Kopp JB, Winkler CA. APOL1 kidney risk alleles: population genetics and disease associations. Adv Chronic Kidney Dis. 2014;21(5):426-33.

102. Gómez-Olivé FX, Angotti N, Houle B, Klipstein-Grobusch K, Kabudula C, Menken J, et al. Prevalence of HIV among those 15 and older in rural South Africa. AIDS Care. 2013;25(9):1122-8.

103. Wudil UJ, Aliyu MH, Ingles DJ, Ahonkhai AA, Masa BM, Muhammad H, et al. Apoliproprotein-1 (APOL1) Risk Variants and Associated Kidney Phenotypes in an Adult HIV Cohort in Nigeria. Kidney Int. 2021:Apr 23:S0085-2538.

104. Peralta CA, Bibbins-Domingo K, Vittinghoff E, Lin F, Fornage M, Kopp JB, et al. APOL1 genotype and race differences in incident albuminuria and renal function decline. J Am Soc Nephrol. 2016;27(3):887-93.

105. Freedman BI, Burke W, Divers J, Eberhard L, Gadegbeku CA, Gbadegesin R, et al. Diagnosis, Education, and Care of Patients with APOL1-Associated Nephropathy: A Delphi Consensus and Systematic Review. J Am Soc Nephrol. 2021. Apr 14:ASN.2020101399.

106. Madala ND, Dubula T, Assounga AGH, Naicker S. Association of Kidney Function and Waist Circumference with Uric Acid Levels in South Africans. Metab Syndr Relat Disord. 2017;15(10):500-6.

107. Nakagawa T, Tuttle KR, Short RA, Johnson RJ. Hypothesis: fructose-induced hyperuricemia as a causal mechanism for the epidemic of the metabolic syndrome. Nat Clin Pract Nephrol. 2005;1(2):80-6.

108. Beighton P, Solomon L, Soskolne C, Sweet B, Robin G. Serum uric acid concentrations in an urbanized South African Negro population. Ann Rheum Dis. 1974;33(5):442.

109. Beighton P, Solomon L, Soskolne C, Sweet B. Serum uric acid concentrations in a rural Tswana community in Southern Africa. Ann Rheum Dis. 1973;32(4):346.

110. Palmer I, Schutte A, Huisman H, Van Rooyen J, Schutte R, Malan L, et al. A comparison of uric acid levels in Black African vs Caucasian women from South Africa: the POWIRS study. Ethn Dis. 2007;17(4):676.

111. Borghi C, Agabiti-Rosei E, Johnson RJ, Kielstein JT, Lurbe E, Mancia G, et al. Hyperuricaemia and gout in cardiovascular, metabolic and kidney disease. Eur J Intern Med. 2020.

112. Fernandez-Prado R, Ortiz A, Perez-Gomez MV. Allopurinol and Chronic Kidney Disease. N Engl J Med. 2020;383(17):1690-.

113. Wells JC. Body composition of children with moderate and severe undernutrition and after treatment: a narrative review. BMC Med. 2019;17(1):1-9.

114. Ntenda PAM. Association of low birth weight with undernutrition in preschool-aged children in Malawi. Nutr J. 2019;18(1):1-15.

115. Adebisi YA, Ibrahim K, Lucero-Prisno III DE, Ekpenyong A, Micheal AI, Chinemelum IG, et al. Prevalence and socio-economic impacts of malnutrition among children in Uganda. Nutr Metab Insights. 2019;12:1178638819887398.

116. Wasserman P, Segal-Maurer S, Rubin DS. High prevalence of low skeletal muscle mass associated with male gender in midlife and older HIV-infected persons despite CD4 cell reconstitution and viral suppression. J Int Assoc Provid AIDS Care. 2014;13(2):145-52.

117. Villamor E, Saathoff E, Mugusi F, Bosch R, Urassa W, Fawzi W. Wasting and body composition of adults with pulmonary tuberculosis in relation to HIV-1 coinfection, socioeconomic status, and severity of tuberculosis. Eur J Clin Nutr. 2006;60(2):163-71.

118. Spearman CW, Sonderup MW. Health disparities in liver disease in sub-Saharan Africa. Liver Int. 2015;35(9):2063-71.

119. Paruk IM, Pirie FJ, Motala AA. Non-alcoholic fatty liver disease in Africa: a hidden danger. Glob Health Epidemiol Genom. 2019;4.

120. Zanetti D, Rao A, Gustafsson S, Assimes TL, Montgomery SB, Ingelsson E. Identification of 22 novel loci associated with urinary biomarkers of albumin, sodium, and potassium excretion. Kidney Int. 2019;95(5):1197-208.

121. Fatumo S, Chikowore T, Kalyesubula R, Nsubuga RN, Asiki G, Nashiru O, et al. Discovery and fine-mapping of kidney function loci in first genome-wide association study in Africans. Hum Mol Genet. 2021 Mar30:ddab088.

122. Cha'on U, Pongskul C, Thinkhamrop B, Limwattananon C, Sharma A, Anutrakulchai S. CKDNET: a program for fighting with chronic kidney disease in the Northeast Thailand. BMC Public Health. 2020;27;20(1):1299.

123. Powe NR. Black Kidney Function Matters: Use or Misuse of Race? JAMA. 2020 25;324(8):737-738.

124. Levey AS, Titan SM, Powe NR, Coresh J, Inker LA. Kidney disease, race, and GFR estimation. Cl J Am Soc Nephrol. 2020;15(8):1203-12.

125. Grubbs V. Precision in GFR reporting: let's stop playing the race card. Am Soc Nephrol; 2020.

126. Biörk J, Bäck SE, Sterner G, Carlson J, Lindström V, Bakoush O, et al. Prediction of relative glomerular filtration rate in adults: New improved equations based on Swedish Caucasians and standardized plasma-creatinine assays. Scand J Clin Lab Invest. 2007;67(7):678-95.

127. Poggio ED, Wang X, Greene T, Van Lente F, Hall PM. Performance of the modification of diet in renal disease and Cockcroft-Gault equations in the estimation of GFR in health and in chronic kidney disease. J Am Soc Nephrol. 2005;16(2):459-66.

Appendices

Declaration by Student and Co-authors Agreement for

Published Papers

Papers 1-4

Human Research Ethics Committee

Approval Certificate Turn it in letter of

approval

Turn it in report

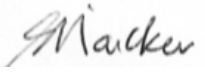
Declaration: Student's contribution to article(s) and agreement of co-author(s)

I, June Fabian, student number 8600952/N, declare that this thesis is my own work and that I contributed adequately towards research findings published in the article(s) stated below which are included in my Thesis/Dissertation/Research Report.

Signature of Student

Date: 27th April 2021

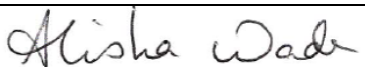
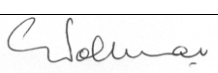
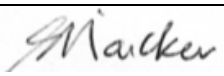
Name of Primary Supervisor: Saraladevi Naicker

Signature of Primary Supervisor  **Date** 27/04/2021

Agreement by co-authors: By signing this declaration, the co-authors listed below agree to the use of the article(s) by the student as part of his/her Thesis/Dissertation/Research Report. In cases where the student is not the 1st author of a published article, the primary supervisor must explain (under comments) why the student is entitled to use the paper for his/her degree purposes.

Article 1: Title: Methods and reporting of kidney function: a systematic review of studies from sub-Saharan Africa

Journal name, year, volume and page numbers: Clinical Kidney Journal, 2019, vol. 12, no. 6, 778–787

Authors	Name	Signature	Date
1st author	Jaya A. George		28/04/2021
2nd author	Robert Kalyesubula		28/04/2021
3rd author	Alisha N. Wade		27//04/21
4th author	Laurie A. Tomlinson		29/04/2021
5th author	Stephen Tollman		28/04/2021
6th author	Saraladevi Naicker		27/04 2021
7th author	Manuel van Deventer		28/04/2021

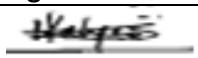
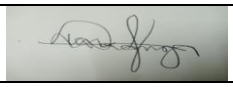
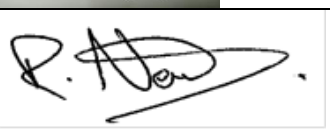
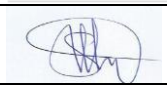
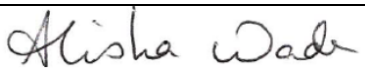
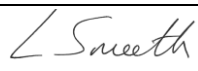
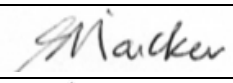
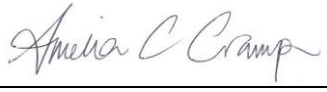
8th author	Harriet R. Etheredge		28/04/2021

Comments by primary supervisor: The student initiated and carried out the systematic review, collated all the data, and wrote the first draft of the manuscript, and there after collated all input from co-authors and finalized and submitted the manuscript for publication.

……

Article 2: Title: How to estimate glomerular filtration rate in sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study

Journal name, year, volume and page numbers: BMC Nephrology (2020) 21:20

Authors	Name	Signature	Date
1st author	Robert Kalyesubula		28/04/2021
2nd author	Wisdom Nakanga		29/4/2021
3rd author	Robert Newton		27/04/2021
4th author	Billy Ssebunya		29/04/2021
5th author	Josephine Prynn		28/04/2021
6th author	Jaya A George		28/04/2021
7th author	Alisha N Wade		27/04/2021
8th author	Janet Seeley		27/4/2021
9th author	Dorothea Nitsch		29/4/2021
10th author	Christian Holm Hansen		28/4/21
11th author	Moffat Nyirenda		29/4/2021
12th author	Liam Smeeth		28/04/2021
13th author	Saraladevi Naicker		27/07/2021
14th author	Amelia Crampin		29/04/2021

15th author	Laurie Tomlinson		29/04/2021

Comments by primary supervisor: The student wrote the study protocol for the Agincourt arm of the study, and was integral in harmonizing the protocol across the ARK Collaboration. She co-wrote the first draft of the manuscript and reviewed all input and corrections.

……..

Prevalence and associated risk factors for chronic kidney disease in a rural South African setting

Journal:	*Kidney International*
Manuscript ID	Draft
Article Type:	Clinical Investigation
Date Submitted by the Author:	n/a
Complete List of Authors:	Fabian, June; University of the Witwatersrand Faculty of Health Sciences, Wits Donald Gordon Medical Centre, School of Clinical Medicine; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt) Gondwe, Mwawi; University of the Witwatersrand School of Public Health, 1Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health Mayindi, Nokthula; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health Chipungu, Shingirai; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health Khoza, Bongekile; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health Wade, Alisha ; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health Gómez-Olivé, F. Xavier; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health Tomlinson, Laurie; London School of Hygiene and Tropical Medicine, Department of Epidemiology and Population Health Ramsay, Michele; University of the Witwatersrand, Sydney Brenner Institute for Molecular Bioscience Tollman, Stephen; University of Witwatersrand, MRC/Wits Agincourt Unit, Rural Public Health and Health Transitions Research Unit Winkler, Cheryl; National Cancer Institute, Basic Research Laboratory; Frederick National Laboratory for Cancer Research, SAIC-Frederick George, Jaya; University of the Witwatersrand Faculty of Health Sciences Naicker, Saraladevi; University of the Witwatersrand, Department of Internal Medicine
Subject Area:	Clinical Nephrology, Chronic Kidney Injury, Epidemiology and Statistics
Keywords:	chronic kidney disease, albuminuria, diabetes

Prevalence and associated risk factors for chronic kidney disease in a rural South African setting

Cohort

Our study partnered with a rural South African community from the northeastern part of South Africa. Embedded within the community is the Agincourt Health and socio-Demographic Surveillance System site, which includes 31 research villages, comprising 90 000 people from the Bushbuckridge sub-district in the Mpumalanga Province

Methods

In this cross-sectional study, we screened adults between 20-80 years for chronic kidney disease and associated risk factors. Using albuminuria (albumin:creatinine ratio (ACR) >=3mg/mmol)) and estimated GFR (eGFR<60ml/min/1.73m^2) and we confirmed persistent albuminuria or eGFR <60ml/min/1.73m^2 with repeat measurement after a minimum of 3 months. Genotyping was performed to determine *APOL1* renal risk alleles

Outcomes

The age-adjusted population prevalence of temporally confirmed CKD (a composite of albuminuria, eGFR <60ml/min/1.73m^2, or both) was 4.0% (95% CI 3.1-4.9). Follow-up confirmed eGFR<60ml/min/1.73m^2 in 41% (12/29) of participants, and albuminuria in 53% (118/221) of participants. In a multivariable model, a high risk *APOL1* genotype (two copies of G1 and/or G2) (OR 1.95; 95% CI 1.20-3.17), hypertension (OR 3.34; 95% CI 1.92-5.79), diabetes mellitus (OR 3.28 95% CI 1.61-6.70), HIV infection (OR 1.89; 95% CI 1.18-3.03), and hyperuricaemia (OR 1.85; 95% CI 1.13-3.05) were risk factors independently associated with CKD. Specifically, the high risk *APOL1* genotype was associated with persistent albuminuria, not reduced GFR.

Fabian, 2021

CKD is prevalent in a relatively young rural South African population undergoing rapid epidemiological transition. Risk for CKD is associated with heritable high risk *APOL1* genotypes, HIV infection, and cardiometabolic diseases, including hypertension, diabetes, and hyperuricaemia. Our findings contribute evidence towards the development of public health policy on screening those at risk for CKD in South Africa and potentially elsewhere in the region.

Prevalence and associated risk factors for chronic kidney disease in a rural South African setting

June Fabian MD,[1,2] Mwawi Gondwe,[1] Nokthula Mayindi,[1] Shingirai Chipungu,[1] Bongekile Khoza,[1] Alisha N. Wade PhD,[1] F. Xavier Gómez-Olivé PhD,[1,3] Laurie A. Tomlinson PhD,[4] Michele Ramsay,[5] Stephen Tollman PhD,[1,3] Cheryl Winkler PhD,[6] *Jaya George PhD,[7,8*]Saraladevi Naicker PhD,[9] on behalf of the ARK Consortium. *Joint senior authors

[1]Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2]Wits Donald Gordon Medical Centre, School of Clinical Medicine, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa

[3]International Network for the Demographic Evaluation of Populations and their Health (INDEPTH), Accra, Ghana

[4]Department of Non-communicable disease epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

[5]Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[6]Molecular Genetic Epidemiology Section, Basic Research Laboratory, Frederick National Laboratory for Cancer Research, NCI, Frederick, MD 21701, USA

[7]Department of Chemical Pathology, University of Witwatersrand, Johannesburg, South Africa

[8]Department of Chemical Pathology, National Health Laboratory Service, South Africa

[9]Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Corresponding author

June Fabian: primary email: june.fabian@mweb.co.za; secondary email: june.fabian@wits.ac.za;

mobile: +2782 785 3579

Abstract

Background

The prevalence and associated risk factors for chronic kidney disease (CKD) are poorly understood in South Africa and Sub-Saharan Africa. This study aimed to determine the population prevalence of CKD and evaluate genotypic and phenotypic risk for CKD in a rural South African setting.

Methods

In this cross-sectional study, 2021 adult participants aged 20 to 80 years were screened for CKD risk factors from the Agincourt Health and socio-Demographic Surveillance System site in rural Bushbuckridge, Mpumalanga Province. We determined CKD prevalence by measuring albuminuria (albumin: creatinine ratio (ACR) >=3mg/mmol)) and estimating GFR (eGFR<60ml/min/1.73m^2) using the CKD-EPI (creatinine) equation without ethnicity coefficient, confirmed with repeat measurement after a minimum three-month period. Genotyping was performed to determine *APOL1* renal risk alleles.

Results

The age-adjusted population prevalence of temporally confirmed CKD (a composite of albuminuria, eGFR <60ml/min/1.73m^2, or both) was 4.0% (95% CI 3.1-4.9). Follow-up confirmed eGFR<60ml/min/1.73m^2 in 41% (12/29) of participants, and albuminuria in 53% (118/221) of participants. In a multivariable model, a high risk *APOL1* genotype (two copies of G1 and/or G2) (OR 1.95; 95% CI 1.20-3.17), hypertension (OR 3.34; 95% CI 1.92-5.79), diabetes mellitus (OR 3.28 95% CI 1.61-6.70), HIV infection (OR 1.89; 95% CI 1.18-3.03), and hyperuricaemia (OR 1.85; 95% CI 1.13-3.05) were risk factors independently associated with CKD.

Conclusion

CKD is prevalent in a relatively young rural South African population undergoing rapid epidemiological transition. Risk for CKD is associated with heritable high risk *APOL1* genotypes, HIV

infection, and cardiometabolic diseases, including hypertension, diabetes, and hyperuricaemia. Our findings contribute evidence towards the development of public health policy on screening those at risk for CKD in South Africa and potentially elsewhere in the region.

Introduction

The prevalence of chronic kidney disease (CKD) in Sub-Saharan Africa (SSA) is unknown but predicted to be high due to combined burdens of infectious and non-communicable disease (NCD) expected to contribute to increased risk.[1] Furthermore, heterogeneity among studies due to differences in sampling and diagnostic criteria for CKD limits interpretation and comparisons between countries.[1] Morbidity and mortality associated with CKD are disproportionately higher than high-income settings in the absence of public health strategies for screening and early detection of CKD, together with minimal access to kidney replacement therapy (KRT) for those with severe CKD.[2] Furthermore, outcomes remain unknown for those with less severe CKD who do not progress to end-stage kidney disease (ESKD) yet remain at high risk of premature death.[3]

In South Africa, isolated non-population-based studies have identified hypertension, diabetes mellitus, and HIV infection as risk factors associated with CKD.[4,5] In those receiving KRT, the South African Renal Registry reported hypertension-associated kidney disease, diabetes mellitus, and primary glomerular disease as the most common identified causes of kidney disease, while the cause was unknown in one third.[6] To date, the only population-based study in South Africa to evaluate impaired kidney function comprised a mixed-ancestry cohort from the Western Cape with female sex, older age, and hypertension as significant risk factors.[7] Subsequently, associations between high risk genotypes (G1/G1; G2/G2; G1/G2) for the apolipoprotein L1 (*APOL1*) gene and (i) systolic hypertension and (ii) reduced GFR was described in the same cohort, although frequencies of the high risk genotypes were lower (1.01%) than those observed in the African diaspora population now resident in the USA (10-15%)[8]. A strong association between high risk *APOL1* genotypes and HIV-associated nephropathy has also been reported from Johannesburg, South Africa.[9] Aside from the Western Cape, the population burden of CKD and its associated risk profile remains unknown in South Africa.

In this study, our primary objective was to determine the population prevalence of CKD in a rural South African setting using the Kidney Disease: Improving Global Outcomes (KDIGO) criteria.[10] Our secondary objective was to determine associated risk factors for CKD, including high risk *APOL1* genotypes and a combination of infectious and non-communicable disease phenotypes.

Methods

Study Setting and Sampling Strategy

This cross-sectional study was conducted in the Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt) in Bushbuckridge, a rural subdistrict of the Mpumalanga province, in the northeastern part of South Africa.[11] The Agincourt Health and Socio-Demographic Surveillance System (HDSS) site (90 000 people) was established in 1992. Annual updates of vital information on residents have facilitated a longitudinal platform supporting observational and interventional work along the life course involving, among others, individual and population health, migration, and urbanization.[11]

A minimum sample size of 1800 was required to provide at least 80% power to determine CKD prevalence >5%, provided the true prevalence was ≥6.5%. To ensure the study sample was population-representative for adults aged 20 to 80 years, proportional allocation of participants by sex and age-strata was performed based on the Agincourt HDSS population census round conducted in 2014. The sample size was increased proportionately to 2759 individuals to accommodate non-participation and high levels of migration for work.

Participant Recruitment and Study Procedures

Ethics approval was obtained from the Medical Human Research Ethics Committee, University of the Witwatersrand (certificate number M170583). From November 2017 to September 2018, a team of locally trained fieldworkers and nurses performed home visits in 31 research villages in the Agincourt HDSS.[12] For each participant recruited to the study, fieldworkers obtained written,

informed consent in the participant's first language (primarily Shangaan), administered a CKD risk questionnaire, performed anthropometry which included height, weight, hip and waist circumference measurement, and measured blood pressure.

For all anthropometric measures, participants were asked to remove heavy clothing and footwear. Weight was measured to the nearest 0,1kg using a digital flat scale (Seca, Germany), and weighing scales were calibrated weekly. Standing height was measured to the nearest 0,1cm with a portable stadiometer (Seca, Germany). Height and weight were used to calculate body mass index (BMI) as weight(kg)/height(m)^2. Waist and hip circumference were measured to the nearest 0.1cm with a circumferential tape measure (Seca, Germany). Waist circumference was measured at the midpoint between the lowest palpable rib and the iliac crest in the mid-axillary plane. Hip circumference was measured at the most protruding part of the buttocks with the participant standing. Waist and hip circumferences were measured three times, and the average of each metric was used for analyses. Blood pressure (BP) was measured using automated upper arm devices (Omron M6W, Intellisense BP785 large cuff, Japan). Participants were asked to sit comfortably with legs uncrossed for five minutes before readings were taken. Three measurements were taken using an appropriate-sized cuff on the left arm at two-minute intervals. Of three BP readings, the first was discarded, and an average of the second and third readings used for analyses.

Nurses performed capillary point of care (POC) random cholesterol and random glucose testing (Cardiochek® PA analyzer, PTS Panels® test strips, PTS Diagnostics, USA), and POC hemoglobin testing (HemoCue® Hb 201+ analyzer and test strips, Norway). Their respective suppliers calibrated all POC and blood pressure monitoring devices before commencing the study.

If a participant knew their HIV status as positive, this was recorded, and they were asked about adherence to antiretroviral therapy (ART).[13] If participants did not know their HIV status or had previously tested negative, nurses offered voluntary POC screening and testing (Alere™ HIV Combo,

Abbott, USA) according to South African Department of Health guidelines.[14] A positive test result was confirmed with a second test (Uni-Gold™ Recombigen® HIV-1/2, Trinity BioTech, USA). Nurses collected blood samples and a freshly voided urine sample for laboratory testing. A spot urine pregnancy test was performed for premenopausal women (Abon™ One Step Pregnancy Test, Pharmaland, UAE). All samples collected in the field were stored in isothermally padded bags prepacked with ice-bricks to maintain temperatures between 2 and 6 °C and temperatures in each bag were monitored (Easylog, Lascar Electronics, UK). Each day, after completion of fieldwork, the team delivered all samples to the Agincourt Research laboratory for onsite processing and storage at -80°C according to standardized protocols.

Laboratory Procedures

For DNA, buffy coats from centrifuged EDTA specimens were lifted and stored in the Agincourt Research Laboratory and shipped to the Sydney Brenner Institute for Molecular Biosciences (SBIMB) for DNA extraction and biobanking. A 20µl aliquot of DNA was shipped to the Frederick National Laboratory at the National Cancer Institute, USA, for *APOL1* genotyping as described in David et al.[15] DNA was genotyped using TaqMan assays (ThermoFisher Scientific, USA). *APOL1* G1 allele comprises two missense variants, rs73885319 (G1g) and rs60910145 (G1m); however, the presence of only the G1g variant is sufficient to define the G1 risk allele. The *APOL1* G2 allele consists of a six bp in-frame deletion, rs717185313. High risk genotypes were defined as those containing two high risk alleles (G1/G1, G1/G2, or G2/G2); low risk genotypes were defined as those containing zero or one risk allele. The number of *APOL1* risk alleles (either G1 or G2 allele) carried by each subject was then coded as 0 for the G0/G0 genotype, 1 for the G0/G1 or G0/G2 genotype, or 2 for the G1/G1, G1/G2, or G2G2 genotypes. The *APOL1* genotypes were further coded as “high renal risk (HR) genotype” if the subject carried any combination of 2 risk alleles or “low renal risk (LR) genotype” if the subject has 0 or 1 risk alleles.[15]

For serum and urine testing, all specimens were shipped at -80°C to the Central Laboratory Services (CLS) in Johannesburg. Serum creatinine was measured by an IDMS-traceable modified Jaffe method using a Cobas© 6000/c501 analyzer. Uric acid was measured by enzymatic colorimetric testing on a Roche/Hitachi Cobas© c analyzer and hepatitis B surface antigen titers (I1000SR Abbott, USA)) using a Roche/Hitachi Cobas© c analyzer, defined as positive or negative based on a prespecified laboratory cut-off value. Urine microalbumin was measured using a colorimetric method on the Cobas© 6000/c501 analyzer, and urine creatinine by the modified Jaffe method on the same analyzer and ACR was calculated. The CLS laboratory adhered to standard daily internal quality control procedures and complied with the requirements of the external quality control program through the College of American Pathologists (CAP) with approved certification for urine and serum chemistry for the study period.

Study definitions

The CKD risk factor questionnaire was designed using the WHO Stepwise approach to surveillance (STEPS) as a framework.[16] For individual participants, the highest level of education and a household assets score were received from the Agincourt HDSS.[17] Increased waist circumference was defined as >=80cm for women and >=94cm for men. Increased waist to hip ratio was defined as >0.85 for women and >1.0 for men.[18] BMI was used to classify: underweight (<18.5); normal weight (18.5-24.99); overweight (25.0-29.99); obese (≥30.0): obese class I (30.0-34.99); obese class II (35.0-39.99); obese class III (>=40.00).[19] Participants were classified according to the 7th Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure as: normotensive (systolic SBP<120mmHg and DBP<80mmHg); prehypertensive (SBP≥120mmHg and <140mmHg or DBP≥80mmHg and <90mmHg); or hypertensive (SBP≥140mmHg or DBP≥90mmHg).[20] Diabetes was defined as a non-fasting glucose >=11.1mmol/L,[21] hypercholesterolaemia as a non-fasting total cholesterol >5.0mmol/L,[22] hyperuricaemia as uric acid >0.43mmol/L in men and >0.36mmol/L in women, [23] and anaemia as haemoglobin (Hb) levels <11.0g/dL in pregnant women;

<12.0 g/dL in non-pregnant women and <13.0 g/dL in men.[24] For apolipoprotein L1 (*APOL1*) genotyping, analyses were performed using high vs. low renal risk stratification, where individuals with zero or one risk allele were defined as low renal risk; and individuals with two risk alleles defined as high renal risk.

Characterizing Kidney Disease

As per the KDIGO guidelines[10] CKD was measured using estimated glomerular filtration rate (eGFR <60ml/min/1.73m^2) calculated using the CKD-EPI (creatinine) equation without ethnicity coefficient, [25] and albuminuria (albumin: creatinine ratio (ACR) >=3.0ml/mmol). Based on our initial screen, all participants with an eGFR<60 ml/min/1.73m^2 or urine ACR>=3.0mg/mmol were revisited by the study team for repeat sampling, where possible, after a minimum period of 3 months to fulfill the temporal requirements for the definition of CKD. Once confirmed, CKD was further classified[10] by eGFR stage (ml/min/1.73m^2) (G1: eGFR >=90; G2: eGFR 60-89; G3a: eGFR 45-59; G3b: eGFR 30-44; G4: eGFR 15-29; G5: eGFR <15), and ACR stage (mg/mmol) (A1: <3; A2: 3-30; A3: >30) to stratify severity and risk for adverse outcomes. All data were age-adjusted using the revised WHO World Standard Population Distribution.[26]

Data Analysis

Continuous variables were represented as mean (± standard deviation) if normally distributed and median (interquartile range) if non-normally distributed. Categorical variables were expressed as frequencies (percentage). To identify factors independently associated with CKD, we defined CKD as a dichotomous variable which included those in whom albuminuria (ACR>=3mg/mmol) and /or eGFR <60ml/min/1.73m^2 were confirmed with follow up. To determine risk factors associated with CKD we used logistic regression to estimate odds ratios (OR), with corresponding 95% confidence intervals (95% CIs) and a parsimonious, forward stepwise approach to develop a multivariable model adjusting for age, sex, phenotypic, genotypic (*APOL1* high vs. low renal risk), infectious, and metabolic comorbidities. Using the same approach, we determined whether the same risk factors

profile was associated with markers of impaired kidney function at first screening (as opposed to confirmed at follow-up as CKD) using ACR>=3.0mg/mmol and eGFR <60ml/min/1.73m^2 as dichotomous variables; and we performed bivariate analyses using ACR and eGFR as continuous and dichotomous variables for the screening and follow up visits to determine any association with *APOL1* high vs. low renal risk genotypes. Missing data were reported as missing and not imputed. Statistical analyses were performed using STATA 15 SE (Stata Corp, Texas, USA).

Participant Referral for Care

Participants with abnormal tests detected by the study team were referred to a local primary health care clinic for confirmatory testing and further management. Referrals were facilitated through an active, ongoing engagement program with the communities and the Agincourt HDSS, health care workers in local clinics, regional hospitals, and the district Department of Health.[27]

Results

Sociodemographic and clinical characteristics of ARK participants in Agincourt, South Africa

Between 2nd November 2017 and 5th September 2018, we approached 2759 potential participants and of these, we recruited 2021 participants representing a response rate of 73%. The final study sample was representative of the Agincourt HDSS population, comparing age and sex (Figure S1). Derivation of the study sample, reasons for non-participation, and sampling for first and repeat screening to diagnose CKD are detailed in Figure 1. Participant sociodemographic and clinical characteristics are summarised in Table 1, stratified by sex. Overall, our participants mean age was 38 (SD ±14) years, almost ten percent had no formal education (predominantly women), and approximately one third were in the lowest quartile of the household assets score. The population frequency of the high risk *APOL1* genotype (two alleles: G1/G1; G1/G2; G2/G2) was 11%. One-quarter of men and more than one-third of women were overweight, with disproportionate obesity in women approaching 40%. Many more men smoked and had higher rates of hypertension and

hyperuricaemia than women. On the contrary, the prevalence of diabetes mellitus, hypercholesterolemia, and HIV infection was higher in women.

Kidney function in Agincourt, South Africa

First screening: impaired kidney function

On initial screening, complete data for markers of kidney function (eGFR and ACR) were available in 2004 of 2021 consented participants (Figure 1) and used for analyses. Overall, the unadjusted prevalence of eGFR <60ml/min/1.73m^2 was 1.6% (95% CI 1.1-2.2), albuminuria was 12.3% (95% CI 10.9-13.9), and a combination (eGFR <60ml/min/1.73m^2 and/or albuminuria) was 13.1% (95% CI 11.7-14.7) (Table S1, and stratified by sex in Tables S2, S3). After age-adjustment, the prevalence of eGFR <60ml/min/1.73m^2 was 1.6% (95% CI 1.1-2.2), albuminuria was 8.2% (95% CI 7.1-9.5), and a combination (eGFR <60ml/min/1.73m^2 and/or albuminuria) was 9.1% (95% CI 7.9-10.4). With older age, creatinine increased, and eGFR declined. However, the age-related rise in creatinine was attenuated in women compared to men (Figures S2-S4).

Repeat screening: confirming CKD

To confirm chronicity, thus defining CKD, the study team rescreened participants with eGFR <60ml/min/1.73m^2 (29/32; 91% rescreened) and albuminuria (221/247; 89% rescreened). Follow-up confirmed eGFR<60ml/min/1.73m^2 in 41% (12/29) of participants, and albuminuria in 53% (118/221) of participants. We tabulated CKD prevalence by eGFR stage and ACR stage with associated risk prediction for participants who had complete data for eGFR and albuminuria, and who underwent repeat screening to confirm CKD (N=1975) (Table 2, and stratified by sex in Tables S4, S5). For this group (N=1975), the unadjusted CKD prevalence (a composite of albuminuria, eGFR <60ml/min/1.73m^2, or both) was 6.0% (95% CI 5.0-7.1), and the age-adjusted CKD prevalence was 4.0% (95% CI 3.1-4.9).

Risk factors associated with CKD in Agincourt, South Africa

For those fulfilling the temporal definition of CKD in our study, the age- and sex-adjusted associations with CKD are shown in Table 3. In our final multivariable model, the following risk factors were independently associated with CKD: a high risk *APOL1* genotype (OR 1.95; 95% CI 1.20-3.17), hypertension (OR 3.34; 95% CI 1.92-5.79), diabetes mellitus (OR 3.28 95% CI 1.61-6.70), HIV infection (OR 1.89; 95% CI 1.18-3.03), and hyperuricaemia (OR 1.85; 95% CI 1.13-3.05) (Table 4). We conducted sensitivity analyses using the same modelling approach for impaired kidney function measured with first screening, as albuminuria and /or eGFR <60ml/min/1.73m^2 (Table S6). As seen in those with confirmed CKD, impaired kidney function on first screening was associated with hypertension (OR 2.24; 95% CI 1.52-3.35), diabetes mellitus (OR 3.34 95% CI 1.93-5.78), HIV infection (OR 1.58; 95% CI 121.-2.22), and hyperuricaemia (OR 1.95; 95% CI 1.36-2.81), however the association with high risk *APOL1* genotypes was not significant (OR 1.26; 95% CI 0.85-1.87). We conducted bivariate analyses comparing high vs. low risk *APOL1* genotypes with albuminuria and eGFR as continuous and dichotomous variables for first and second screening (confirmed CKD). Examining the association between high risk *APOL1* genotypes and eGFR or ACR (as continuous variables): there was no significant association on first screening for either eGFR (p=0.29) or ACR (p=0.50); with repeat screening there was still no association with eGFR (p=0.35), however, the association with albuminuria was significant (p=0.001). When stratifying eGFR by albuminuria, the odds of albuminuria with eGFR<90 (compared to eGFR>=90) were 2.39 (95%CI 1.69-3.36); and for eGFR<60 (compared to >=60) were 5.81 (95% CI 2.89-11.68).

Discussion

In our large, prospective population-based study from rural South Africa we confirmed persistent kidney impairment in half our participants at follow-up and the overall age-adjusted prevalence of CKD was 4.0% (95% CI 3.1-4.9). Multivariable analysis showed that prevalent CKD was associated with high risk *APOL1* genotypes, HIV infection, hypertension, diabetes mellitus, and hyperuricaemia.

The CKD risk profile reflects syndemic genetic predisposition and combined burdens of infectious and NCD in a community undergoing rapid epidemiological transition. Although temporal confirmation of impaired kidney function is feasible for an individual in a clinical setting, it is less practical and expensive for large epidemiological studies. Aside from our study, there is only one other African study from Morocco that repeated eGFR and urine albumin measurements to confirm CKD. They showed similar findings - at least one-third of participants with dipstick positive for proteinuria or with low GFR were not confirmed at follow-up.[28]

Recent single-screen population-based studies from ARK study sites in Malawi,[29] Uganda,[30] and the Africa Wits-International Network for the Demographic Evaluation of Populations and their Health Partnership for Genomic Studies (AWI-Gen)[31] reported low GFR rates similar to those observed with the first screen in our study.[31] Collectively, in almost 20 000 individuals, the age-adjusted population prevalence of low GFR was 1.4% (95% CI 1.1 – 1.7%) in Malawi, 1.64% (95% CI 1.34–1.99) in Uganda, and 2.4% (95% CI 2.1–2.8) overall for AWI-GEN countries (Ghana, Burkina Faso, Kenya, South Africa). Similar methodologies in the studies, namely IDMS-traceable creatinine assays, eGFR using the CKD-EPI (creatinine) equation without the ethnicity coefficient, and age-standardization allowed regional comparison. Across the SSA region, and including our study, the prevalence of low GFR was substantially less than expected.[1] For comparison, the NHANES US population survey (1999-2006) reported 6.7% of the population with low GFR, and in the Health Survey for England, 7% of the population over 35 years had low eGFR.[32]

Potential explanations might be that defining CKD using and eGFR cut-off of <60ml/min/1.73m^2 is inappropriate for African populations. With reason, the utility of dichotomous thresholds to define CKD across all age groups is a source of intense debate.[33] In the context of aging populations with a senescent decline in kidney function, especially in the absence of albuminuria, a decline in eGFR to 45ml/min/1.73m^2 might be completely normal. At the risk of overdiagnosing CKD in the elderly, the opposite is true in relatively young populations such as ours, in whom we expect well preserved kidney function. Given younger people with low GFR are more likely to have a potentially treatable

glomerular condition, it might be reasonable to consider an eGFR cut-off <75ml/min/1.73m^2 in the presence of albuminuria as reason to suspect CKD. A threshold of eGFR of <<75ml/min/1.73m^2 has been proposed for those <40 years.[33] Following on from this, another potential explanation for the relatively small population percentage with low GFR might be due to a survivor effect, as those with severe progressive chronic kidney disease demise in the absence of appropriate care.

In addition to low GFR, the AWI-Gen study reported an overall single-screen age-adjusted prevalence of albuminuria of 9.2% (95%CI 8.4–10.0), similar to the prevalence in our population at first screen.[31] However, with repeat screening, we confirmed albuminuria in just over half of our participants, affirming caution against overinterpreting single-screen results to define CKD. Most of the transient albuminuria occurred with well-preserved kidney function (eGFR stages G1-G2) and lower levels of albuminuria.

While the population frequency of *APOL1* high risk genotypes in Agincourt was substantially higher than in an admixed population from the Western Cape,[34] it was similar to frequencies observed in the African diaspora population now resident in the USA, in which the high risk genotype confers a lifetime risk of developing kidney disease of approximately 20%.[8,35] *APOL1*-associated kidney disease requires a "second hit" that activates interferon pathways, as occurs with viral infection or autoimmune disease. *APOL1*, an innate immune factor, is potently upregulated by interferons.[36] Potential "second hits" include the high prevalence of HIV infection,[37] as HIV-associated nephropathy has been strongly associated with the high risk *APOL1* genotype in other parts of South Africa.[9] The association we observed between the high risk *APOL1* genotype and CKD was driven by persistent albuminuria and not reduced eGFR, an observation recently reported from a study of HIV infected adults from Nigeria.[38] While ours was a baseline cross-sectional study, longitudinal studies from relatively young cohorts in the USA with mostly preserved kidney function showed that high risk *APOL1* genotypes predispose to an earlier age of onset of incident albuminuria coincident with a decline in eGFR.[39] A subsequent study found that *APOL1* high risk genotypes are associated with a

greater risk of proteinuria and that the onset of proteinuria was followed by a rapid decline in eGFR in persons with CKD attributed to hypertension.[40]

The association we observed between hyperuricaemia and CKD is the first to be described from a rural South African population. Mechanistically, hyperuricaemia is a pro-inflammatory state, inducing oxidative stress and endothelial dysfunction, with consequent vascular stiffness and vasoconstriction. This explains associations with hypertension, obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease - all known risk factors for CKD.[41] Furthermore, high fructose intake, precipitated by increased ingestion of sugar-sweetened beverages, induces hyperuricaemia and insulin resistance.[42] Studies from the 1970s describe higher rates of hyperuricaemia in urban than rural South Africans,[43,44] yet we observed similar concentrations in men from Agincourt. Paradoxically, women in Agincourt have lower rates of hyperuricaemia despite relatively higher obesity and diabetes, also noted in studies from the North-West province of South Africa.[45] Whether hyperuricaemia is a predictor of cardiovascular, metabolic, or renal outcomes, and whether urate-lowering therapy (which is affordable and readily available) modifies these outcomes has not been proven and remains to be determined in our context.[41,46]

Our baseline study was limited by the absence of longitudinal follow up thus precluding outcomes analyses critical to understanding the impact of CKD on cardiovascular and all-cause mortality in our setting. Additionally, the association with hyperuricaemia and high risk *APOL1* genotypes requires longitudinal follow-up to ascertain the risk for CKD progression. Numerous life course and environmental exposures potentially impact kidney function in our rural communities, and these remain unquantified. While we evaluated traditional risk factors for CKD, we did not explore the impact of endemic and other infections (aside from HIV), exposure to traditional and over-the-counter medicines, and environmental toxins (agricultural and heavy metal). However, our baseline findings and the solid HDSS clinical research platforms in Agincourt and other ARK and AWI-Gen countries create the opportunity for creating a longitudinal CKD cohort in SSA for future work.

Conclusion

In our study, the age-adjusted population prevalence of CKD in rural South Africa was 4.0% (95% CI 3.1-4.9), primarily driven by persistent albuminuria. *APOL1* high risk genotype, hypertension, diabetes, HIV, and hyperuricaemia were associated risk factors associated with CKD. This reflects hereditable risk and combined burdens of disease in a relatively young community undergoing rapid epidemiological transition. Several potential context-specific risk factors for CKD along the life course remain to be explored in vulnerable communities in SSA.

Disclosure Statement

The authors have nothing to disclose.

References

1. Kaze AD, Ilori T, Jaar BG, Echouffo-Tcheugui JB. Burden of chronic kidney disease on the African continent: a systematic review and meta-analysis. BMC nephrology 2018;19:1-11.

2. Ashuntantang G, Osafo C, Olowu WA, et al. Outcomes in adults and children with end-stage kidney disease requiring dialysis in sub-Saharan Africa: a systematic review. The Lancet Global Health 2017;5:e408-e17.

3. Consortium CKDP. Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: a collaborative meta-analysis. The Lancet 2010;375:2073-81.

4. Madala ND, Thusi GP, Assounga AG, Naicker S. Characteristics of South African patients presenting with kidney disease in rural KwaZulu-Natal: a cross sectional study. BMC nephrology 2014;15:61.

5. Adeniyi AB, Laurence CE, Volmink JA, Davids MR. Prevalence of chronic kidney disease and association with cardiovascular risk factors among teachers in Cape Town, South Africa. Clinical Kidney Journal 2017:sfw138.

6. Davids MR, Jardine T, Marais N, Zunza M, Jacobs JC, Sebastian S. South African Renal Registry Annual Report 2017. African Journal of Nephrology 2019;22:60-71.

7. Matsha TE, Yako YY, Rensburg MA, Hassan MS, Kengne AP, Erasmus RT. Chronic kidney diseases in mixed ancestry south African populations: prevalence, determinants and concordance between kidney function estimators. BMC nephrology 2013;14:1.

8. Limou S, Nelson GW, Kopp JB, Winkler CA. APOL1 kidney risk alleles: population genetics and disease associations. Adv Chronic Kidney Dis 2014;21:426-33.

9. Kasembeli AN, Duarte R, Ramsay M, et al. APOL1 risk variants are strongly associated with HIV-associated nephropathy in black South Africans. Journal of the American Society of Nephrology 2015;26:2882-90.

10. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int Suppl 2013;3:1-150.

11. Kahn K, Collinson MA, Gómez-Olivé FX, et al. Profile: Agincourt health and socio-demographic surveillance system. International journal of epidemiology 2012;41:988-1001.

12. Site AHaDS. MRC/Wits-Agincourt Unit HDSS Study Area - Research Villages 20202020.

13. Rohr JK, Xavier Gómez-Olivé F, Rosenberg M, et al. Performance of self-reported HIV status in determining true HIV status among older adults in rural South Africa: a validation study. Journal of the International AIDS Society 2017;20:21691.

14. National Department of Health. National consolidated guidelines for the prevention of mother-to-child transmission of HIV (PMTCT) and the management of HIV in children, adolescents and adults. Pretoria, South Africa: National Department of Health; 2015.

15. David VA, Binns-Roemer EA, Winkler CA. Taqman Assay for Genotyping CKD-Associated APOL1 SNP rs60910145: A Cautionary Note. Kidney Int Rep 2019;4:184-5.

16. World Health Organization. WHO STEPwise approach to surveillance (STEPS). STEPS manual. Geneva: World Health Organization 2017.

17. Kabudula CW, Houle B, Collinson MA, Kahn K, Tollman S, Clark S. Assessing changes in household socioeconomic status in rural South Africa, 2001–2013: a distributional analysis using household asset indicators. Social indicators research 2017;133:1047-73.

18. Alberti K, Eckel RH, Grundy SM, et al. Harmonizing the metabolic syndrome: a joint interim statement of the international diabetes federation task force on epidemiology and prevention; national heart, lung, and blood institute; American heart association; world heart federation; international atherosclerosis society; and international association for the study of obesity. Circulation 2009;120:1640-5.

19. World Health Organisation. Managing the Global Epidemic of Obesity: Report of the World Health Organization Consultation of Obesity. Geneva, June 1997.

20. Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. Jama 2003;289:2560-72.

21. American Diabetes Association. 2. Classification and diagnosis of diabetes: standards of medical care in diabetes—2018. Diabetes Care 2018;41:S13-S27.

22. Klug E, Raal F, Marais A, et al. South African dyslipidaemia guideline consensus statement: 2018 update A joint statement from the South African Heart Association (SA Heart) and the Lipid and Atherosclerosis Society of Southern Africa (LASSA). South African Medical Journal 2018;108:973-1000.

23. Alan H. Tietz clinical guide to laboratory tests. Missouri: WB Saunders Company 2006;1138.

24. World Health Organization. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity: World Health Organization; 2011.

25. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Annals of internal medicine 2009;150:604-12.

26. Ahmad OB, Boschi-Pinto C, Lopez AD, Murray CJ, Lozano R, Inoue M. Age standardization of rates: a new WHO standard. Geneva: World Health Organization 2001;9.

27. Twine R. Public engagement throughout the research process in a health and socio-demographic surveillance system study area in rural South Africa.

28. Gharbi MB, Elseviers M, Zamd M, et al. Chronic kidney disease, hypertension, diabetes, and obesity in the adult population of Morocco: how to avoid “over”-and “under”-diagnosis of CKD. Kidney international 2016;89:1363-71.

29. Nakanga WP, Prynn JE, Banda L, et al. Prevalence of impaired renal function among rural and urban populations: findings of a cross-sectional study in Malawi. Wellcome Open Res 2019;4:92.

30. Kalyesubula R, Hau JP, Asiki G, et al. Impaired renal function in a rural Ugandan population cohort. Wellcome Open Res 2018;3:149.

31. George JA, Brandenburg J-T, Fabian J, et al. Kidney damage and associated risk factors in rural and urban sub-Saharan Africa (AWI-Gen): a cross-sectional population study. The Lancet Global Health 2019;7:e1632-e43.

32. Fat LN, Mindell J, Roderick P. Health Survey for England 2016 Kidney and liver disease. Health and Social Care Information Centre, London 2017.

33. Glassock R, Delanaye P, El Nahas M. An age-calibrated classification of chronic kidney disease. Jama 2015;314:559-60.

34. Matsha TE, Kengne AP, Masconi KL, Yako YY, Erasmus RT. APOL1 genetic variants, chronic kidney diseases and hypertension in mixed ancestry South Africans. BMC genetics 2015;16:69.

35. Yusuf AA, Govender MA, Brandenburg J-T, Winkler CA. Kidney disease and APOL1. Human Molecular Genetics 2021.

36. Nichols B, Jog P, Lee JH, et al. Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1. Kidney international 2015;87:332-42.

37. Gómez-Olivé FX, Angotti N, Houle B, et al. Prevalence of HIV among those 15 and older in rural South Africa. AIDS care 2013;25:1122-8.

38. Wudil UJ, Aliyu MH, Ingles DJ, et al. Apoliproprotein-1 (APOL1) Risk Variants and Associated Kidney Phenotypes in an Adult HIV Cohort in Nigeria.

39. Peralta CA, Bibbins-Domingo K, Vittinghoff E, et al. APOL1 genotype and race differences in incident albuminuria and renal function decline. Journal of the American Society of Nephrology 2016;27:887-93.

40. Freedman BI, Burke W, Divers J, et al. Diagnosis, Education, and Care of Patients with APOL1-Associated Nephropathy: A Delphi Consensus and Systematic Review. Journal of the American Society of Nephrology 2021.

41. Borghi C, Agabiti-Rosei E, Johnson RJ, et al. Hyperuricaemia and gout in cardiovascular, metabolic and kidney disease. European Journal of Internal Medicine 2020.

42. Nakagawa T, Tuttle KR, Short RA, Johnson RJ. Hypothesis: fructose-induced hyperuricemia as a causal mechanism for the epidemic of the metabolic syndrome. Nature clinical practice Nephrology 2005;1:80-6.

43. Beighton P, Solomon L, Soskolne C, Sweet B, Robin G. Serum uric acid concentrations in an urbanized South African Negro population. Annals of the rheumatic diseases 1974;33:442.

44. Beighton P, Solomon L, Soskolne C, Sweet B. Serum uric acid concentrations in a rural Tswana community in Southern Africa. Annals of the rheumatic diseases 1973;32:346.

45. Palmer I, Schutte A, Huisman H, et al. A comparison of uric acid levels in Black African vs Caucasian women from South Africa: the POWIRS study. Ethnicity and disease 2007;17:676.

46. Fernandez-Prado R, Ortiz A, Perez-Gomez MV. Allopurinol and Chronic Kidney Disease. The New England journal of medicine 2020;383:1690-.

47. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics 2009;42:377-81.

48. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. Journal of biomedical informatics 2019;95:103208.

Acknowledgment of Support

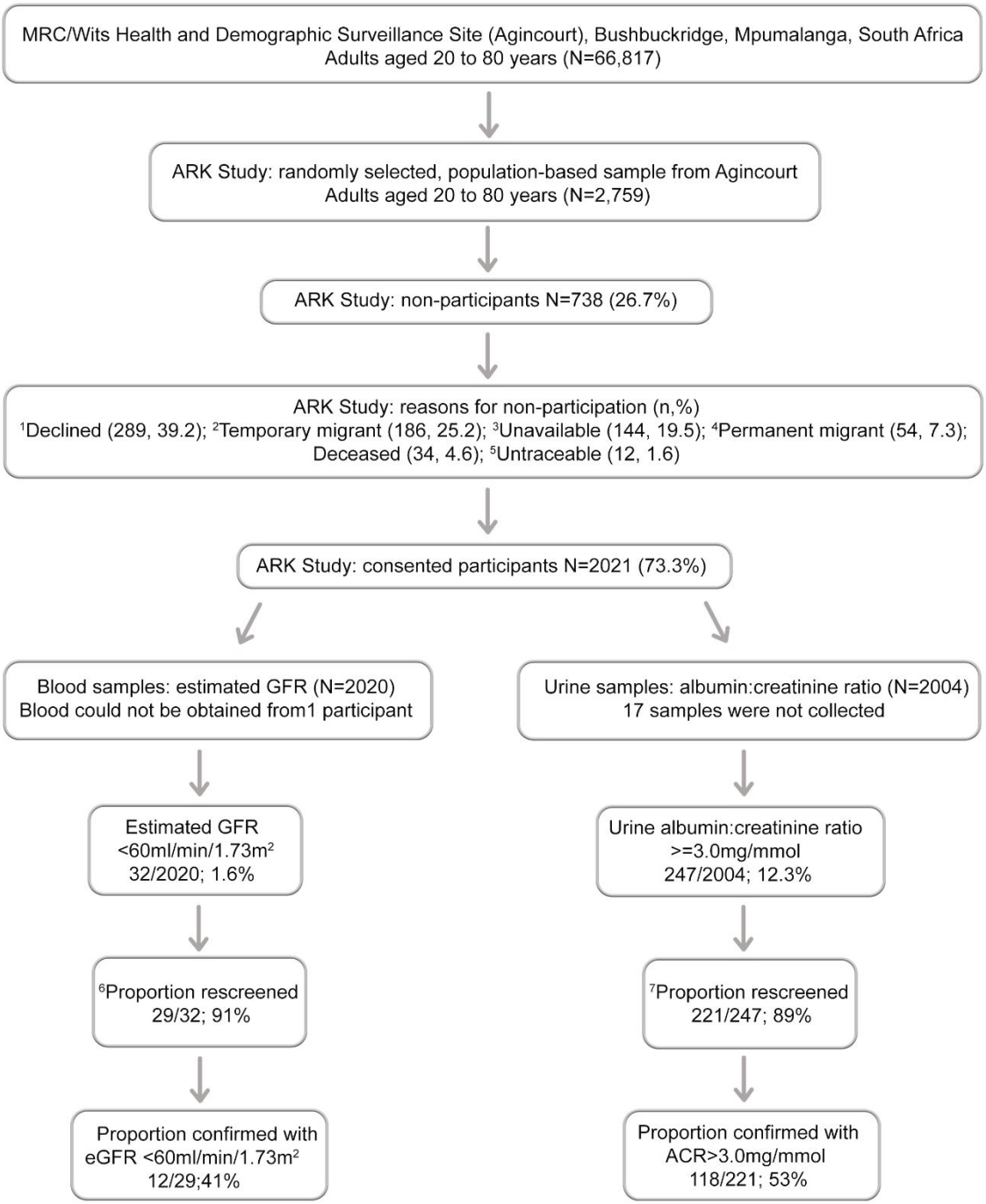
Study data were collected and managed using REDCap electronic data capture tools hosted at the University of the Witwatersrand.[47,48]

Declaration of Funding

The ARK South Africa Study was jointly funded by the South African MRC, MRC UK (via the Newton Fund), and GSK Africa Non-Communicable Disease Open Lab (via a supporting grant Project Number: 074, and the Faculty of Health Sciences Research Incentive Grant; Grant number: 001283843420351211050000000000000004550; University of the Witwatersrand; FHS (Wits).

The MRC/Wits Rural Public Health and Health Transitions Research Unit and Agincourt Health and Socio-Demographic Surveillance System, a node of the South African Population Research Infrastructure Network (SAPRIN), is supported by the Department of Science and Innovation, the University of the Witwatersrand, and the Medical Research Council, South Africa, and previously the Wellcome Trust, UK (grants 058893/Z/99/A; 069683/Z/02/Z; 085477/Z/08/Z; 085477/B/08/Z).

Figure 1: Flow diagram for the ARK Study, Agincourt HDSS site, South Africa



[1]Participants who declined were referred to the study supervisor, who subsequently confirmed non-participation and the reason; [2]Temporary migrants included those who intermittently returned to their homes in the HDSS, having migrated for various reasons, such as employment or post-school education; [3]Fieldworkers categorized potential participants as "unavailable" if they could not locate them after three visits to the homestead; [4]Permanent migrants included those who had relocated their permanent residence outside the HDSS; [5]Fieldworkers categorized potential participants as "untraceable" if they could not trace their whereabouts after three attempts; [6] Rescreened GFR median duration between first and second visit was 117 days (IQR 80-161); [7] Rescreened Albuminuria median duration between first and second visit was 193 days (IQR 123-244).

Table 1: Characteristics of participants in the ARK Study, Agincourt, South Africa

Variable	Men, n (%) (n=851)	Women, n (%) (n=1170)	Total*, n (%) (n=2021)
Mean age (±SD) years	36.5 (±13.4)	39.0 (±14.5)	37.9 (±14.1)
Mean serum creatinine (µmol/L)	73.4 (15.5)	56.2 (12.9)	63.5 (16.4)
Mean estimated GFR (ml/min/1.73m^2)	112 (17)	112 (19)	112 (18)
Age group (years)			
20-29	327 (38.43)	376 (32.14)	703 (34.78)
30-39	230 (27.03)	329 (28.12)	559 (27.66)
40-49	149 (17.51)	210 (17.95)	359 (17.76)
50-59	78 (9.17)	119 (10.17)	197 (9.75)
60-69	48 (5.64)	80 (6.84)	128 (6.33)
70-79	19 (2.23)	56 (4.79)	75 (3.71)
Household assets score (quintiles)	**(n=718)**	**(n=1003)**	**(n=1721)**
Quartile 1 (poorest)	212 (29.53)	299 (29.81)	511 (29.69)
Quartile 2	200 (27.86)	279 (27.82)	479 (27.83)
Quartile 3	161 (22.42)	273 (27.22)	434 (25.22)
Quartile 4 (wealthiest)	145 (20.19)	152 (15.15)	297 (17.26)
Highest education level	**(n=845)**	**(n=1157)**	**(n=2002)**
No formal education	47 (5.56)	135 (11.67)	182 (9.09)
Less than six years completed	65 (7.69)	94 (8.12)	159 (7.94)
Six or more years completed	733 (86.75)	928 (80.21)	1661 (82.97)
Family history	**(n=851)**	**(n=1169)**	**(n=2020)**
Hypertension	260 (30.55)	467 (39.95)	727 (35.99)
Diabetes mellitus	95 (11.16)	191 (16.34)	286 (14.16)
Chronic kidney disease	20 (2.35)	60 (5.13)	80 (3.96)
Current smoking			
Yes	223 (26.20)	6 (0.51)	229 (11.33)
[1]***APOL1* risk genotypes**	**(n=844)**	**(n=1158)**	**(n=2002)**
Low risk	745 (88.27)	1031 (89.03)	1776 (88.71)
High risk	99 (11.73)	127(10.97)	226 (11.29)
[2]Body mass index	**(n=850)**	**(n=1114)**	**(n=1964)**
Normal weight	493 (58.00)	277 (24.87)	770 (39.21)
Underweight	55 (6.47)	28 (2.51)	83 (4.23)
Overweight	206 (24.24)	365 (32.77)	571 (29.07)
Obese	96 (11.29)	444 (39.86)	540 (27.50)
Obese class I	*67 (7.88)*	*243 (21.81)*	*310 (57.41)*
Obese class II	*19 (2.24)*	*132 (11.85)*	*151 (27.96)*
Obese class III	*10 (1.18)*	*69 (6.19)*	*79 (14.63)*
[3]Waist circumference	**(n=851)**	**(n=1117)**	**(n=1968)**
Increased	170 (19.98)	861 (77.08)	1031 (52.39)

[4]Waist-hip ratio	(n=851)	(n=1117)	(n=1968)
Increased	261 (30.67)	541 (48.43)	802 (40.75)
[5]Blood pressure			
Normal	212 (24.91)	543 (46.41)	755 (37.36)
Pre-hypertension	452 (53.11)	445 (38.03)	897 (44.38)
Hypertension	187 (21.97)	182 (15.56)	369 (18.26)
[6]Diabetes mellitus			
Present	21 (2.47)	50 (4.27)	71 (3.51)
[7]Hypercholesterolaemia			
Present	144 (16.92)	321 (27.44)	465 (23.01)
[8]Anaemia			
Present	170 (19.98)	397 (33.93)	567 (28.05)
[9]Hyperuricaemia	**(n=851)**	**(n=1169)**	**(n=2020)**
Present	150 (17.63)	111 (9.50)	261 (12.92)
Hepatitis B infection	**(n=851)**	**(n=1169)**	**(n=2020)**
Positive	40 (4.70)	40 (3.42)	80 (3.96)
HIV infection			
Positive	102 (11.99)	278 (23.76)	380 (18.8)

*Column total percentages may not sum to 100 due to rounding; categories presented as frequency (%).

[1]Apolipoprotein L1 (APOL1) genotyping: High risk genotypes were defined as those containing two high risk alleles (G1/G1, G1/G2, or G2/G2); low risk genotypes were defined as those containing zero or one risk allele.[15]

[2]BMI (kg/(height)^2): underweight (<18.5); normal weight (18.5-24.99); overweight (25.0-29.99); obese (≥30.0): obese class I (30.0-34.99); obese class II (35.0-39.99); obese class III (>=40.00).[19]

[3]Increased waist circumference: >=80cm for women; >=94cm for men.[18]

[4]Increased waist to hip ratio: >0.85 for women; >1.0 for men.[18]

[5]Classified according to the 7th Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure as: normotensive (systolic SBP<120mmHg and DBP<80mmHg); prehypertensive (SBP≥120mmHg and <140mmHg or DBP≥80mmHg and <90mmHg); or hypertensive (SBP≥140mmHg or DBP≥90mmHg).[20]

[6]Diabetes mellitus: non-fasting glucose >=11.1mmol/L.[21]

[7]Hypercholesterolaemia: non-fasting total cholesterol >5.0mmol/L.[22]

[8]Anaemia: haemoglobin (Hb) <11.0g/dL in pregnant women; <12.0 g/dL in non-pregnant women and <13.0 g/dL in men.[24]

[9]Hyperuricaemia: uric acid >0.43mmol/L in men; >0.36mmol/L in women.[23]

Table 2: KDIGO CKD and risk for adverse outcomes[10]: All participants in the ARK Study, Agincourt, South Africa[1]

low risk / moderate risk / high risk / very high risk		ACR (mg/mmol)			
		<3	3-30	>30	Total
		A1	A2	A3	
>=90	[4]G1	1755 [2]99.94% [3]94.51%	0	1 0.06% 7.69%	1,756 100.00% 88.91%
60-89	G2	90 43.48% 4.85%	105 50.72% 100.00%	12 5.80% 92.31%	207 100.00% 10.48%
45-59	G3a	9 100.00% 0.48%	0	0	9 100.00% 0.46%
30-44	G3b	2 100.00% 0.11%	0	0	2 100.00% 0.10%
15-29	G4	1 100.00% 0.05%	0	0	1 100.00% 0.05%
<15	G5	0	0	0	0
Total		1857 94.03% 100.00%	105 5.32% 100.00%	13 0.66% 100.00%	1,975 % %

[1]Only including participants with complete data for eGFR and ACR, and fulfilling temporal requirements for the definition of CKD with follow up screening for low eGFR and / or albuminuria (N=1,975); [2]*Row percentage*; [3]*column percentage*; [4]eGFR calculated using CKD-EPI (creatinine) without ethnicity coefficient (ml/min/1.73m^2)

Table 3: Age and sex-adjusted risk factors independently associated with CKD in Agincourt, South Africa*

Variable	Total participants n (%)	Participants with CKD n (%)	Unadjusted OR (95% CI)	Age and sex adjusted OR (95% CI)
Sex			P=0.60	P=0.49
Male	851 (42.13)	55 (6.46)	*reference*	*reference*
Female	1169 (57.87)	69 (5.90)	0.91 (0.63-1.31)	0.88 (0.61-1.27)
Age group (years)			P=0.04	P=0.03
20-29	702 (34.75)	41 (5.84)	*reference*	*reference*
30-39	559 (27.67)	29 (5.19)	0.88 (0.54-1.44)	0.89 (0.55-1.45)
40-49	359 (17.77)	19 (5.29)	0.90 (0.52-1.58)	0.91 (0.52-1.59)
50-59	197 (9.75)	15 (7.61)	1.33 (0.72-2.46)	1.34 (0.73-2.48)
60-69	128 (6.34)	11 (8.59)	1.52 (0.76-3.03)	1.54 (0.77-3.08)
70-79	75 (3.71)	9 (12.00)	2.20 (1.02-4.72)	2.26 (1.05-4.89)
Household assets score			P=0.54	P=0.54
Quartile 1	511 (29.69)	29 (5.68)	*reference*	*reference*
Quartile 2	479 (27.83)	33 (6.89)	1.23 (0.74-2.06)	1.24 (0.74-2.08)
Quartile 3	434 (25.22)	25 (5.76)	1.02 (0.59-1.76)	1.02 (0.59-1.77)
Quartile 4	297 (17.26)	14 (4.71)	0.82 (0.43-1.58)	0.82(0.43-1.58)
Highest education level			P=0.09	P=0.48
No education	182 (9.10)	14 (7.69)	*reference*	*reference*
Less than six years	159 (7.95)	15 (9.43)	1.25 (0.58-2.68)	1.29 (0.60-2.78)
Six or more years	1660 (82.96)	93 (5.60)	0.71 (0.40-1.28)	0.84 (0.42-1.70)
Family history hypertension			P=0.75	P=0.72
No	1293 (64.01)	81 (6.26)	*reference*	*reference*
Yes	727 (35.99)	43 (5.91)	0.94 (0.64-1.38)	0.93 (0.63-1.37)
Family history diabetes			P=0.03	P=0.04
No	1734 (85.84)	98 (5.65)	*reference*	*reference*
Yes	286 (14.16)	26 (9.09)	1.67 (1.06-2.62)	1.62 (1.03-2.56)
Family history CKD			P=0.97	P=0.88
No	1940 (96.04)	119 (6.13)	*reference*	*reference*
Yes	80 (3.96)	5 (6.25)	1.02 (0.41-2.57)	1.08 (0.43-2.72)
Current smoker			P=0.99	P=0.82
No	1791 (88.66)	110 (6.14)	*reference*	*reference*
Yes	229 (11.34)	14 (6.11)	1.00 (0.56-1.77)	0.93 (0.50-1.73)
[1]*APOL1* risk genotypes			P=0.003	P=0.003
Low risk	1776 (88.71)	99 (5.57)	*reference*	*reference*
High risk	226 (11.29)	24 (10.62)	2.01 (1.26-3.22)	2.05 (1.28-3.28)
[2]Body mass index			P=0.51	P=0.46
Normal weight	785 (38.94)	45 (5.73)	*reference*	*reference*

Underweight	83 (4.12)	9 (10.84)	2.00 (0.94-4.25)	1.95 (0.91-4.15)
Overweight	585 (29.02)	36 (6.15)	1.08 (0.69-1.70)	1.05 (0.66-1.68)
Obese	563 (27.93)	33 (5.86)	1.02 (0.65-1.63)	1.00 (0.61-1.66)
[3]Blood pressure			P<0.001	P<0.001
Normal	754 (37.33)	33 (4.38)	*reference*	*reference*
Pre-hypertension	897 (44.41)	47 (5.24)	1.21 (0.77-1.91)	1.19 (0.74-1.89)
Hypertension	369 (18.27)	44 (11.92)	2.96 (1.85-4.73)	2.79 (1.69-4.62)
[4]Diabetes mellitus			P<0.001	P<0.001
No	1949 (96.49)	110 (5.64)	*reference*	*reference*
Yes	71 (3.51)	14 (19.72)	4.11 (2.22-7.60)	3.70 (1.91-7.16)
[5]Hyperuricaemia			P=0.001	P=0.007
No	1759 (87.08)	96 (5.46)	*reference*	*reference*
Yes	261 (12.92)	28 (10.73)	2.08 (1.34-3.24)	1.90 (1.20-3.01)
Hepatitis B infection			P=0.60	P=0.72
No	1940 (96.04)	118 (6.08)	*reference*	*reference*
Yes	80 (3.96)	6 (7.50)	1.25 (0.53-2.94)	1.17 (0.50-2.75)
HIV infection			P=0.12	P=0.13
No/unknown	1640 (81.19)	94 (5.73)	*reference*	*reference*
Yes	380 (18.81)	30 (7.89)	1.41 (0.92-2.16)	1.40 (0.91-2.16)

*Column total percentages may not sum to 100 due to rounding; categories presented as frequency (%).

[1]Apolipoprotein L1 (APOL1) genotyping: High risk genotypes were defined as those containing two high risk alleles (G1/G1, G1/G2, or G2/G2); low risk genotypes were defined as those containing zero or one risk allele.[15]

[2]BMI (kg/(height)^2): underweight (<18.5); normal weight (18.5-24.99); overweight (25.0-29.99); obese (≥30.0): obese class I (30.0-34.99); obese class II (35.0-39.99); obese class III (>=40.00).[19]

[3]Classified according to the 7th Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure as: normotensive (systolic SBP<120mmHg and DBP<80mmHg); prehypertensive (SBP≥120mmHg and <140mmHg or DBP≥80mmHg and <90mmHg); or hypertensive (SBP≥140mmHg or DBP≥90mmHg)[20].

[4]Diabetes mellitus: non-fasting glucose >=11.1mmol/L.[21]

[5]Hyperuricaemia: uric acid >0.43mmol/L in men; >0.36mmol/L in women.[23]

Table 4: Phenotypic, genotypic, infectious and metabolic risk factors associated with CKD In Agincourt, South Africa

Variable	Age and sex adjusted OR (95% CI)
Sex	P=0.73
Men	*reference*
Women	1.09 (0.68-1.72)
Age group (years)	P=0.63
20-29	*reference*
30-39	0.77 (0.46-1.30)
40-49	0.61 (0.33-1.13)
50-59	0.76 (0.39-1.49)
60-69	0.78 (0.36-1.69)
70-79	0.92 (0.38-2.23)
Current smoker	P=0.78
No	*reference*
Yes	0.91 (0.48-1.74)
[1]***APOL1*** **risk genotypes**	P=0.007
Low risk	*reference*
High risk	1.95 (1.20-3.17)
[2]**Body Mass Index**	P=0.022
Normal	*reference*
Underweight	2.11 (0.97-4.59)
Overweight	0.93 (0.57-1.51)
Obese	0.69 (0.39-1.19)
[3]**Blood pressure**	P<0.001
Normal	*reference*
Pre-hypertension	1.31 (0.80-2.14)
Hypertension	3.34 (1.92-5.79)
[4]**Diabetes Mellitus**	P=0.001
No	*reference*
Yes	3.28 (1.61-6.70)
HIV infection	P=0.008
No	*reference*
Yes	1.89 (1.18-3.03)
Hepatitis B infection	P=0.43
No	*reference*

Yes	1.43 (0.59-3.44)
[5]Hyperuricaemia	P=0.015
No	*reference*
Yes	1.85 (1.13-3.05)

*Column total percentages may not sum to 100 due to rounding; categories presented as frequency (%).

[1] Apolipoprotein L1 (APOL1) genotyping: High risk genotypes were defined as those containing two high risk alleles (G1/G1, G1/G2, or G2/G2); low risk genotypes were defined as those containing zero or one risk allele.[15] High risk *APOL1* genotypes were associated with albuminuria, not low GFR.

[2]BMI (kg/(height)^2): underweight (<18.5); normal weight (18.5-24.99); overweight (25.0-29.99); obese (≥30.0): obese class I (30.0-34.99); obese class II (35.0-39.99); obese class III (>=40.00).[19]

[3]Classified according to the 7th Report of the Joint National Committee on Prevention, Detection, Evaluation and Treatment of High Blood Pressure as: normotensive (systolic SBP<120mmHg and DBP<80mmHg); prehypertensive (SBP≥120mmHg and <140mmHg or DBP≥80mmHg and <90mmHg); or hypertensive (SBP≥140mmHg or DBP≥90mmHg)[20].

[4]Diabetes mellitus: non-fasting glucose >=11.1mmol/L.[21]

[5]Hyperuricaemia: uric acid >0.43mmol/L in men; >0.36mmol/L in women.[23]

Supplementary Appendix

For the paper titled: Prevalence and associated risk factors for chronic kidney disease in a rural South African setting

Table of contents

***Figure S1*: Comparison of the ARK Study Sample to the Agincourt HDSS population by age, sex**

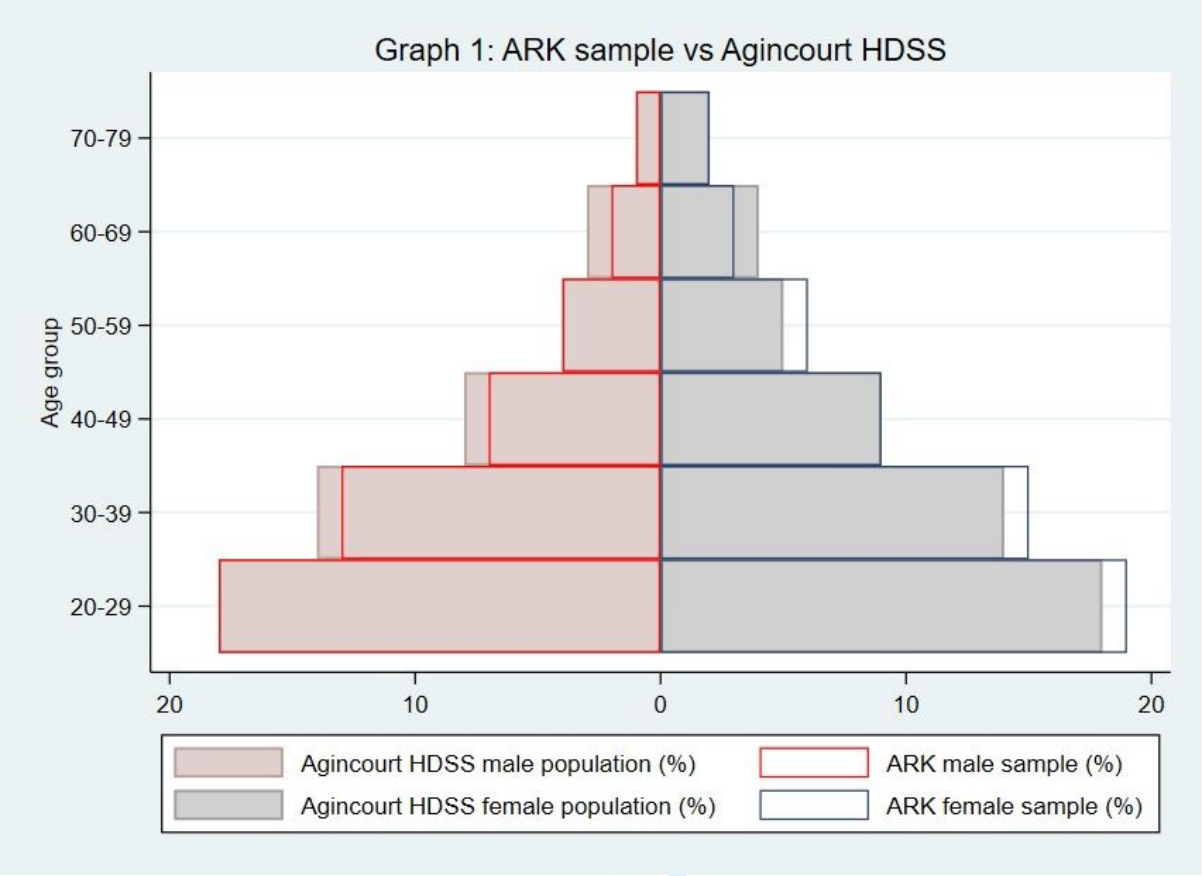


***Figure S2*: Box plots of serum creatinine by age category (years) and sex**

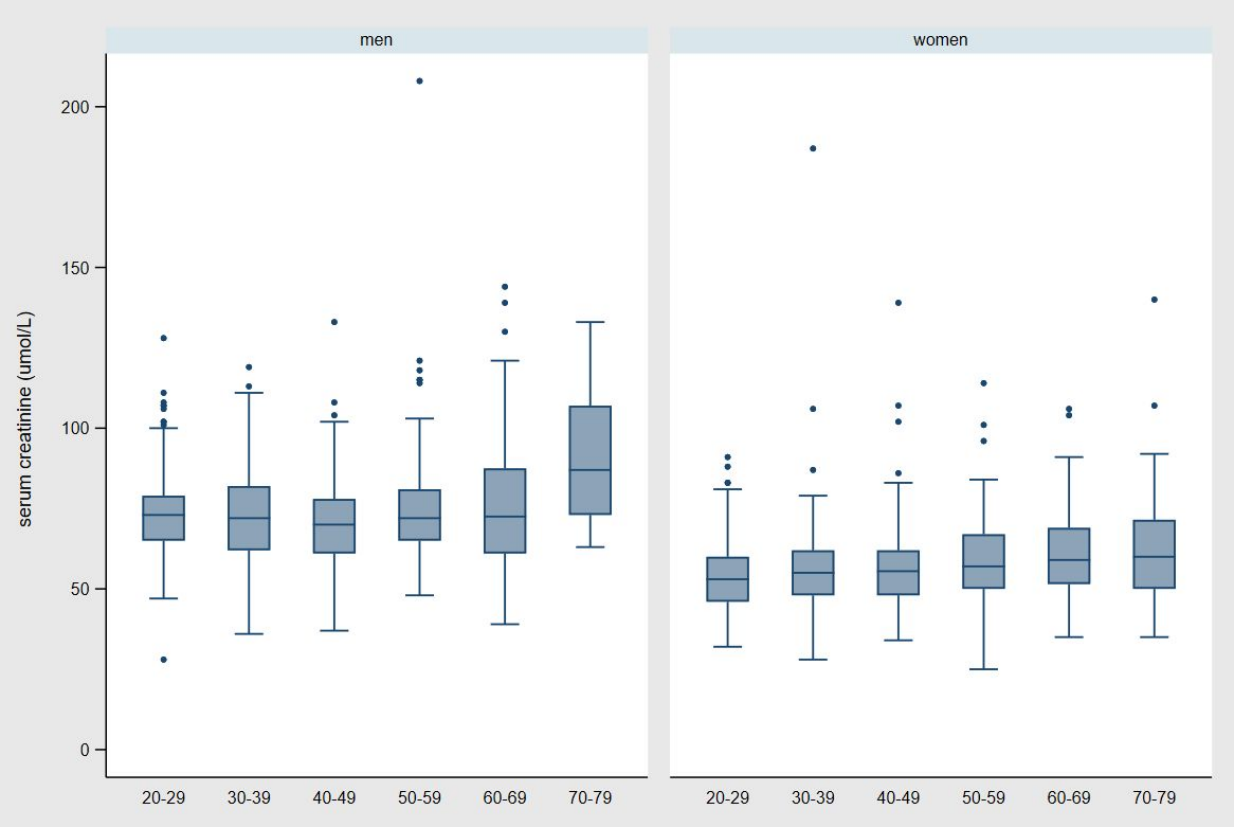


***Figure S3*: eGFR (CKD-EPI creatinine, no ethnicity coefficient) by age (years) and sex**

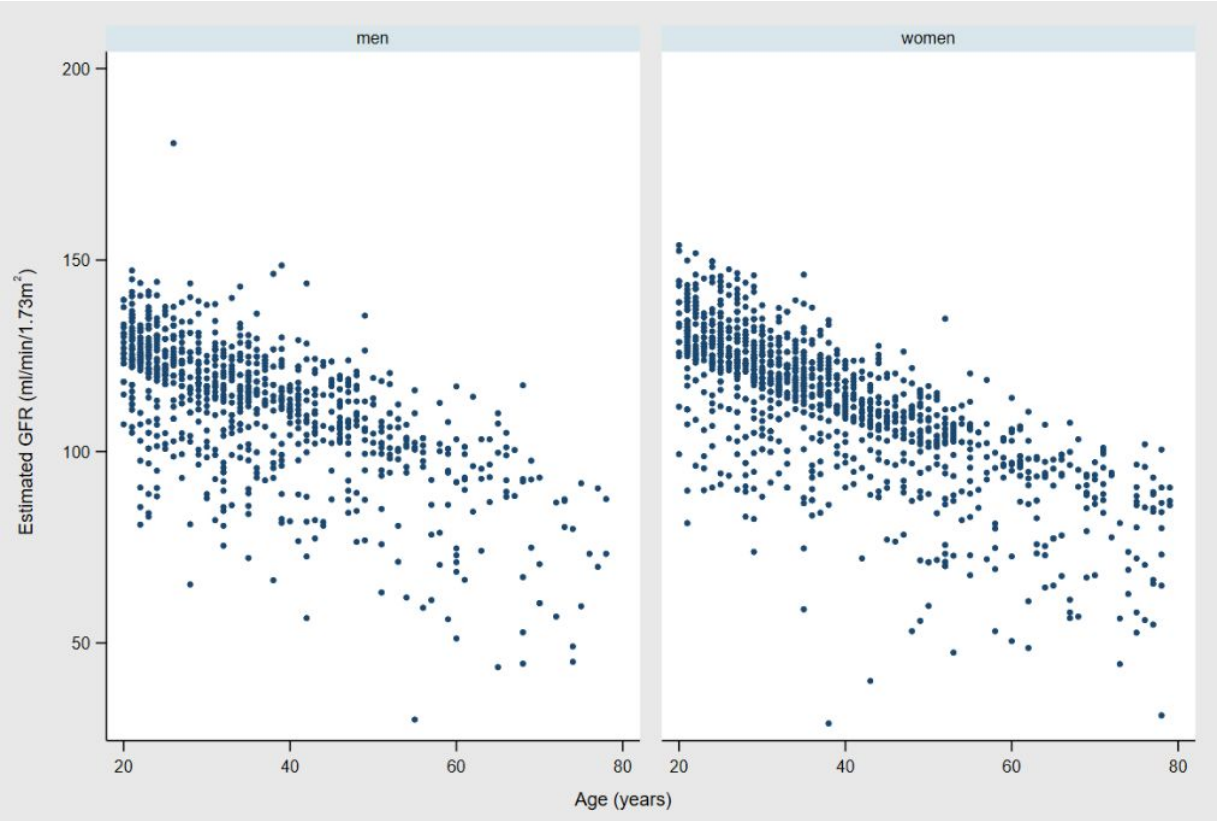


***Figure S4*: eGFR stages (G1, G2, G3-5) by age category (years) and sex**

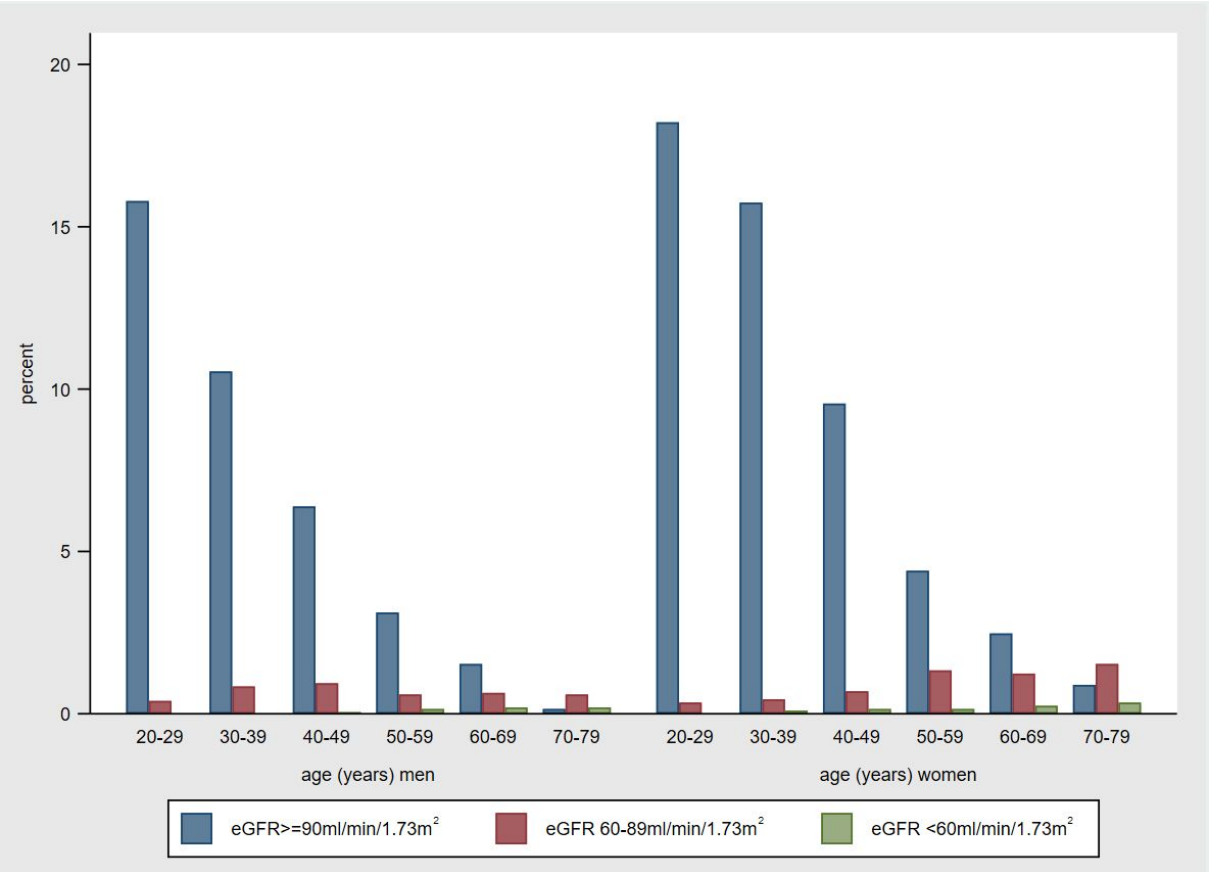
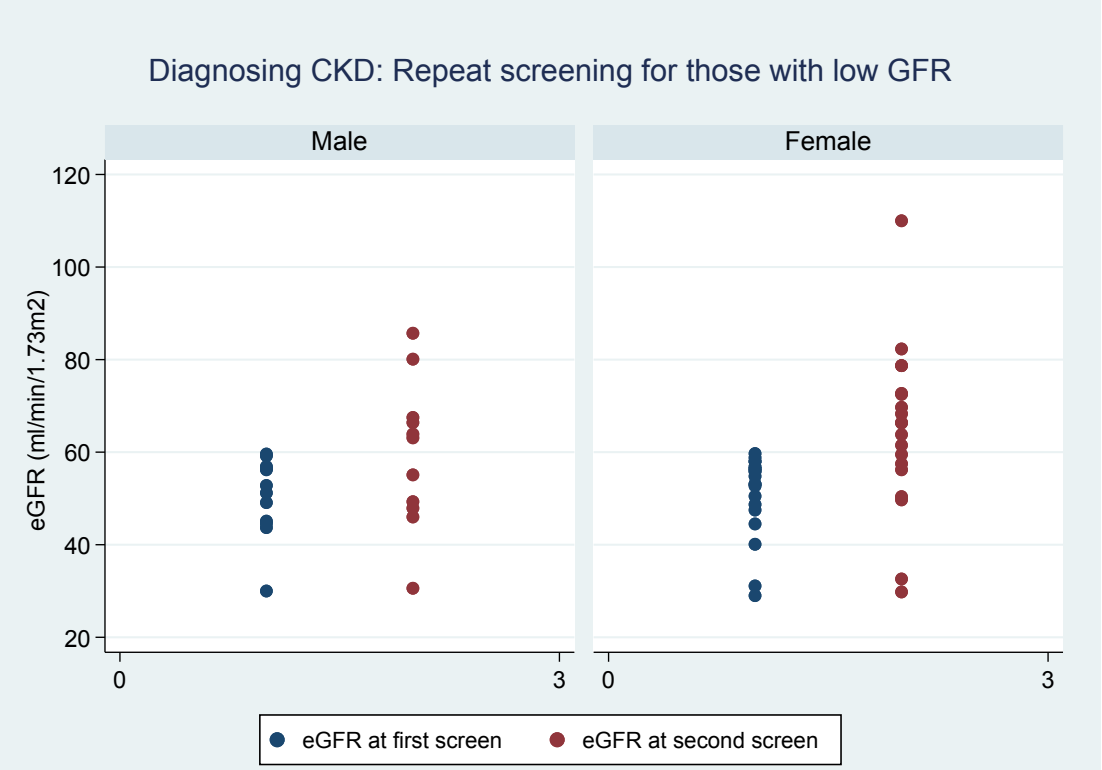


***Figure S5*: Comparison of the first and second screening for low GFR and albuminuria in ARK participants, stratified by sex**



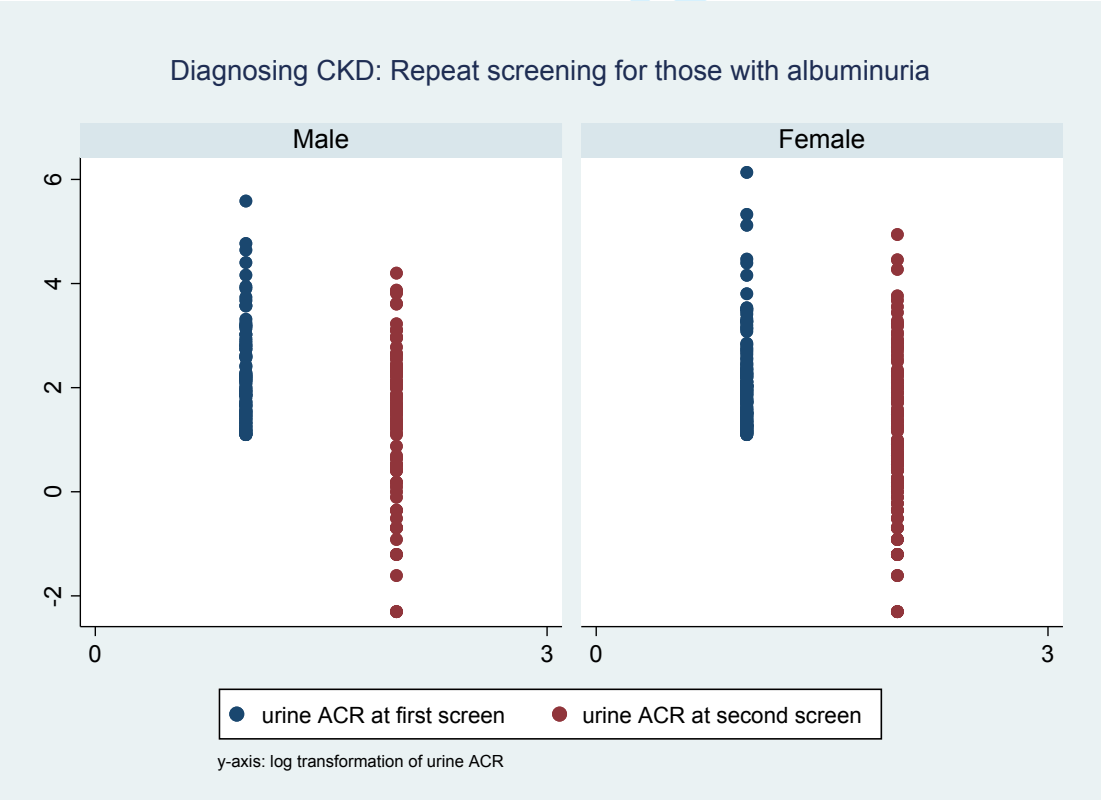


Table S1: First screen: Kidney function for all participants in the ARK Study, Agincourt, South Africa*

ACR = albumin: creatinine ratio (mg/mmol) G1-5: estimated GFR (ml/min/1.73m²)		<3 A1	3-30 A2	>30 A3	Total
>=90	[4]G1	1,584 [2]88.99% [3]90.15%	182 10.22% 81.25%	14 0.79% 60.87%	1,780 100.00% 88.82%
60-89	G2	155 80.73% 8.82%	30 15.63% 13.39%	7 3.65% 30.43%	192 100.00% 9.58%
45-59	G3a	17 68.00% 0.97%	8 32.00% 3.57%	0	25 100.00% 1.25%
30-44	G3b	1 16.67% 0.06%	4 66.67% 1.79%	1 16.67% 4.35%	6 100.00% 0.30%
15-29	G4	0	0	1 100.00% 4.35%	1 100.00% 0.05%
<15	G5	0	0	0	0
Total		1,757 87.67% 100.00%	224 11.18% 100.00%	23 1.15% 100.00%	2004 100.00% 100.00%

*Only including participants with complete data for eGFR and ACR (N=2004); [2]*Row percentage*; [3]*column percentage*; [4]eGFR calculated using CKD-EPI(creatinine) without ethnicity coefficient

Table S2: First screen: Kidney function for women in the ARK Study, Agincourt, South Africa*

ACR = albumin: creatinine ratio (mg/mmol) G1-5: estimated GFR (ml/min/1.73m²)		<3 A1	3-30 A2	>30 A3	Total
>=90	[4]G1	913 [2]88.90% [3]90.40%	107 10.42% 78.10%	7 0.68% 58.33%	1,027 100.00% 88.61%
60-89	G2	86 76.79% 8.51%	23 20.54% 16.79%	3 2.68% 25.00%	112 100.00% 9.66%
45-59	G3a	11 68.75% 1.09%	5 31.25% 3.65%	0	16 100.00% 1.38%
30-44	G3b	0	2 66.67% 1.46%	1 33.33% 8.33%	3 100.00% 0.26%
15-29	G4	0	0	1 100.00% 8.33%	1 100.00% 0.09%
<15	G5	0	0	0	0
Total		1,010 87.14% 100.00%	137 11.82% 100.00%	12 1.04% 100.00%	1,159 100.00% 100.00%

*Only including participants with complete data for eGFR and ACR (N=1159); [2]*Row percentage*; [3]*column percentage*; [4]eGFR calculated using CKD-EPI(creatinine) without ethnicity coefficient

Table S3: First screen: Kidney function for men in the ARK Study, Agincourt, South Africa*

ACR = albumin: creatinine ratio (mg/mmol) G1-5: estimated GFR (ml/min/1.73m²)		<3 A1	3-30 A2	>30 A3	Total
>=90	[4]G1	671 [2]89.11% [3]89.83%	75 9.96% 86.21%	7 0.93% 63.64%	753 100.00% 89.11%
60-89	G2	69 86.25% 9.24%	7 8.75% 8.05%	4 5.00% 36.36%	80 100.00% 9.47%
45-59	G3a	6 66.67% 0.80%	3 33.33% 3.45%	0	9 100.00% 1.07%
30-44	G3b	1 33.33% 0.13%	2 66.67% 2.30%	0	3 100.00% 0.36%
15-29	G4	0	0	0	0
<15	G5	0	0	0	0
Total		747 88.40% 100.00%	87 10.30% 100.00%	11 1.30% 100.00%	845 100.00% 100.00%

*Only including participants with complete data for eGFR and ACR (N=845); [2]*Row percentage*; [3]*column percentage*; [4]eGFR calculated using CKD-EPI(creatinine) without ethnicity coefficient

Table S4: KDIGO[1] CKD and risk for adverse outcomes: Women in the ARK Study, Agincourt, South Africa*

low risk / moderate risk / high risk / very high risk		ACR (mg/mmol)			
		<3	3-30	>30	Total
		A1	A2	A3	
>=90	[4]G1	1,017 [2]100.00% [3]94.17%	0	0	1,017 100.00% 88.74%
60-89	G2	56 45.90% 5.19%	57 46.72% 100.00%	9 7.38% 100.00%	122 100.00% 10.65%
45-59	G3a	5 100.00% 0.46%	0	0	5 100.00% 0.44%
30-44	G3b	1 100.00% 0.09%	0	0	1 100.00% 0.09%
15-29	G4	1 100.00% 0.09%	0	0	1 100.00% 0.09%
<15	G5	0	0	0	0
Total		1,080 94.24% 100.00%	57 4.97% 100.00%	9 0.79% 100.00%	1,146 100.00% 100.00%

*Only including participants with complete data for eGFR and ACR, and fulfilling temporal requirements for the definition of CKD with follow up screening for low eGFR and / or albuminuria (N=1146); [2]*Row percentage*; [3]*column percentage*; [4]eGFR calculated using CKD-EPI(creatinine) without ethnicity coefficient (ml/min/1.73m^2)

Table S5: KDIGO[1] CKD and risk for adverse outcomes: Men in the ARK Study, Agincourt, South Africa*

low risk / moderate risk / high risk / very high risk		ACR (mg/mmol)			
		<3	3-30	>30	Total
		A1	A2	A3	
>=90	[4]G1	738 [2]99.86% [3]94.98%	0	1 0.14% 25.00%	739 100.00% 89.14%
60-89	G2	34 40.00% 4.38%	48 56.47% 100.00%	3 3.53% 75.00%	85 100.00% 10.25%
45-59	G3a	4 100.00% 0.51%	0	0	4 100.00% 0.48%
30-44	G3b	1 100.00% 0.13%	0	0	1 100.00% 0.12%
15-29	G4	0	0	0	0
<15	G5	0	0	0	0
Total		777 93.73% 100.00%	48 5.79% 100.00%	4 0.48% 100.00%	829 100.00% 100.00%

*Only including participants with complete data for eGFR and ACR, and fulfilling temporal requirements for the definition of CKD with follow up screening for low eGFR and / or albuminuria (N=829); [2]*Row percentage*; [3]*column percentage*; [4]eGFR calculated using CKD-EPI(creatinine) without ethnicity coefficient (ml/min/1.73m^2)

Table S6: Phenotypic, genotypic, infectious and metabolic risk factors associated with eGFR <60ml/min/1.73m^2 or albuminuria on first screen in the Agincourt HDSS population

Variable	Age and sex adjusted OR (95% CI)
Sex	P=0.237
Male	*reference*
Female	1.22 (0.88-1.70)
Age group (years)	P=0.44
20-29	*reference*
30-39	0.72 (0.49-1.04)
40-49	0.74 (0.49-1.13)
50-59	1.02 (0.64-1.62)
60-69	0.87 (0.50-1.53)
70-79	1.38 (0.74-2.56)
Current smoker	0.65
No	*reference*
Yes	0.90 (0.55-1.45)
APOL1 risk genotypes	P=0.26
Low risk	*reference*
High risk	1.26 (0.85-1.87)
Body Mass Index	P=0.049
Normal	*reference*
Underweight	2.13 (1.18-3.85)
Overweight	1.01 (0.71-1.43)
Obese	0.90 (0.61-1.32)
[6] **Blood pressure**	P<0.001
Normal	*reference*
Pre-hypertension	1.25 (0.90-1.75)
Hypertension	2.25 (1.52-3.35)
[7]**Diabetes Mellitus**	P<0.001
No	*reference*
Yes	3.34 (1.93-5.78)
HIV infection	P=0.01
No	*reference*
Yes	1.58 (1.12-2.22)
Hepatitis B infection	P=0.50
No	*reference*
Yes	1.25 (0.65-2.40)
Hyperuricaemia	P<0.001
No	*reference*
Yes	1.95 (1.36-2.81)

[1]albuminuria defined as albumin:creatinine ratio (ACR) >=3.0mg/mmol (N=2004)

[2]estimated glomerular filtration rate (eGFR) <60ml/min/1.73m^2 (using the CKD-EPI equation without adjusting for African American ethnicity) (N=2020)

References

1. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int Suppl 2013;3:1-150.

R14/49 «Tit init name»

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M170583

NAME: (Principal Investigator)	Dr June Fabian et al
DEPARTMENT:	Internal Medicine Wits Donald Gordon Medical Centre
PROJECT TITLE:	Determining the Prevalence of Chronic Kidney Disease (CKD) in a Rural Population in South Africa
DATE CONSIDERED:	Adhoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-Study under Primary Studies M1403107, M160939, M160937 and M160939 Principal Investigator Prof Saraladevi Naicker et al
SUPERVISOR:	
APPROVED BY:	Professor P. Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	05/06/2017

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand. I/We fully understand the conditions under which I am/we are authorised to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. in this case, the study was initially review in May and will therefore be due in the month of May each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Reporting checklist for cross sectional study.

Based on the STROBE cross sectional guidelines.

Instructions to authors

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cross sectionalreporting guidelines, and cite them as:

von Elm E, Altman DG, Egger M, Pocock SJ, Gotzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational studies.

		Reporting Item	Page Number
Title and abstract			
Title	#1a	Indicate the study's design with a commonly used term in the title or the abstract	3

Abstract	#1b	Provide in the abstract an informative and balanced summary of what was done and what was found	3-4
Introduction			
Background / rationale	#2	Explain the scientific background and rationale for the investigation being reported	5-6
Objectives	#3	State specific objectives, including any prespecified hypotheses	6
Methods			
Study design	#4	Present key elements of study design early in the paper	6
Setting	#5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	6-8
Eligibility criteria	#6a	Give the eligibility criteria, and the sources and methods of selection of participants.	6
	#7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	9-10
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. Give information separately for for exposed and unexposed groups if applicable.	6-10

Bias	#9	Describe any efforts to address potential sources of bias	10
Study size	#10	Explain how the study size was arrived at	6
Quantitative variables	#11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen, and why	10
Statistical methods	#12a	Describe all statistical methods, including those used to control for confounding	n/a
Statistical methods	#12b	Describe any methods used to examine subgroups and interactions	n/a
Statistical methods	#12c	Explain how missing data were addressed	11
Statistical methods	#12d	If applicable, describe analytical methods taking account of sampling strategy	n/a
Statistical methods	#12e	Describe any sensitivity analyses	10-11
Results			
Participants	#13a	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. Give information separately for for exposed and unexposed groups if applicable.	11-12, 24, 25-31
Participants	#13b	Give reasons for non-participation at each stage	24

Participants	#13c	Consider use of a flow diagram	24
Descriptive data	#14a	Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.	25-26
Descriptive data	#14b	Indicate number of participants with missing data for each variable of interest	24, 25-31
Outcome data	#15	Report numbers of outcome events or summary measures. Give information separately for exposed and unexposed groups if applicable.	27-31
Main results	#16a	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	12-13, 28-31
Main results	#16b	Report category boundaries when continuous variables were categorized	8-10, 25-26, 28-31
Main results	#16c	If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	n/a
Other analyses	#17	Report other analyses done—e.g., analyses of subgroups and interactions, and sensitivity analyses	10-11
Discussion			
Key results	#18	Summarise key results with reference to study objectives	12

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.	15
Interpretation	#20	Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.	15
Generalisability	#21	Discuss the generalisability (external validity) of the study results	12-15
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	22

Clinical Kidney Journal, 2019, vol. 12, no. 6, 778–787

doi: 10.1093/ckj/sfz089
Advance Access Publication Date: 19 August 2019
CKJ Review

CKJ REVIEW

Methods and reporting of kidney function: a systematic review of studies from sub-Saharan Africa

June Fabian[1,2,3], Jaya A. George[4], Harriet R. Etheredge[1,3], Manuel van Deventer[4,5], Robert Kalyesubula[6,7], Alisha N. Wade[2], Laurie A. Tomlinson[8], Stephen Tollman[2,9] and Saraladevi Naicker[3] on behalf of the African Research in Kidney Disease (ARK) Working Group

[1]Wits Donald Gordon Medical Centre, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa, [2]Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, [3]Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, [4]Department of Chemical Pathology, National Health Laboratory Services, University of Witwatersrand, Johannesburg, South Africa, [5]Lancet Laboratories, Johannesburg, South Africa, [6]Medical Research Council/UVRI, London School of Hygiene and Tropical Medicine Research Unit, Entebbe, Uganda, [7]Department of Internal Medicine and Department of Physiology, Makerere University College of Health Sciences, Kampala, Uganda, [8]Department of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK and [9]International Network for the Demographic Evaluation of Populations and their Health (INDEPTH) Network, Accra, Ghana

Correspondence and offprint requests to: June Fabian; E-mail: june.fabian@mweb.co.za

ABSTRACT

Globally, chronic kidney disease (CKD) is an emerging public health challenge but accurate data on its true prevalence are scarce, particularly in poorly resourced regions such as sub-Saharan Africa (SSA). Limited funding for population-based studies, poor laboratory infrastructure and the absence of a validated estimating equation for kidney function in Africans are contributing factors. Consequently, most available studies used to estimate population prevalence are hospital-based, with small samples of participants who are at high risk for kidney disease. While serum creatinine is most commonly used to estimate glomerular filtration, there is considerable potential bias in the measurement of creatinine that might lead to inaccurate estimates of kidney disease at individual and population level. To address this, the Laboratory Working Group of the National Kidney Disease Education Program published recommendations in 2006 to standardize the laboratory measurement of creatinine. The primary objective of this review was to appraise implementation of these recommendations in studies conducted in SSA after 2006. Secondary objectives were to assess bias relating to choice of estimating equations for assessing glomerular function in Africans and to evaluate use of recommended diagnostic criteria for CKD. This study was registered with Prospero (CRD42017068151), and using PubMed, *African Journals Online* and Web of Science, 5845 abstracts were reviewed and 252 full-text articles included for narrative analysis. Overall, two-thirds of studies

Received: 15.1.2019; Editorial decision: 18.6.2019

did not report laboratory methods for creatinine measurement and just over 80% did not report whether their creatinine measurement was isotope dilution mass spectroscopy (IDMS) traceable. For those reporting a method, Jaffe was the most common (93%). The four-variable Modification of Diet in Renal Disease (4-v MDRD) equation was most frequently used (42%), followed by the CKD Epidemiology Collaboration (CKD-EPI) equation for creatinine (26%). For the 4-v MDRD equation and CKD-EPI equations, respectively, one-third to one half of studies clarified use of the coefficient for African-American (AA) ethnicity. When reporting CKD prevalence, <15% of studies fulfilled Kidney Disease: Improving Global Outcomes criteria and even fewer used a population-based sample. Six studies compared performance of estimating equations to measured glomerular filtration rate (GFR) demonstrating that coefficients for AA ethnicity used in the 4-v MDRD and the CKD-EPI equations overestimated GFR in Africans. To improve on reporting in future studies, we propose an 'easy to use' checklist that will standardize reporting of kidney function and improve the quality of studies in the region. This research contributes some understanding of the factors requiring attention to ensure accurate assessment of the burden of kidney disease in SSA. Many of these factors are difficult to address and extend beyond individual researchers to health systems and governmental policy, but understanding the burden of kidney disease is a critical first step to informing an integrated public health response that would provide appropriate screening, prevention and management of kidney disease in countries from SSA. This is particularly relevant as CKD is a common pathway in both infectious and non-communicable diseases, and multimorbidity is now commonplace, and even more so when those living with severe kidney disease have limited or no access to renal replacement therapy.

Keywords: albuminuria, chronic kidney disease, creatinine, estimated and measured glomerular filtration rate, prevalence, systematic review

INTRODUCTION

In sub-Saharan Africa (SSA), a double burden of infectious and non-communicable diseases contributes to a potentially high risk for chronic kidney disease (CKD). However, the true prevalence of CKD remains difficult to quantify [1, 2]. A systematic review of the epidemiology of CKD in SSA concluded that most studies were of medium to low quality due to poor sampling methods and unreliable laboratory measurements of kidney function [1]. Many of these studies were conducted in urban hospitals with participants who had multiple risk factors for CKD, and proteinuria was the most common measure of kidney function [1]. Given that convenience sampling may lead to overestimating the burden of CKD at population level, and that the diagnostic criteria used to define CKD were sub-optimal in many of the studies from the review, the reasons for this low research quality merit consideration.

First, the funding and infrastructure required for population-based CKD prevalence studies is substantial and, even in the most well-resourced environments, conducting these studies is challenging [2]. Second, in the absence of validated, affordable, point of care diagnostics for kidney disease, for studies to conform to the internationally accepted Kidney Disease: Improving Global Outcomes (KDIGO) guidelines requires (i) laboratory measurement of serum creatinine using an isotope dilution mass spectroscopy (IDMS)-traceable standard reference material for creatinine; (ii) calculation of the glomerular filtration rate (GFR) using the serum creatinine; (iii) measurement of urine protein or albumin from a spot urine sample, preferably laboratory quantified as an albumin:creatinine ratio; and (iv) in the absence of clinical evidence that confirms chronicity, repeating these serum and urine measurements after a minimum period of 3 months [3, 4]. These requirements impose substantial logistic, infrastructural and financial hurdles for clinical researchers, particularly in resource-limited settings where the burden of CKD is projected to be highest.

For any study using KDIGO criteria, access to accurate diagnostic laboratory services is essential as even small variations in creatinine measurement can impact on population prevalence estimates [2]. This is underappreciated by clinician scientists and may reflect poor interdisciplinary collaboration with chemical pathologists. For example, the older colorimetric picric acid (Jaffe) method is less accurate but cheaper than the recently developed enzymatic method. While the enzymatic method is recommended, in SSA most laboratories use the Jaffe method, for which a correction factor should be applied [1]. To reduce inter-laboratory measurement bias, the Laboratory Working Group of the National Kidney Disease Education Program (NKDEP) in the USA published guidelines in 2006 that recommended use of an IDMS-traceable standard reference material for creatinine measurement [5]. Consequently, the four-variable Modification of Diet in Renal Disease (4-v MDRD) equation was re-expressed for IDMS-traceability and the CKD Epidemiology Collaboration (CKD-EPI) equations are based on the use of an IDMS-traceable method [6, 7]. Adherence to standard internal and external quality assurance procedures are additional obligatory steps that mitigate laboratory-induced bias. In SSA and other low and middle income settings, reliable diagnostic laboratory infrastructure cannot be assumed. The lack of study-specific standardization in sampling and measurement of creatinine is considered the greatest obstacle to determining the global prevalence of CKD [2, 8].

Further bias may be introduced through the choice of estimating equation for GFR. Initially, the National Kidney Foundation guidelines recommended the 4-v MDRD equation for estimating GFR [3]. This equation was derived from a relatively small study sample in the USA with a coefficient that corrected for a higher creatinine observed in African-Americans (AAs). A similar correction coefficient for AA ethnicity was derived for use with the CKD-EPI equation, which is now the preferred equation in the revised KDIGO guidelines for 2014 [4]. In studies from SSA, when the coefficients for AA ethnicity are used for either of these two equations, they consistently overestimate GFR [9–15]. In the absence of appropriate, validated estimating equations for African populations, some studies have omitted the coefficients for AA ethnicity when using the 4-v MDRD and CKD-EPI equations [16, 17]. Researchers need to be cognizant of the limitations imposed by these equations and the bias that could be introduced into their results.

The primary objective of this review was to appraise reporting of laboratory methods for creatinine measurement in studies utilizing creatinine to assess kidney function in SSA. The secondary objectives were to assess bias relating to: choice of

equations used to estimate GFR; use of coefficients for AA ethnicity for the 4-v MDRD and CKD-EPI equations; criteria used to diagnose CKD; and study design and sampling strategies.

MATERIALS AND METHODS

Search strategy and selection criteria

This narrative systematic review was registered with the International Prospective Register of Systematic Reviews (PROSPERO) (CRD42017068151) and completed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) together with the revised quality assessment of diagnostic accuracy studies (QUADAS-2) guidelines [18–20]. The period selected for review included all original research studies published from 31 January 2008 to 31 December 2018. This was based on the assumption that after the NKDEP guidelines were published in 2006, this would set the standard for the widespread implementation of IDMS-traceable creatinine assays in diagnostic laboratories by 2008—and we would see a similar trend in studies from SSA [5]. For studies that determined the prevalence of CKD, the recommended criteria for diagnosis of CKD were first published in 2002 and subsequently updated in 2012 [3, 4]. Likewise, we anticipated that these guidelines would inform the choice of criteria for clinical studies.

The online databases for PubMed, *African Journals Online* and Web of Science were searched using the relevant medical subject headings (Supplementary data, Appendix S1). Based on the title and abstract, all studies from the SSA region that assessed kidney function in adults were evaluated according to inclusion and exclusion criteria agreed upon by the team conducting the systematic review (Supplementary data, Appendix S1). Only those abstracts with studies with full-text articles, available in English, were selected for the final analysis.

Quality assessment and data extraction

The online database search generated abstracts for screening. J.A.G. and M.v.D. independently screened abstracts from *African Journals Online* and Web of Science. There were no additional references identified through bibliography searches. Results were compared and differences of opinion referred to J.F. or H.R.E. Similarly, J.F. and H.R.E. independently screened abstracts from PubMed and followed the same process with differences referred to J.A.G. and M.v.D. Duplicates were removed and the team agreed on a list of eligible abstracts for which full-text articles were sourced by J.A.G. and H.R.E. For each full-text article, data were extracted by two authors who worked independently (J.F., J.A.G., H.R.E., M.v.D. and R.K.). Each author then compared their data extraction to the second author and discrepancies were resolved or referred to the remaining authors for review. If needed, J.F. contacted authors via email for clarification of study-specific information and incorporated the feedback for those who responded.

The QUADAS-2 assessment sheet (Supplementary data, Appendix S2) was used as a template by the team to pilot the data extraction process. After conducting the pilot, a revised form, adapted for the specific needs of this study, was generated and accepted by the group for use. This was formulated into a study-specific data extraction sheet (Supplementary data, Appendix S3). In summary, where appropriate for each full-text study, the data were abstracted as follows: (i) was GFR measured or estimated, and by which method; (ii) was creatinine measured, and by which method, and was the method IDMS-traceable to a standard reference material for creatinine; (iii) was cystatin C measured; (iv) did the study determine CKD prevalence—if so, by which criteria, was chronicity confirmed with repeat measurements; and (v) study design, sample selection and sample size. The estimating equations for creatinine, cystatin and creatinine-cystatin, and CKD criteria are defined in Supplementary data, Appendix S4. Data were analyzed using simple descriptive statistics including frequency and percentage tabulation for continuous variables. A narrative analysis was considered more appropriate for the aim of this systematic review and a meta-analysis was therefore not conducted.

Performance of eGFR equations

For those studies with measured GFR (mGFR) and comparisons of performance of the different estimated GFR (eGFR) equations, the relative performance of these equations was assessed using: (i) bias: median of difference between estimated and mGFR; 95% confidence interval (CI), when available; and (ii) P30: percentage of eGFR values within 30% of the (gold standard) mGFR; 95% CI, when available.

RESULTS

From the online database searches, there were 5845 records published during the review period. The procedure followed for study identification and selection is summarized in Figure 1. The final number of full-text articles assessed as eligible for inclusion into this systematic review totaled 252 (Supplementary data, Appendix S5). The results are presented as follows: (i) laboratory method for creatinine measurement; (ii) estimating equations for eGFR; (iii) mGFR using a gold standard method, comparison of the performance of eGFR equations in relation to mGFR and the impact of coefficients for AA ethnicity; (iv) diagnostic criteria for CKD; and (v) quality of the studies.

Laboratory method for creatinine measurement

The laboratory method for creatinine measurement was not reported in two-thirds of studies (159/252; 63.1%). For those that included a method, the Jaffe method was by far the most common (80/86; 93.0%). Only six studies described an enzymatic method. Most studies (206/252; 81.7%) did not state whether their laboratory was using an IDMS-traceable standard reference material for creatinine measurement. When stipulated, IDMS-traceable assays were more common (34/39; 87.2%) than non-IDMS traceable assays, of which there were only five (Table 1).

Estimating equations for GFR

Most studies used an estimating equation to assess kidney function (231/252; 91.6%) and some used more than one, so that eGFR equations were used in 363 instances. Of the available equations, the 4-v MDRD was most frequently used (146/363; 40.2%), followed by the CKD-EPI equation for creatinine (94/363; 25.9%) and Cockcroft–Gault (85/363; 23.4%). When the 4-v MDRD equation was used, one-third used the 4-v MDRD re-expressed for IDMS traceability (46/146; 31.5%), half did not stipulate which version was used (74/146; 50.7%) and the remainder used the original 4-v MDRD (prior to the introduction of an IDMS-traceable standard reference material for creatinine). The AA coefficient was used for a third of the 4-v MDRD equations (45/146; 30.8%) and almost half (72/146; 49.3%) did not stipulate if this was used. For the CKD-EPI creatinine equation, more studies used the AA coefficient (39/94; 41.5%) and a quarter did not clarify if this was used (23/94; 24.5%) (Table 1 and Supplementary data, Appendix S4).

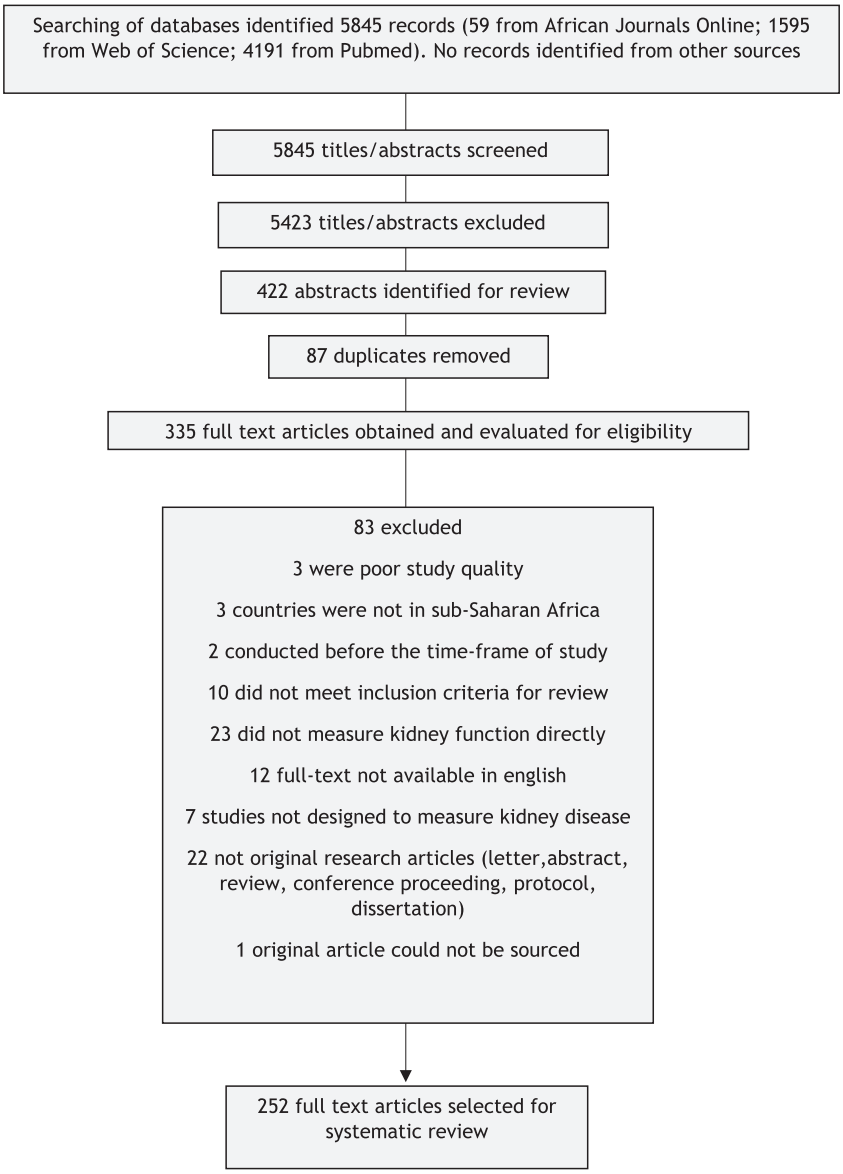


FIGURE 1: Flowchart for study identification and selection.

mGFR using a gold standard method, comparison of the performance of eGFR equations in relation to mGFR and the impact of coefficients for AA ethnicity

There were 10 studies that measured GFR using a gold standard reference method, 4 from South Africa and 1 each from the Democratic Republic of the Congo, Ghana, Ivory Coast, Kenya, Seychelles and Sudan [9–13, 15, 22–25]. The Kenyan study and one study from South Africa focused on adult participants with HIV infection [12, 13]. This was particularly relevant for the SSA region because of widespread use of tenofovir-containing antiretroviral therapy regimens as first-line treatment. The methods used to measure GFR included iohexol on dried blood spots, iohexol plasma excretion, technetium-99m diethylenetriamine penta-acetic acid (DTPA), chromium-51 ethylenediamine tetra-acetic acid (EDTA), inulin and 24-h urinary creatinine clearance. Six studies compared the performance of currently recommended eGFR equations to a reference mGFR method. The impact of using the coefficient for AA ethnicity for the 4-v MDRD and CKD-EPI equations was also evaluated. Overall, most studies demonstrated that when compared with the mGFR method that was used, inclusion of the coefficient for AA ethnicity resulted in overestimation of eGFR in Africans (Table 2). Some studies compared the relative performance of one or more eGFR equations without reference to a gold standard mGFR. Since the scientific validity of this practice is unsubstantiated, no further analysis was conducted.

Diagnostic criteria for CKD

Most studies (162/252; 64.5%) defined CKD using a broad range of criteria that have been summarized in Table 3. Fewer than 15% of studies confirmed chronicity by repeating the out of range test for eGFR and/or albuminuria/proteinuria after the minimum recommended period of 3 months, a requirement for diagnosing CKD as per KDIGO guidelines. Notably, most studies that fulfilled the KDIGO requirements were published in the latter half of the period under review [26–38].

Table 1. Reporting of laboratory creatinine measurement and eGFR equations

Laboratory creatinine measurement ($n = 252$ studies[a])	
Creatinine method	Jaffe/enzymatic 80 (31.7%)/6 (2.4%)
	Not stated 159 (63.1%)
	Not measured 7 (2.8%)
Creatinine method	IDMS-traceable 34 (13.5%)
	Non-IDMS-traceable 5 (2.0%)
	Not stated 206 (81.7%)
	Not measured 7 (2.8%)
eGFR equations[a] ($n = 363$ eGFR equations[b])	
Creatinine clearance ($n = 9$; 2.5%)	BSA[d] normalized (4/9)
	BSA not normalized (3/9)
	BSA not stated (2/9)
Cockcroft–Gault ($n = 85$; 23.4%)	BSA normalized (29/85)
	BSA not normalized (28/85)
	BSA not stated (28/85)
Non-IDMS-traceable 4-v MDRD ($n = 26$; 7.2%)	+ Coefficient for AA ethnicity (14/26)
	− Coefficient for AA ethnicity (5/26)
	Coefficient not stated (7/26)
IDMS-traceable 4-v MDRD ($n = 46$; 12.7%)	+ Coefficient for AA ethnicity (23/46)
	− Coefficient for AA ethnicity (15/46)
	Coefficient not stated (8/46)
4-v MDRD (not stated) ($n = 74$; 20.4%)	+ Coefficient for AA ethnicity (8/74)
	− Coefficient for AA ethnicity (9/74)
	Coefficient not stated (57/74)
CKD-EPI[c] (creatinine) ($n = 94$; 25.9%)	+ Coefficient for AA ethnicity (39/94)
	− Coefficient for AA ethnicity (32/94)
	Coefficient not stated (23/94)
CKD-EPI[c] (creatinine + cystatin C) ($n = 6$; 1.7%)	+ Coefficient for AA ethnicity (3/6)
	− Coefficient for AA ethnicity (3/6)
Other ($n = 23$; 6.3%)	eGFR equation not specified (17/23)
	Different eGFR equation used (6/23)

[a]Percentages rounded to one decimal point and may not sum to 100%.
[b]Of 252 studies, 21 did not use an eGFR method. Of the remaining studies (231), some evaluated more than one eGFR method, thus totaling 363 eGFR equations.
[c]CKD-EPI equation for creatinine alone, or creatinine and cystatin C, or cystatin C alone.
[d]BSA (body surface area), normalized to 1.73 m^2 [21].

Quality of studies

Most study designs were cross-sectional (172/252; 68.3%), fewer were case-controlled (41/252; 16.3%) and the methodology was unclear in the remainder (39/252; 15.4%). Many studies reported 'prevalent' CKD, but this was restricted to hospital- or clinic-based convenience samples and often focused on participants at high risk for kidney disease, for example, those with sickle cell disease, diabetes, hypertension, obesity, HIV and cardiac failure. For true population prevalence of CKD, there were eight randomly sampled population-based studies that determined prevalent CKD [10, 17, 39–44]. None of the population-based studies reported incident CKD.

DISCUSSION

There is scope for improving the quality of studies on kidney function in SSA. This includes aspects of study design and sampling, reporting of laboratory methods for creatinine measurement and IDMS traceability, detailing choices for GFR equations that include coefficients for AA ethnicity, and diagnosing CKD using appropriate criteria. While this is the first review of its kind for SSA, similar findings have been reported in Europe, Mexico, Uruguay and India [8, 45].

There are published NKDEP guidelines for the laboratory reporting of creatinine; however, two-thirds of our studies did not report the method of creatinine measurement and even fewer reported whether the method was IDMS-traceable to a standard reference material for creatinine. This limited insights into the extent of implementation of laboratory standards for creatinine-based testing of kidney function, but our findings reflect a need for better study-specific interdisciplinary collaboration between chemical pathologists, epidemiologists, public health and clinical scientists. Collaboration would deepen the scientific rationale and strengthen the methodological components of studies that characterize kidney disease, particularly at population level.

Many studies applied an estimating equation for GFR, but it is noteworthy that Cockcroft–Gault was still relatively frequently used, despite its omission from clinical practice guidelines since 2002. Perhaps this is due to the long history of this equation and its ease of use in clinical settings. Because of the time period for this review, it would be expected, as we have confirmed, that the 4-v MDRD equation was the most frequently used equation, but most studies did not identify which version of the equation was used in relation to IDMS traceability. Depending on which equation was used [CKD-EPI(serum creatinine, SCr) and/or 4-v MDRD], up to half of studies did not state whether the coefficients for AA ethnicity were used. In this regard, two recent studies from the Democratic Republic of the Congo and Ivory Coast (both using iohexol plasma excretion to measure GFR) deserve mention. The former, supporting prior findings from South and Eastern Africa, confirmed that omitting coefficients for AA ethnicity with CKD-EPI(SCr) and 4-v MDRD equations improved accuracy. Together, these studies highlight the critical importance of validating eGFR equations in populations in which they are recommended for use [9, 12, 13, 15]. The Ivorian study goes a few steps further, where for the first time in West Africans, normal reference ranges for GFR are now available and the performance of the Full Age Spectrum for creatinine (FAScreat) equation was validated. Originally derived from northern Caucasian populations, the FAScreat equation incorporates adjustment for age and gender as a Q value. In the Ivorian study, using Q values derived for West Africans, the performance of FAScreat was better than CKE-EPI(SCr) and requires no adjustment for ethnicity [25, 46]. If the FAScreat equation is validated in other regions of SSA and proven to be superior to the currently recommended CKD-EPI(SCr) equation, this provides strong support for changing current eGFR equations recommended for use in KDIGO clinical practice guidelines [4]. More broadly, remembering that performance of CKD-EPI(SCr) has been questioned in Asia, it is incumbent on policy makers—including KDIGO, to prioritize use of validated population-appropriate eGFR equations particularly when relating to diagnosis of CKD—given the global health implications [46, 47].

The gold standard reference methods used for mGFR also require some reflection due to the potential biases that might arise from their use [49–51]. In Europe, iohexol is the preferred method while in the USA, iothalamate is most commonly used. Recently, it has been suggested that iohexol may be the most practical and accurate measure of GFR for a few reasons: it is stable at a wide range of temperatures and has a very good clinical safety profile; iohexol is more accurate than iothalamate (which can systematically overestimate GFR by 10–15% due to tubular secretion); there is very little difference between

Table 2. Studies from SSA with mGFR comparing performance of eGFR equations with/without coefficients for AA ethnicity

Study	Sample size	Self-reported ethnicity; HIV status	mGFR method	mGFR (mL/min/1.73 m^2)	eGFR equation	Bias with AA ethnicity coefficient % (95% CI)	Bias without AA ethnicity coefficient % (95% CI)	P30 with AA ethnicity coefficient % (95% CI)	P30 without AA ethnicity coefficient % (95% CI)
South Africa 2008 [9]	100	Black African; 20 HIV positive	^{51}Cr-EDTA	61.5	4-v MDRD[a]	27	5	52	74
South Africa 2012 [11]	148	Black African; HIV negative	^{99}mTc-DTPA	38	4-v MDRD[b]	Not reported	Not reported	50	55
Kenya 2013 [12]	99	Black African; HIV positive	Iohexol blood spot	115	4-v MDRD[b]	18	−3	73	83
					CKD-EPI(SCr)[c]	10	−4	82	85
South Africa 2016 [13]	100	Black African; HIV positive	^{51}Cr-EDTA	92.5	4-v MDRD[a]	38.4 (27.5; 49.3)	14.2 (5.2; 23.2)	43.3	59.8
					CKD-EPI(SCr)[c]	33.7 (25.0; 42.4)	15.3 (7.8; 22.8)	41.2	62.9
					CKD-EPI(SCrCys)[d]	11.5 (5.4; 17.6)	2.9 (−2.9; 8.8)	73.0	78.0
Ivory Coast 2018 [24]	237	Black African; HIV negative	Iohexol plasma excretion	103	CKD-EPI(SCr)[c]	NA	NA	NA	NA
					[e]FAScreat	NA	NA	NA	NA
Democratic Republic of the Congo 2018 [15]	93	Black African; HIV status not declared	Iohexol plasma excretion	92.0	4-v MDRD[a]	13.6 (8.0; 19.2)	−4.9 (−9.6; −0.2)	79.6 (71.2; 87.9)	86.0 (78.8; 94.0)
					CKD-EPI(SCr)[c]	17.2 (13.3; 21.2)	2.3 (−1.3; 5.8)	73.1 (63.9; 82.3)	81.7 (73.7; 89.7)
					CKD-EPI(SCrCys)[d]	9.0 (5.9; 12.0)	1.5 (−1.4; 4.4)	87.1 (80.2; 94.0)	92.5 (87.0; 97.9)
					CKD-EPI(Cys)[f]	1.5 (−1.8; 4.7)		91.4 (85.6; 97.2)	

[a]Re-expressed IDMS traceable 4-v MDRD equation.
[b]Original 4-vMDRD equation.
[c]CKD-EPI equation for serum creatinine.
[d]CKD-EPI equation for serum creatinine and serum cystatin C [52].
[e]Full Age Spectrum (FAS) creatinine equation; using Q values derived for age and gender [25, 46].
[f]CKD-EPI equation for serum cystatin C (no adjustment for AA ethnicity) [53].

Table 3. Studies evaluating CKD in SSA ($n = 162$)

Number of studies[a,b] $n = 162$, n (%)	Parameter used to define CKD	Concordance with KDIGO definition of CKD: Creatinine-based eGFR equation	Urine protein or albumin	Confirmation of chronicity
4 (2.5)	Serum creatinine	☹	☹	☹
1 (0.6)	Serum creatinine + urine protein	☹	☺	☹
67 (41.3)	Serum creatinine-based eGFR	☺	☹	☹
8 (4.9)	Serum creatinine-based eGFR + follow up measurement at ≥3 months	☺	☹	☺
22 (13.6)	Serum creatinine-based eGFR + urine albumin	☺	☺	☹
38 (23.5)	Serum creatinine-based eGFR + urine protein	☺	☺	☹
1 (0.6)	Serum creatinine-based eGFR + urine albumin + urine protein	☺	☺	☹
12 (7.4)	Serum creatinine-based eGFR + urine protein + follow up measurement/s at ≥3 months	☺	☺	☺
3 (1.9)	Urine albumin	☹	☺	☹
5 (3.0)	Urine protein	☹	☺	☹
1 (0.6)	Urine protein + follow up measurement at ≥3 months	☹	☺	☺

[a]One study was excluded as the definition used for CKD was unclear.
[b]Percentages have been rounded to one decimal point and might not sum to 100%.

laboratory techniques for measuring iohexol, for example when comparing high-performance liquid chromatography with ultraviolet detection and liquid chromatography-tandem mass spectrometry, whereas differences have been demonstrated with iothalamate; and there is an external quality control program for iohexol to aid standardizing interlaboratory variation in measurements which does not exist for iothalamate [48]. It is critical to contextualize this for SSA, as the 4-v MDRD and CKD-EPI eGFR equations were all developed using iothalamate and the coefficient for AA ethnicity would further overestimate GFR. For the few studies in this review that compared mGFR with eGFR, the reference methods differed, but none used iothalamate and three used iohexol (two intravenously for measurement of plasma excretion and one with dried blood spots) [12, 15]. While iohexol blood spots may be relatively inferior to intravenous administration of iohexol, the use of dried blood spots was possibly the most pragmatic, on the assumption that access to laboratory support was unlikely in rural Kenya.

This systematic review highlights potential sources of bias inherent in the assessment of kidney function in clinical studies in SSA. This has created the opportunity to increase awareness of the requirements for laboratory-based creatinine assays, the appropriate choice of estimating equations for calculating GFR and the appropriate use of diagnostic criteria for CKD. In response, we propose an 'easy to use' checklist for researchers as a guide for the reporting of kidney function in SSA (Table 4). We hope this will minimize bias and strengthen future studies conducted in the region. In addition, inferring population prevalence of CKD in SSA from small convenience samples of individuals at high risk for developing CKD must be cautioned, as the risk of overestimating CKD is substantial.

While the focus of this systematic review has been clinical research, the implications of our findings have much broader application for the management of kidney disease in SSA. This encompasses individual patient care in the setting of acute and chronic kidney disease, home-based monitoring of CKD as a component of an integrated care model for chronic disease management, and public health policy for screening and prevention of CKD in SSA. None of this can be realized without affordable and accurate diagnostics for kidney disease. To achieve this, validated population-appropriate estimating equations for glomerular function need to be prioritized, as well as innovative approaches to accurate and affordable point of care diagnostics for assessing kidney function and diagnosing CKD.

SUPPLEMENTARY DATA

Supplementary data are available at ckj online.

FUNDING

There was no specific funding to conduct this systematic review. However, this research was undertaken as part of a broader multicenter collaborative study between South Africa, Uganda, Malawi and the London School of Hygiene and Tropical Medicine, which is collectively identified as the African Research in Kidney Disease (ARK) Network. This is jointly funded by the South African Medical Research Council, with funding from the South African National Department of Health, MRC UK (via the Newton Fund) and GlaxoSmithKline Africa Non-Communicable Disease Open Lab Grant [Project Number: 8111 (Uganda and Malawi) and 074 (South Africa)]. The funder had no role in study design, data collection and analysis or decision to publish. Authors retained control of the final content of the publication.

Table 4. Recommendations for reporting kidney function in SSA populations: the African Research of Kidney Disease (ARK) checklist for researchers

mGFR (gold standard reference method)—method and biomarker (^{51}Cr-EDTA; ^{99}mTc-DTPA; inulin, iohexol, iothalamate)
- Urinary clearance of biomarker, state which biomarker; OR
- Plasma clearance of biomarker, state which biomarker

Laboratory creatinine method—include all the following:
- Enzymatic
- Jaffe (alkaline picrate): modified or compensated
- IDMS traceable to a standard reference material
- The external quality control program used by the laboratory for creatinine

Estimating equations for GFR—state which equation was used:
4-v MDRD equation
Original 4-v MDRD equation
[eGFR (mL/min/1.73 m^2) = 186 × SCr$^{-1.154}$ × age (years)$^{-0.203}$ × (0.742 if female) × (1.1212 if AA)] [53]
Use if laboratory method for creatinine was not IDMS-traceable
State whether the coefficient for AA ethnicity was used
Re-expressed[a] *4-v MDRD equation*
[eGFR (mL/min/1.73 m^2) = 175 × SCr$^{-1.154}$ × age (years)$^{-0.203}$ × (0.742 if female) × (1.1212 if AA)] [6]
Use if laboratory method for creatinine was IDMS-traceable
State whether the coefficient for AA ethnicity was used

CKD-EPI equation for creatinine
CKD-EPI equation
[eGFR (mL/min/1.73 m^2) = 141 × min(SCr/κ, 1)$^{\alpha}$ × max(SCr/κ, 1)$^{-1.209}$ × 0.993^{age} × 1.018 (if female) × 1.159 (if black)] [21]
Laboratory method for creatinine measurement must be IDMS-traceable
State whether the coefficient for AA ethnicity was used

Cockcroft–Gault equation
In its original form, this equation does not adjust for body surface area (BSA). To compare this equation to 4-v MDRD or CKD-EPI equations, which are adjusted for BSA, it is necessary to use the duBois formula and adjust for BSA [21]
Cockcroft–Gault equation
{eGFR (mL/min) = [140 − age (years) × weight (kg) × (0.85 if female)]/S-Cr}[54]
Cockcroft–Gault equation with BSA normalized to 1.73 m^2
{eGFR (mL/min/1.73 m^2) = [140 − age (years) × weight (kg) × (0.85 if female) × 1.73 m^2]/[S-Cr × BSA (m^2)]}

Full Age Spectrum (FAS) equation for creatinine
FAScrea = 107.3/(SCr/Qcrea) × {[0.988(Age − 40) when age >40 years]} [25]
Qcrea = derived from reference healthy population

Diagnosis of CKD using KDIGO criteria [4]—include the following:
True prevalence requires a randomized population-based sample: describe the sampling strategy
KDIGO Clinical Practice Guidelines (2012) are recommended for diagnosis of CKD and require testing for:
- Urine albumin/protein—if qualitative, confirm with quantitative test, preferably albumin:creatinine ratio **AND**
- Serum creatinine: use CKD-EPI equation[b] for calculation of eGFR
- In the absence of prior testing or additional supporting evidence[c] that confirms chronicity, demonstrate chronicity with a repeat of the abnormal diagnostic test after a minimum 12 weeks.

Recommendation: In SSA, for CKD-EPI equation—omit coefficient for AA ethnicity [9–13, 24]

[a]Re-expressed using an IDMS-traceable assay to a standard reference material.
[b]Use of 4-v MDRD and Cockcroft–Gault equations not recommended.
[c]Supporting evidence can be findings on renal ultrasound; and/or proof of pre-existing kidney disease from medical records or prior urine and serum test results.

A.N.W. is supported by the Fogarty International Centre of the National Institutes of Health under Award Number K43TW010698. This article describes the views of the authors and does not necessarily represent the official views of the National Institutes of Health (USA).

L.A.T. is funded by a Wellcome Intermediate Clinical Fellowship (reference: WT101143MA).

AUTHORS' CONTRIBUTIONS

J.A.G., M.v.D., H.R.E., R.K. and J.F. reviewed full-text articles and extracted the data. J.F. conducted the narrative analysis, wrote the first draft of the paper and incorporated revisions into the final draft. H.R.E. compiled recommendations for reporting of studies that assess kidney function. H.R.E., J.A.G., R.K., S.N., S.T., L.A.T., M.v.D. and A.N.W. contributed to revising the first draft, editing and review of the final draft of the article.

CONFLICT OF INTEREST STATEMENT

None declared. The results presented in this article have not been published previously in whole or part, except in abstract format.

REFERENCES

1. Stanifer JW, Jing B, Tolan S *et al.* The epidemiology of chronic kidney disease in sub-Saharan Africa: a systematic

review and meta-analysis. *Lancet Glob Health* 2014; 2: e174–e181
2. Coresh J, Hu JR, Bello AK *et al*. Action plan for determining and monitoring the prevalence of chronic kidney disease. *Kidney Int Suppl* 2017; 7: 63–70
3. Levey AS, Coresh J, Bolton K *et al*. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis* 2002; 39: S1–S266
4. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney Int Suppl* 2013; 3: 1–150
5. Myers GL, Miller WG, Coresh J *et al*. Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program. *Clin Chem* 2006; 52: 5–18
6. Levey A, Coresh J, Greene T *et al*. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. *Ann Int Med* 2006; 145: 247–254
7. Levey AS, Stevens LA, Schmid CH *et al*. A new equation to estimate glomerular filtration rate. *Ann Int Med* 2009; 150: 604–612
8. Brück K, Jager KJ, Dounousi E *et al*. Methodology used in studies reporting chronic kidney disease prevalence: a systematic literature review. *Nephrol Dial Transplant* 2015; 30: iv6–iv16
9. van Deventer HE, George JA, Paiker JE *et al*. Estimating glomerular filtration rate in black South Africans by use of the modification of diet in renal disease and Cockcroft-Gault equations. *Clin Chem* 2008; 54: 1197–1202
10. Eastwood JB, Kerry SM, Plange-Rhule J *et al*. Assessment of GFR by four methods in adults in Ashanti, Ghana: the need for an eGFR equation for lean African populations. *Nephrol Dial Transplant* 2010; 25: 2178–2187
11. Madala ND, Nkwanyana N, Dubula T *et al*. Predictive performance of eGFR equations in South Africans of African and Indian ancestry compared with 99mTc-DTPA imaging. *Int Urol Nephrol* 2012; 44: 847–855
12. Wyatt CM, Schwartz GJ, Owino Ong'or W *et al*. Estimating kidney function in HIV-infected adults in Kenya: comparison to a direct measure of glomerular filtration rate by iohexol clearance. *PLoS One* 2013; 8: e69601
13. Seape T, Gounden V, van Deventer HE *et al*. Cystatin C-and creatinine-based equations in the assessment of renal function in HIV-positive patients prior to commencing highly active antiretroviral therapy. *Ann Clin Biochem* 2016; 53: 58–66
14. Moodley N, Hariparshad S, Peer F *et al*. Evaluation of the CKD-EPI creatinine based glomerular filtration rate estimating equation in Black African and Indian adults in KwaZulu-Natal, South Africa. *Clin Biochem* 2018; 59: 43–49
15. Bukabau JB, Sumaili EK, Cavalier E *et al*. Performance of glomerular filtration rate estimation equations in Congolese healthy adults: the inopportunity of the ethnic correction. *PLoS One* 2018; 13: e0193384
16. Glaser N, Deckert A, Phiri S *et al*. Comparison of various equations for estimating GFR in Malawi: how to determine renal function in resource limited settings? *PLoS One* 2015; 10: e0130453
17. Stanifer JW, Maro V, Egger J *et al*. The epidemiology of chronic kidney disease in Northern Tanzania: a population-based survey. *PLoS One* 2015; 10: e0124506
18. Liberati A, Altman DG, Tetzlaff J *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009; 6: e1000100
19. Whiting PF, Rutjes AWS, Westwood ME *et al*. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Int Med* 2011; 155: 529–536
20. Systematic Review of the Assessment of Kidney Function Using Measured and Estimated Glomerular Filtration Rate in Sub-Saharan Africans [Internet], 2017. https://www.crd.york.ac.uk/PROSPERO/display_record.php? RecordID=68151 (11 March 2019, date last accessed)
21. Du Bois D, Du Bois E. A formula to estimate the approximate surface area if height and weight are known. *Arch Int Med* 1917; 17: 863–871
22. van Deventer HE, Paiker JE, Katz IJ *et al*. A comparison of cystatin C-and creatinine-based prediction equations for the estimation of glomerular filtration rate in black South Africans. *Nephrol Dial Transplant* 2011; 26: 1553–1558
23. Wuerzner G, Pruijm M, Maillard M *et al*. Marked association between obesity and glomerular hyperfiltration: a cross-sectional study in an African population. *Am J Kidney Dis* 2010; 56: 303–312
24. Osman AO, Elmadani AE. Comparison of slope-intercept with single plasma sample methods in estimating glomerular filtration rate using radionuclides. *Saudi J Kidney Dis Transplant* 2014; 25: 321
25. Yayo E, Ayé M, Yao C *et al*. Measured (and estimated) glomerular filtration rate: reference values in West Africa. *Nephrol Dial Transplant* 2017; 33: 1176–1180
26. Madala ND, Dubula T, Assounga AGH *et al*. Association of kidney function and waist circumference with uric acid levels in South Africans. *Metab Syndr Relat Disord* 2017; 15: 500–506
27. Adugna T, Merga H, Gudina EK. Impaired glomerular filtration rate, high grade albuminuria and associated factors among adult patients admitted to tertiary Hospital in Ethiopia. *BMC Nephrol* 2018; 19: 345
28. Amoako YA, Laryea DO, Bedu-Addo G *et al*. Clinical and demographic characteristics of chronic kidney disease patients in a tertiary facility in Ghana. *Pan Afr Med J* 2014; 18: 274
29. Arogundade F, Sanusi A, Hassan M *et al*. An appraisal of kidney dysfunction and its risk factors in patients with sickle cell disease. *Nephron Clin Pract* 2011; 118: c225–c231
30. Babua C, Kalyesubula R, Okello E *et al*. Cardiovascular risk factors among patients with chronic kidney disease attending a tertiary hospital in Uganda: cardiovascular topics. *Cardiovasc J Afr* 2015; 26: 177–180
31. Cailhol J, Nkurunziza B, Izzedine H *et al*. Prevalence of chronic kidney disease among people living with HIV/AIDS in Burundi: a cross-sectional study. *BMC Nephrol* 2011; 12: 40
32. Ekiti ME, Zambo JB, Assah FK *et al*. Chronic kidney disease in sugarcane workers in Cameroon: a cross-sectional study. *BMC Nephrol* 2018; 19: 10
33. Hamadou B, Boombhi J, Kamdem F *et al*. Prevalence and correlates of chronic kidney disease in a group of patients with hypertension in the Savanah zone of Cameroon: a cross-sectional study in Sub-Saharan Africa. *Cardiovasc Diagn Ther* 2017; 7: 581
34. Kaze FF, Kengne AP, Magatsing CT *et al*. Prevalence and determinants of chronic kidney disease among hypertensive Cameroonians according to three common estimators of the glomerular filtration rate. *J Clin Hypertens* 2016; 18: 408–414
35. Madala ND, Thusi GP, Assounga AG *et al*. Characteristics of South African patients presenting with kidney disease in

rural KwaZulu-Natal: a cross sectional study. *BMC Nephrol* 2014; 15: 61
36. Namuyimbwa L, Atuheire C, Okullo J *et al*. Prevalence and associated factors of protein-energy wasting among patients with chronic kidney disease at Mulago hospital, Kampala-Uganda: a cross-sectional study. *BMC Nephrol* 2018; 19: 139
37. Onodugo O, Ulasi I, Ijoma C *et al*. Predictors of autonomic dysfunction among predialysis chronic kidney disease patients in Nigeria. *Niger J Clin Pract* 2018; 21: 932–938
38. Marie PH, Moussa O, Francois K *et al*. Prevalence and associated factors of chronic kidney disease among patients infected with human immunodeficiency virus in Cameroon. *Iran J Kidney Dis* 2018; 12: 268–274
39. Kalyesubula R, Nankabirwa JI, Ssinabulya I *et al*. Kidney disease in Uganda: a community based study. *BMC Nephrol* 2017; 18: 116
40. Matsha TE, Yako YY, Rensburg MA *et al*. Chronic kidney diseases in mixed ancestry south African populations: prevalence, determinants and concordance between kidney function estimators. *BMC Nephrol* 2013; 14: 75
41. Kaze FF, Meto DT, Halle MP *et al*. Prevalence and determinants of chronic kidney disease in rural and urban Cameroonians: a cross-sectional study. *BMC Nephrol* 2015; 16: 117
42. Seck SM, Doupa D, Guéye L *et al*. Chronic kidney disease epidemiology in northern Senegal: a cross-sectional study. *Iran J Kidney Dis* 2014; 8: 286
43. Oluyombo R, Ayodele O, Akinwusi P *et al*. A community study of the prevalence, risk factors and pattern of chronic kidney disease in Osun State, South West Nigeria. *West Afr J Med* 2013; 32: 85–92
44. Sumaili EK, Krzesinski JM, Zinga CV *et al*. Prevalence of chronic kidney disease in Kinshasa: results of a pilot study from the Democratic Republic of Congo. *Nephrol Dial Transplant* 2008; 24: 117–122
45. Ruiz-Arenas R, Sierra-Amor R, Seccombe D *et al*. A summary of worldwide national activities in chronic kidney disease (CKD) testing. *EJIFCC* 2017; 28: 302–314
46. Pottel H, Hoste L, Dubourg L *et al*. An estimated glomerular filtration rate equation for the full age spectrum. *Nephrol Dial Transplant* 2016; 31: 798–806
47. Kim BS, Lee YK, Choi HY *et al*. Is the new GFR equation using inulin clearance a more accurate method for Asian patients? *Clin Nephrol* 2015; 84: 331–338
48. Wang J, Xie P, Huang J-M *et al*. The new Asian modified CKD-EPI equation leads to more accurate GFR estimation in Chinese patients with CKD. *Int Urol Nephrol* 2016; 48: 2077–2081
49. Seegmiller JC, Burns BE, Schinstock CA *et al*. Discordance between iothalamate and iohexol urinary clearances. *Am J Kidney Dis* 2016; 67: 49–55
50. Delanaye P, Ebert N, Melsom T *et al*. Iohexol plasma clearance for measuring glomerular filtration rate in clinical practice and research: a review. Part 1: how to measure glomerular filtration rate with iohexol? *Clin Kidney J* 2016; 9: 682–699
51. Inker LA, Levey AS, Coresh J. Estimated glomerular filtration rate from a panel of filtration markers—hope for increased accuracy beyond measured glomerular filtration rate? *Adv Chronic Kidney Dis* 2018; 25: 67–75
52. Inker LA, Schmid CH, Tighiouart H *et al*. Estimating glomerular filtration rate from serum creatinine and cystatin C. *N Engl J Med* 2012; 367: 20–29
53. Levey AS, Bosch JP, Lewis JB *et al*. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. *Ann Int Med* 1999; 130: 461–470
54. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. *Nephron* 1976; 16: 31–41

Kalyesubula et al. BMC Nephrology (2020) 21:20
https://doi.org/10.1186/s12882-020-1688-0

BMC Nephrology

STUDY PROTOCOL Open Access

How to estimate glomerular filtration rate in sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study

Robert Kalyesubula[1,2,3*†], June Fabian[4,5†], Wisdom Nakanga[6], Robert Newton[1], Billy Ssebunnya[1], Josephine Prynn[6], Jaya George[7], Alisha N. Wade[5], Janet Seeley[1,3], Dorothea Nitsch[3], Christian Hansen[1,3], Moffat Nyirenda[1,3,6], Liam Smeeth[3], Saraladevi Naicker[8], Amelia C. Crampin[3,6] and Laurie A. Tomlinson[3]

Abstract

Background: Chronic kidney disease (CKD) is a substantial cause of morbidity and mortality worldwide with disproportionate effects in sub-Saharan Africa (SSA). The optimal methods to estimate glomerular filtration rate (GFR) and therefore to determine the presence of CKD in SSA are uncertain. We plan to measure iohexol excretion to accurately determine GFR in Malawi, South Africa and Uganda. We will then assess the performance of existing equations to estimate GFR and determine whether a modified equation can better improve estimation of GFR in sub-Saharan Africa.

Methods: The African Research on Kidney Disease (ARK) study is a three-country study embedded within existing cohorts. We seek to enrol 3000 adults > 18 years based on baseline serum creatinine. Study procedures include questionnaires on socio-demographics and established risk factors for kidney disease along with anthropometry, body composition, blood pressure, blood chemistry and urine microscopy and albuminuria. We will measure GFR (mGFR) by plasma clearance of iohexol at 120, 180 and 240 min. We will compare eGFR determined by established equations with mGFR using Bland-Altman plots. We will use regression methods to estimate GFR and compare the newly derived model with existing equations.

Discussion: Through the ARK study, we aim to establish the optimal approach to estimate GFR in SSA. The study has the advantage of drawing participants from three countries, which will increase the applicability of the findings across the region. It is also embedded within established cohorts that have longitudinal information and serial measures that can be used to characterize kidney disease over a period of time. This will help to overcome the limitations of previous research, including small numbers, selected population sub-groups, and lack of data on proteinuria.

The ARK collaboration provides an opportunity for close working partnerships across different centres, using standardized protocols and measurements, and shared bio-repositories. We plan to build on the collaboration for this study for future work on kidney disease in sub-Saharan Africa, and welcome additional partners from across the continent.

Keywords: Measurement, Glomerular filtration rate, Chronic kidney disease, Sub-Saharan Africa

* Correspondence: rkalyesubula@gmail.com
†Robert Kalyesubula and June Fabian contributed equally to this work.
[1]Medical Research Council/ UVRI & London School of Hygiene and Tropical Medicine Research Unit, Entebbe, Uganda
[2]Makerere University College of Health Sciences, Kampala, Uganda
Full list of author information is available at the end of the article

Background

Chronic kidney disease (CKD) is a substantial cause of morbidity and mortality worldwide with an estimated prevalence of 8–16% [1]. A recent systematic review estimated the prevalence of CKD in sub-Saharan Africa to be around 14% [2]. However, the optimal equation to accurately estimate glomerular filtration rate (GFR) among sub-Saharan African populations is uncertain. This uncertainty is a major barrier to identifying individuals with CKD and to estimation of disease burden.

Worldwide, several equations have been used to estimate GFR from serum creatinine [3–7]. Unfortunately, most of these have been derived from high-income countries with minimal validation in sub-Saharan African populations. Their accuracy at an individual level is limited, particularly in relation to the adjustments for African-American ethnicity, which might be related to variations in muscle mass. Serum creatinine levels at a given level of renal function vary substantially with ethnicity, age, sex, physical activity and nutritional status [8–14]. Several sources of inaccuracy in estimating GFR have been described, including biological variability in serum creatinine, laboratory induced errors in creatinine measurement and choice of estimating equation [15, 16]. The imprecision of creatinine measurement is more marked at values near the normal range where it is most critical to determine earlier stages of CKD [17–19]. Several methods for direct measurement of GFR are available through measurement of clearance of inulin, iothalamate, iohexol, and radio-active agents such as technetium-99 m diethylenetriamine penta-acetic acid (DTPA) and chromium-51 ethylene diamine tetra-acetic acid (EDTA), and 24-h urinary creatinine collection for estimation of creatinine clearance [20]. However, all of these have their challenges. DTPA, EDTA and inulin are expensive and not readily available in most African countries, while 24-h urinary collection is often inaccurate due to difficulty in ensuring a complete sample, coupled with the additional limitation of tubular creatinine secretion [20–22].

Iohexol is a readily available compound, which can be used to measure the GFR. Its advantages include low cost, excellent safety profile, low protein binding, low levels of toxicity at the dose used for measuring GFR, stability at room temperature (20 to 25 °C), and being able to provide an accurate measure of glomerular filtration [20, 23]. Previous studies using iohexol to measure GFR in sub-Saharan Africa have included few people with CKD [13, 24, 25]. The largest study to date conducted in the Democratic Republic of Congo and Ivory Coast used plasma iohexol to measure GFR in 494 participants. They noted that the African-American race coefficient did not improve the performance of creatinine-based GFR estimates of iohexol GFR. In particular, the Chronic Kidney Disease Epidemiology (CKD-EPI) and Modification of Diet in Renal Disease (MDRD) equations performed better without the race coefficient in participants with GFR ≥60 mL/min/1.73m^2. The authors also evaluated the Full Age Spectrum (FAS) equation and found it to be as accurate as CKD-EPI for GFR ≥60 mL/min/1.73m^2 but better for those with creatinine based GFR < 60 mL/min/1.73m^2. Addition of cystatin C did not improve performance of the equations among this study group [25].

In the proposed study, we will use iohexol excretion in a population sample of 3000 participants with and without CKD to measure GFR and determine the optimal equation to estimate GFR in Uganda, Malawi and South Africa.

Methods

Primary objectives

1. To measure GFR using plasma clearance of iohexol
2. To assess the performance of existing eGFR equations, namely Cockcroft -Gault (CG), Modification of Diet in Renal Disease (4-v MDRD) and Chronic Kidney Disease-Epidemiology collaboration (CKD-EPI) by comparing these to measured GFR, among sub-Saharan Africans.

Secondary objectives

In a Ugandan sub-sample, to assess the feasibility of using dry blood spot collection of iohexol in the determination of GFR.

In a Ugandan sub-sample to explore the current community understanding of kidney disease and available treatments from traditional healers and herbalists.

Study organization and sites

We will have three participant countries across SSA (Fig. 1) and these centres have formed a new collaborative group, African Research on Kidney Disease (ARK) around this initial study.

A working committee will review the protocols for the design, conduct, progress and data collection and analysis plan. The committee will meet in person on several occasions to review protocols and measurement tools and to agree on a minimum dataset. The three centres have unique characteristics that will enrich the study with diversity among participants. The details of each of the different cohorts in which the current study will be nested are described below:

Uganda

In Uganda, the study will be based in the General Population Cohort (GPC), originally established in 1989 [26]. The study area is located in rural south-western Uganda

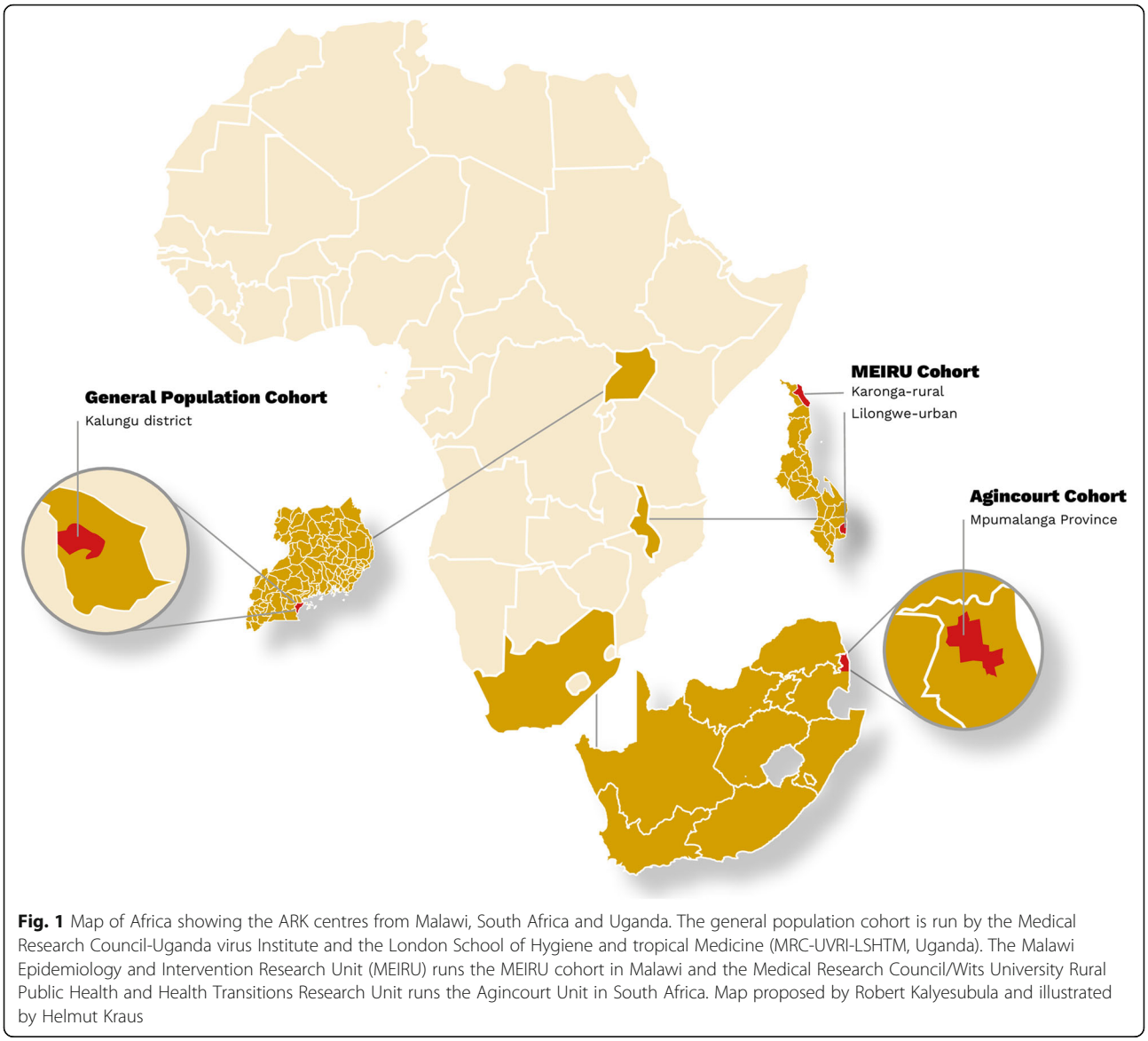


Fig. 1 Map of Africa showing the ARK centres from Malawi, South Africa and Uganda. The general population cohort is run by the Medical Research Council-Uganda virus Institute and the London School of Hygiene and tropical Medicine (MRC-UVRI-LSHTM, Uganda). The Malawi Epidemiology and Intervention Research Unit (MEIRU) runs the MEIRU cohort in Malawi and the Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit runs the Agincourt Unit in South Africa. Map proposed by Robert Kalyesubula and illustrated by Helmut Kraus

in one sub-county of Kalungu district, approximately 120 km from Entebbe town. The cohort comprises all residents of 25 adjacent villages, approximately 22,000 people with men and women in equal proportions, and 52% older than 13 years. The study population for the general population cohort is recruited through annual house-to-house rounds of the census where study participants are selected. Medical care is available through an established general clinic.

Malawi

The Malawi Epidemiology and Intervention Research Unit (MEIRU) conducts ongoing epidemiological studies based at two sites, which will be used for recruiting participants for this study. The first is Karonga District in rural northern Malawi, predominantly subsistence farmers, fishermen and informal traders. The second site is Area 25 of Lilongwe, the capital city of Malawi. The urban area is socio-economically mixed and includes government and industry employees, traders and those in casual employment. We will enrol participants from both sites, providing a true population-wide sample and enabling urban: rural comparisons.

The stratified study sample was drawn from a pool of over 5000 individuals who were tested for serum creatinine as part of a larger survey (n = 33,000) of chronic conditions. Medical care is available through an established chronic care Non-Communicable Disease (NCD) clinic.

South Africa

In South Africa, the study will be located within the Medical Research Council/University of the Witwatersrand

Rural Public Health and Health Transitions Research Unit in Agincourt, Bushbuckridge sub-district of Mpumalanga province. The area includes a high-functioning health and socio-demographic surveillance system (HDSS) which covers approximately 115,000 people based on an annual update of resident status and vital events. The study setting also serves as a national pilot site for the development of integrated chronic disease care. The population still has gaps in access to electricity, water and tarred roads and unemployment rates are high, leading to high rates of labour migration. Patients will be managed through the primary health care system consisting of six clinics, two health centres and three district hospitals. The HDSS has a central clinic with advanced laboratory capacity, which will be used for this study.

Study design

This will consist of cross-sectional studies embedded within ongoing cohorts in Malawi and Uganda and a longitudinal study in South Africa, all with serial biobanking.

Data collection, sample size and selection of sample

We intend to recruit 3000 participants in total. We have based the power calculation is based on the number of study participants needed to examine the accuracy of GFR estimating equations for 'true' GFR, at each of the three study sites separately. We have specified this as having 90% power to detect whether eGFR is within 5% of the iohexol value at an eGFR of 60mls/min, assuming a standard deviation of 25mls/min [9] and with a two-sided alpha of 0.05. This gives an estimated required sample size of 730 participants. Allowing for participants who do not wish to participate in this element of the study, and technical failures, we intend to recruit 1000 participants in each country.

In order to fulfil the aims of the study, a structured approach is needed. This will differ between South Africa and the other two sites (Malawi and Uganda).

Details of site recruitment

We plan to recruit from each of the sites guided by the previous recruitment protocols. Each of the sites will however, have two phases of recruitment.

For the Uganda GPC, the first phase will consist of measurement of creatinine from stored sera.

Participant selection for this part of the study has been described elsewhere [27]. Briefly, participants were selected from the general population cohort, which is a community-based open cohort study of residents of 25 neighbouring villages. The participants are selected through house-to-house mobilisation and community surveys conducted through village-based hubs. For Uganda, 5979 individuals were identified for baseline creatinine measurement, which included all the adults surveyed in that round. We will use creatinine levels from the stored samples to stratify participants for the second phase of the study. Here, we will select participants according to eGFR cut offs of normal (>90mls/min/1.73m^2); impaired renal function (60-90mls/min/1.73m$^{2)}$ and low GFR (<60mls/min/1.73m^2), all additionally stratified by age and sex. Selected participants will be approached by the community engagement lead (community mobilizer) for each village. Once participants assemble, the study team will take time to explain the details of the study to them, after which individual informed consent will be obtained. Participants will undertake a questionnaire to collect demographic data as well as history of known risk factors for kidney disease. Participants will also undergo detailed biophysical measurements and blood draws as detailed in the next sections. We will give participants an appointment to attend the medical clinic at the central Research Station, where iohexol measurements and other study tests will be performed.

For Malawi, for the first phase of the study, we will use creatinine assays from a previous survey to guide the selection of participants for the second phase of the study using a similar approach to that of Uganda. However, we will recruit participants 18 years and above from Malawi through household visits where similar procedures to those outlined above will be undertaken. We will carry out iohexol testing and other physical tests at a research clinic (phase 2). Details of participant recruitment have been detailed elsewhere [28].

In South Africa, we will determine baseline creatinine and the prevalence of CKD using an age and gender stratified random sample of 2250 members of the rural Agincourt HDSS. The number of participants sampled in each stratum will be determined in order to ensure proportional allocation, based on the population demographic distribution. We will visit participants in their homes for the first phase of the study where the questionnaires, biophysical measurements and take blood draws to determine Human immunodeficiency virus (HIV) status, blood sugar level, lipid profile as well as the baseline creatinine level. We will repeat creatinine and urine albumin measures after a minimum period of 3 months to confirm CKD. For the second phase of the study, we will recruit participants from Phase 1 through household visits and give them a referral date to the clinic for iohexol measurement along with other tests (See Table 1 and Fig. 2 for details).

Inclusion criteria

- Adults aged 18 years and above from the three population cohorts

Table 1 Clinical and Laboratory measurements within the ARK study, by country

Type of information	Equipment	Field procedures	Participant type	Definition of variable	Site participation		
					Malawi	S. Africa	Uganda
Demographics	Questionnaire	Face-to-face interview	all	Sex, age, education status, marital status	x	x	x
Chronic disease history and family history	Questionnaire / medical records	Face-to-face interview	all	HTN, DM, CKD, Stroke, heart disease	x	x	x
Treatment history	Patient report/ records	Face-to-face interview	all	HTN, DM, CKD, Stroke, HIV,TB, Heart disease, Cancer, Backache	x	x	x
Health behaviours	Patient report	Face-to-face interview	all	Tobacco use, alcohol consumption, Physical activity, vegetable, fruits, salt and water intake	x		x
Traditional medicine use	Questionnaire	Face-to-face interview	all	Use of herbs and traditional medicines	x	x	x
Physical examination	Omron® SECA® Flexible tape SECA© scales	Clinical Examination in the field, machines calibrated weekly	all	BP (mmHg), waist and hip circumferences (cms), height (m), weight (Kgs)	x	x	x
ABPM	ABPM spacelabs®	24 h BP, 30 min interval day, 1 h interval night	Sample per eGFR quarter	BP wake periods and sleep periods	x		x
BIA	Bodystat®	Tested in the field or clinic with participant, calibration weekly	all	Fat mass (kg), Lean mass (kg), Dry lean mass(kg), Total body water (L), Impedance at 50 KHZ (Ω)	x	x	x
DXA SCAN	Hologic Discovery A. QDR 4500 Series	Whole body scan performed in clinical research clinic during iohexol testing	all	Lean Mass (g), Fat mass (g) Fat % and BMI		x	
Ultra sound scan	Logiq e	Performed at clinic	all	Probe 4c for kidneys and bladder and probe12L CIMT probe		x	
ECG	ECG 300A®/ MAC600®	ECG protocol followed	all	LVH using the Sokolow-Lyon criteria	x	x	x
CBC	BC®	Venous blood draw	all	Total cell count, HB, MCV	x		x
24 HR URINE	Cobas Roche®	urine using small bucket.	Select patients for 24 h proteinuria, salt and feasibility.	Volume (mls) Protein (mg) Creatinine (mmol/L) Na^+ mmol/L			x
Malaria screen	Malaria RDT®	Done in the community	all	Malaria infection	x		x
CRP	Cobas® & BC®	Blood draw	all	CRP	x		x
Hepatitis B	Cobas® ABBOTT (i1000SR)	Blood draw	all	HepB SAg	x	x	x
Hepatitis C	Cobas® (Uganda), HCV antibody rapid test (Malawi)	Blood draw	all	HepC Ab	x		x
ASOT	Cobas® & BC®	Blood draw	All	> 300			x
HIV screen	Alere Dertemine-Stat-pak-Bio line & Abbott Determine	MOH serial testing algorithm	all	HIV status	x	x	x
Schistosomiasis	Microscopy	Examination of centrifuged urine	all	Schistosoma hematobium	x	x	x
Microalbuminuria	CLINITEK® + analyser & Cobas® & BC®	ACR	all	Urine ACR	x	x	x
Urine analysis	Clinitek® & UroColor	Early morning urine	all	Protein, blood, glucose, WBCs, QuickVue Hcg Urine for pregnancy	x	x	x
Lipid profile	Cobas® & BC®	Blood draw	all	Cholesterol, LDL,	x		x

Table 1 Clinical and Laboratory measurements within the ARK study, by country *(Continued)*

Type of information	Equipment	Field procedures	Participant type	Definition of variable	Site participation		
					Malawi	S. Africa	Uganda
				HDL, TGs			
RBS	BC®	Point of care testing (Uganda $ SA)	all	Diabetes > 11.1 or on medical treatment	x	x	x
HBA1C	Cobas® & BC®	Blood draw	All	6.5% or on treatment	x		x
Creatinine	Cobas® & BC®	Jaffe and IDMS	all	For eGFR estimation	x	x	x
Cystatin-C	Cobas®		all	For eGFR estimation	x	x	x
Iohexol	Omnipaque 300 mg I/ml Healthcare® HPLC®	Clinic based blood draws at 5, 120, 180 and 240 min after administration	all	Measured GFR	x	x	x
Iohexol, DBS	DBS®	DBS by finger prick at 120 and 240 min	300 selected by eGFR	Validation of measured GFR technique			x
Aldosterone/Renin	Cobas Roche®		Selected population		x		x

All biochemistry was tested using Cobas equipment in Uganda and South Africa and Beckman Coulter equipment in Malawi
Abbreviations: *ABPM* Ambulatory Blood Pressure Monitor, *ACR* Albumin Creatinine Ratio, *ASOT* Antistreptococcal antibody titres, *BC®* Beckman Coulter®, *BIA* Bioimpedance Analysis, *BP* Blood Pressure, *CBC* Cell blood count, *CKD* Chronic Kidney Disease, *CRP* C-Reactive protein, *CIMT* Carotid intima-medial thickness, *DBS* Dry blood sample, *DM* Diabetes Mellitus, *ECG* Electrocardiography, *eGFR* Estimated glomerular filtration rate, *HB* Haemoglobin, *HTN* Hypertension, *HPLC* High liquid pressure chromatography, *IDMS* Isotope dilution mass spectrometry, *LVH* Left ventricular hypertrophy, *MCV* Mean corpuscular volume, *MOH* Ministry of Health, *RBS* Random blood sugar, *SA* South Africa, *TB* Tuberculosis

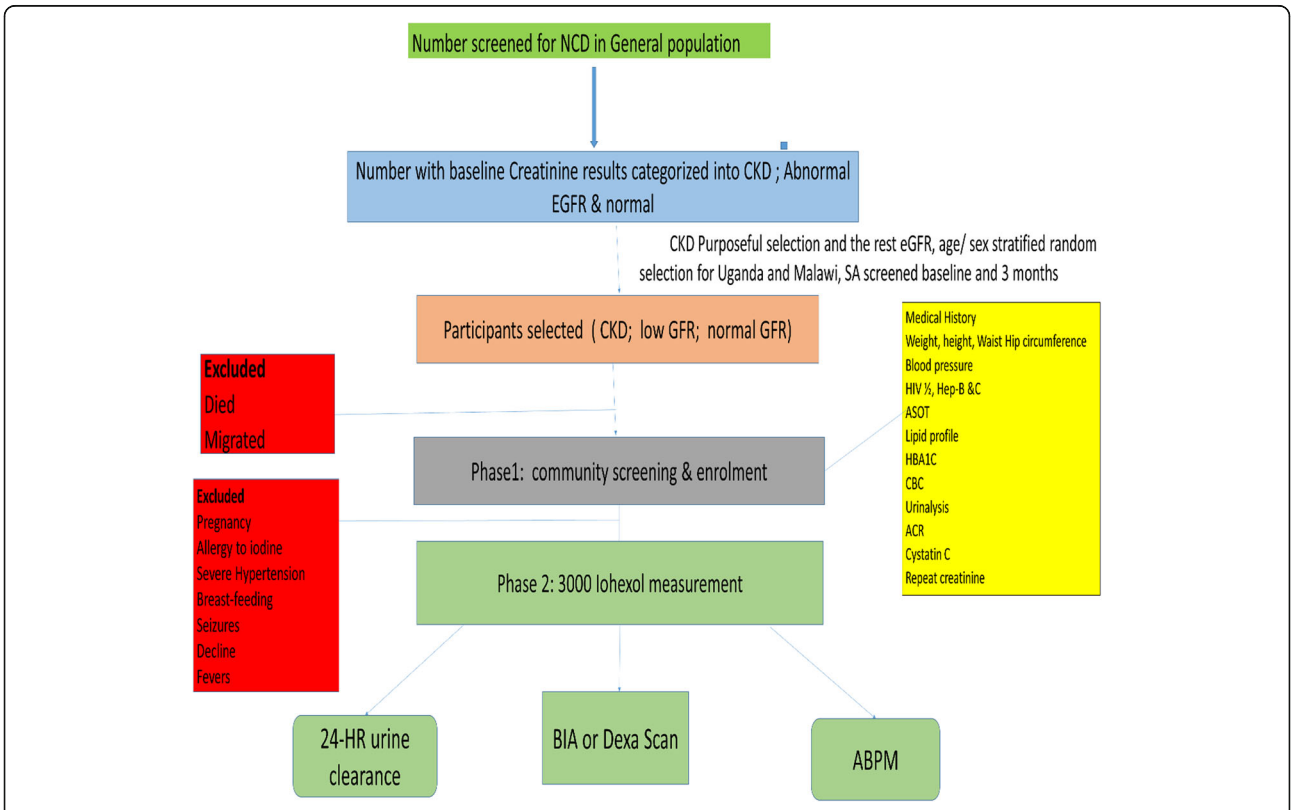


Fig. 2 Recruitment flowchart for ARK. ABPM- Ambulatory Blood Pressure Monitor; ACR- Albumin: creatinine Ratio; ASOT- Antistreptococcal antibody titres; BIA- Bioimpedance Analysis; BP-Blood Pressure; CBC- Cell blood count; GFR- glomerular filtration rate; CKD- Chronic kidney disease; HB-Haemoglobin; NCD-Non-communicable diseases; SA-South Africa. Credit for the ARK study map and copyright go to Helmut Kraus

- Able to give informed consent

Exclusion criteria

- Blood pressure > 180/110mmhg
- Pregnancy
- Breast-feeding mothers
- Known allergy to iodine–containing substances
- Uncontrolled seizures (defined as a seizure within the last 12 months).
- Acute febrile illness

Demographic factors

We will collect data on demographic factors including age, sex, place of residence, education, occupation and livelihood, tobacco use, alcohol use, dietary history, physical activity. We will also take a medical history and treatment for chronic diseases including HIV, diabetes mellitus, hypertension, heart disease and CKD as well as previous and current use of traditional medicine and drugs.

Physical examination

We will measure height, weight, waist circumference and hip circumference and calculate the body mass index (BMI) and the waist hip ratio accordingly. We will classify BMI according to WHO categories (weight/height^2: kg/m^2): underweight (< 18.5 kg/m^2), normal weight (18.5–24.99 kg/m^2), overweight (25.0–29.99 kg/m^2) and obese (> 30.0 kg/m^2).

We will undertake cardiovascular assessment through blood pressure and 24-h ambulatory blood pressure (BP) measurements (ABPM) on a sub-sample of participants (Malawi and Uganda), and electrocardiography (ECG). We will measure BP using Omron® M6 (for small, medium and large participants) and Omron HBP 110 machines (for obese participants). We will measure BP in triplicate after at least 5 min of rest and take the mean of the last two readings as the true blood pressure. We will derive BP classification from the National Institute of Health guidelines: where participants with a systolic BP greater than 120 mmHg but less than 140 mmHg, and/or a diastolic BP greater than 80 mmHg but less than 90 mmHg will be classified as pre-hypertensive. We defined hypertension as having a diastolic BP greater than or equal to 90 mmHg, systolic BP greater than or equal to 140 mmHg or being on treatment for high blood pressure. 24-h ambulatory blood pressure will be undertaken on a selected number of participants with no hypertension, pre-hypertension and hypertension across the spectrum of eGFR ranges and will capture wake and sleep periods. We will use the ECG for assessment of LVH using the Sokolow-Lyon criteria [29].

We will perform bioimpedance analysis (BIA) using the Bodystat® machine to measure body fat in relation to lean body mass for parameters outlined in Table 1. We will also perform a Dual-energy X-ray absorptiometry (DXA) scan for body composition in South Africa and use it to examine validity of the BIA measurements across countries.

Laboratory investigations

The key laboratory measurement for this study will be iohexol (Omnipaque 300 and 350 mg I/mL, GE Healthcare) clearance. The iohexol measured GFR will serve as the gold standard for comparison with other methods.

Using aseptic technique study nurses will insert an IV line into the non-dominant arm of the participant and use the line for subsequent blood draws. We will administer a single dose of 5 ml of iohexol over 30 s through an intravenous cannula inserted in the dominant arm contralateral to the established IV line, followed by a 10 ml normal saline flush. We will weigh the iohexol dose in grams to a specificity of 0.01 g by weighing the syringe before and after administration of the iohexol. The research nurses will draw four milliliters of blood for the iohexol assay again at approximately 5, 120, 180 and 240 min after injection of iohexol, and record the exact time. We will calculate the GFR using a single compartment model based on iohexol clearance between 120 and 240 min.

In Uganda at the time points of 120 and 240 min, we will also collect a capillary sample of dried blood spots to compare to intravenous sampling of iohexol GFR as a modification from previous studies [11, 30]. Should this method prove to be accurate, it will greatly facilitate further studies in resource-limited settings. In order to avoid inter-laboratory variations, all plasma and blood spot samples of iohexol will be measured at the National Health Laboratory, Johannesburg, South Africa by ultra-pressure liquid chromatography-tandem mass spectrometry with identification and quantitation based on MRM transitions. The laboratory is accredited for the International Organization for Standardization (ISO) standard 15,189 and also participates in the internal quality control run at a high and a low level every 20 samples. The laboratory has participated in Equalis external control since 2017.

We will measure creatinine using the enzymatic method in Uganda and Jaffe for South Africa and Malawi standardized against an isotope dilution mass spectrometry (IDMS) method across all sites. We will adopt the recommended estimation and reporting format or measured renal function proposed by Fabian et al [31]. We will measure serum cystatin C to determine the accuracy of the cystatin-based CKD-EPI equations. We will measure Cystatin C using the standardized Roche Gen2 assay, which has been standardized against certified reference material (ERM-DA471/IFCC). To minimise batch variation, samples will be

analysed at the end of the study at one central laboratory at MRC/UCRI&LSHTM for Uganda and Malawi while the samples in South Africa will be analysed at the National Health Laboratory, Johannesburg, South Africa. To understand laboratory differences and enable calibration we will cross-validate and measure 50-paired samples for creatinine in Uganda, Malawi and South Africa, and cystatin between South Africa and Uganda. Additional tests include full blood count (including haemoglobin level), tests for screening of infections (malaria, streptococcal infections, Hepatitis B, Hepatitis C and HIV Infection), C-reactive protein, random blood glucose, HbA1c and the fasting lipid profile as markers of metabolic risk.

We will perform four urine tests. We will take an early morning spot urine sample for quantitative urine albumin: creatinine ratio, an early morning urine dipstick analysis to qualitatively assess haematuria, leucocyturia and proteinuria, and microscopy will be performed on the centrifuged urine to screen for urinary schistosomiasis within 1–2 h of collection. We will take a 24-h urine collection for urinary albumin excretion, sodium and 24-h creatinine clearance in a sample of 300 participants from Uganda during the second phase of the study.

Details of the equipment used for the measurements are outlined in Table 1 below.

Qualitative sub-study in Uganda

The qualitative sub-study in Uganda will have two components: a study of people with abnormal renal function, and an enquiry into medicinal product supply and use sourced from traditional healers and herbalists.

This sub-study of people with abnormal renal function will include up to 50 participants found to have CKD during the recruitment of the 1000 participants for the overarching study. In order to recruit a comparator group for the 50 CKD participants, a further 50 participants without CKD of a similar age living in a neighboring house to each case will be invited to take part. Data from both cases and 'controls' will be collected through two in-depth interviews. These selections were made on an estimated sample to reach saturation.

For the sub-study with traditional healers/herbalists, we will approach ten traditional healers and herbalists and invite them to participate in one or more interviews to discuss the common ailments they treat and the types of herbal and other preparations that they use in their practice.

Bio-banking

Serum, plasma and urine samples will be collected, processed and frozen and transported to the central laboratory for additional tests that cannot be conducted at the originating laboratory. All three centers have quality assurance systems for processing, sample transportation and sample storage management systems [26, 32, 33]. All sites have received ethical approval for testing of specimens for genetic studies and exploration of new markers of CKD.

Data management

Data management will follow local established procedures ensuring quality, confidentiality and proper use of abnormal results to guide patient care when needed. Electronic data capture allows automatic data checks such as double entry of numbers and range checks. We will enter anonymized data into compatible statistical databases from centres in all three countries to for sharing between all investigators.

Quality assurance

All three centres have quality assurance procedures that support the studies at different levels. Study staff are trained and certified according to national and international guidelines governing research. The ARK team holds regular meetings to ensure that study protocols and research standards are adhered to. We will follow standard sample storage protocols for the biorepository.

Statistical and qualitative analysis plan

We will tabulate baseline characteristics stratified by country. We will perform initial analyses in a sample of 1500 participants selected to represent all centres, ages, sex, and eGFR categories for modification of the equation, and allow for subsequent validation using the remaining data. We plan to recombine the training and test sets for fitting a final model. We are also open to alternatives to data-splitting such as re-sampling techniques (cross-validation and bootstrapping), that would allow us to develop the model on the whole dataset and still validate predictive accuracy. Evidence from previous derivation of eGFR equations suggests that errors are multiplicative, so we will do analysis on the log (ioxehol-GFR) scale. To determine the measured GFR (iohexol-GFR); we will calculate the slope from the 3 samples at 120, 180 and 240 min points using the exact time of collection turn these into a GFR normalised to BSA using standard methods [13, 34]. We will also put in the R-value of the fit, which will be used to exclude ones where the fit is particularly poor. In our primary analysis we will evaluate the performance of the CockCroft Gault, 4v-MDRD, CKD-EPIcr, CKD-EPIcyst, CKD-EPIcr/cyst equations with and without ethnic correction factor by calculating the bias, precision and accuracy at 10% (P10) and 30% (P30) compared to iohexol GFR as done by previous studies [13]. Derivation of GFR estimating equations is rapidly developing and we will also consider evaluation against newly developed methods including the full age spectrum (FAS) and the Lund-

Malmo GFR estimating equations [35, 36]. We will assess the performance of the different equations for eGFRs with and without ethnic factors using Lin's Concordance Correlation Coefficient and Spearman correlation coefficients to evaluate the degree to which pairs of observations fall on the 45° line through the origin. We will calculate bias and relative bias. We will evaluate precision by the standard deviation of the bias (random error) and root-mean-square error. We will test the difference in P30 accuracy between eGFRs with the exact McNemar test.

We will use Bland-Altman plots to investigate the measurement error of existing equations of log(MDRD-eGFR), log(CKD-EPI-eGFR), log(CG-eGFR) and log(cystatin C-eGFR) compared to log(ioxehol-GFR). If needed, to develop modified correction factors for the setting, we will carry out regressions to predict log(ioxehol-GFR) as a function of age, gender, and creatinine/cystatin C. Predicted eGFRs will subsequently be compared with measured GFRs in the validation sample. We will investigate what happens if information on weight, height, and bioimpedance measures are added to the regression model and use computed R^2 values to investigate whether anthropometry adds relevant information over and above existing equations.

All statistical analyses will be performed using STATA 15 SE (Stata Corp, Texas, USA).

For the qualitative study, we will record interviews with permission from the participant and transcribe and translate them. We will revise the interview guidance after the initial interviews in order to collect greater depth of data on emerging themes. We will review transcripts for accuracy and enter these into Dedoose, qualitative data analysis software. We will conduct data analysis using an iterative coding process, during the interview period. After data collection, we will perform open coding, and move to more refined codes, subthemes, and then themes. We will analyze themes and code them using a constant comparison method. We will apply the codes to interview transcripts and summary statements with representative quotes developed for each theme.

Ethical considerations

We have obtained ethical approvals from the Uganda Virus Research Institute, Research Ethics Committee (UVRI-REC-#HS 1978) the Uganda National Council for Science and Technology (UNCST-#SS 4283), from the Malawi National Health Sciences Research Committee (#1072) and the University of Witwatersrand Human Research Ethics Committee (#M160938).

We will obtain written informed consent from participants for the collection and analysis of genetic samples as well as iohexol testing and for the use of their clinical records for research purposes. We will seek approval for all study procedures including material transfer agreements from the respective ethical review boards. In case we diagnose medical conditions through study screening, we will refer the participants to appropriate medical facilities. We will refer participants found to have advanced CKD to appropriate nephrology services as directed by the study physicians in each centre.

Discussion

This study aims to establish the optimal approach to estimate GFR in sub-Saharan Africa. Our study has the advantage of drawing participants from three countries and both rural and urban settings within sub-Saharan Africa, which will increase the applicability of the findings across the region. Furthermore, we embedded the ARK study within established cohorts that have background data and longitudinal serial measures that we can use to characterize kidney disease over a period. The study will seek to overcome a number of the limitations of previous research.

There are challenges with measuring iohexol GFR, with the major one being accuracy of measurement of iohexol [23, 37]. Stringent measures, shared standard operating procedures across the centres, and analysis in one certified laboratory will help minimise errors.

Measurement of creatinine is another challenge. Creatinine is influenced by many factors and these can greatly influence the estimation, particularly at low eGFR [19, 37, 38]. It is particularly important that the methods used for creatinine measurements are the same across sites and are validated across the centres. We opted to use the Jaffe method for South Africa and Malawi and enzymatic method in Uganda, standardized against an isotope dilution mass spectrometry method across all sites to improve the accuracy in measurement. We will measure Cystatin C in order to assess whether the addition of cystatin C by itself or in combination with creatinine assays further refines the accuracy of estimating GFR within SSA. Several studies have shown that the addition of cystatin C to the estimation formula improves its accuracy [39, 40] and using CKD-EPI-creatinine/cystatin C improved the accuracy in a study from Congo [13]. Using a unified approach, these issues will be addressed by this study. We also plan to measure bioimpedance to define body composition and its contribution to measured GFR variation across the sites, and to assess the correlation of body composition measured by BIA with the gold standard, DXA – which will be done in South Africa [41]. Bioimpedance measurement was included in this study for measurement of lean mass and total body water and we hope to use it to optimize GFR estimation on an individual patient basis, for example in a patient with significant oedema. While this

has not proven useful in high-income settings, such technologies may be useful in low-resource settings [42]. Bio-impedance will help in incorporating more parameters allowing more flexibility to work with the data to get the best formula.

Assessment of risk factors for CKD in sub-Saharan Africa has previously been hampered by inaccurate measurement of GFR. Although, not exhaustive, we included a number of risk factors for CKD in our study. We will examine the role of traditional cardiovascular risk factors including diabetes mellitus, obesity and lipid profiles along with conventional and ambulatory blood pressure monitoring. In addition we will measure the impact of infections which may be of great importance in this region where infectious diseases are common and may play a major role in disease causation alongside genetic factors [43–49].

However, there are limitations. Prolonged blood sampling for iohexol excretion may help to define GFR in patients with very poor renal function [23]. We chose not to include this in our protocol to minimize participant burden. In addition, we do not include participants from other regions of sub-Saharan Africa, including Central and West Africa but we welcome data-sharing with studies in other regions.

The African Research into Kidney disease (ARK) collaboration will provide definitive information about optimal measurement of renal function in sub-Saharan Africa, and is an opportunity for establishing close working relationships across different centres, using standardized protocols and measurements, and shared biorepositories. We plan to build on the collaboration for this study and for future work on kidney disease in sub-Saharan Africa, and welcome additional partners from across the region including North African and Arab populations.

Abbreviations

ARK: African Research into Kidney Diseases; CKD: Chronic kidney disease; CKD-EPI: Chronic Kidney Disease-Epidemiology collaboration; DTPA: Technetium-99 m diethylenetriamine penta-acetic acid; EDTA: chromium-51 ethylene diamine tetra-acetic acid; EGFR: Estimated glomerular filtration rate; IDMS: Isotope dilution mass spectrometry; MDRD: Modification of Diet in Renal Disease; MGFR: Measured glomerular filtration rate

Acknowledgements

We would like to thank all the participants for the primary studies as well as the staff for the ARK cohorts. We will like to thank Professor Anatoli Kamali for his involvement in the initial stages of setting up the study.

Authors' contributions

RK, JN, RN, DN, MN, LS, SN, ACC and LAT conceived the study, designed the study, drafted work and substantially revised the manuscript. CH, WN, BS, JP, JG, ANW, JS designed the study, drafted work and substantially revised the manuscript. All authors have approved submission of the manuscript and subscribe to the contents as accurate.

Funding

The study in Uganda and Malawi (RK) is funded by the GlaxoSmithKline Africa Non-Communicable Disease (NCD) Open Lab (Project Number: 8111). The study in South Africa (SN) is funded by the South African Medical Research Council, MRC UK (via the Newton Fund) and the GlaxoSmithKline Africa NCD Open Lab (Project Number: 074). LAT is funded by a Wellcome Trust Intermediate Clinical Fellowship (101143/Z/13/Z). The original cohort studies at MEIRU, Malawi were funded by the Wellcome Trust. ANW is supported by the Fogarty International Centre (NIH) under Award Number K43TW010698. This paper describes the views of the authors and does not necessarily represent the official views of the National Institutes of Health (USA). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Availability of data and materials

Not applicable

Ethics approval and consent to participate

We have obtained ethical approvals from the Uganda Virus Research Institute, Research Ethics Committee (UVRI-REC) the Uganda National Council for Science and Technology (UNCST), from the Malawi National Health Sciences Research Committee and the University of Witwatersrand Human Research Ethics Committee (Medical).
We will obtain written informed consent from participants for the collection and analysis of genetic samples as well as iohexol testing and for the use of their clinical records for research purposes. We will seek approval for all study procedures including material transfer agreements from the respective ethical review boards. In case we diagnose medical conditions through study screening, we will refer the participants to appropriate medical facilities. We will refer participants found to have advanced CKD to appropriate nephrology services as directed by the study physicians in each centre.

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Medical Research Council/ UVRI & London School of Hygiene and Tropical Medicine Research Unit, Entebbe, Uganda. [2]Makerere University College of Health Sciences, Kampala, Uganda. [3]London School of Hygiene & Tropical Medicine, London, UK. [4]Wits Donald Gordon Medical Centre, Parktown, Johannesburg, South Africa. [5]Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa. [6]Malawi Epidemiology and Intervention Research Unit, Lilongwe, Malawi. [7]Department of Chemical Pathology, National Health Laboratory Services, University of Witwatersrand, Johannesburg, South Africa. [8]Department of Internal Medicine, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa.

Received: 17 May 2019 Accepted: 8 January 2020
Published online: 15 January 2020

References

1. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, Saran R, Wang AY, Yang CW. Chronic kidney disease: global dimension and perspectives. Lancet. 2013;382(9888):260–72.
2. Stanifer JW, Jing B, Tolan S, Helmke N, Mukerjee R, Naicker S, Patel U. The epidemiology of chronic kidney disease in sub-Saharan Africa: a systematic review and meta-analysis. Lancet Glob Health. 2014;2(3):e174–81.
3. Ji M, Lee YH, Hur M, Kim H, Cho HI, Yang HS, Navarin S, Di Somma S, Network G. Comparing results of five glomerular filtration rate-estimating equations in the Korean general population: MDRD study, revised Lund-Malmo, and three CKD-EPI equations. Ann Lab Med. 2016;36(6):521–8.
4. Feng JF, Qiu L, Zhang L, Li XM, Yang YW, Zeng P, Guo XZ, Qin Y, Liu HC, Han XM, et al. Multicenter study of creatinine- and/or cystatin C-based equations for estimation of glomerular filtration rates in Chinese patients with chronic kidney disease. PLoS One. 2013;8(3):e57240.

5. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron. 1976;16(1):31–41.
6. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of diet in renal disease study group. Ann Intern Med. 1999;130(6):461–70.
7. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, Kusek JW, Eggers P, Van Lente F, Greene T, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604–12.
8. Shemesh O, Golbetz H, Kriss JP, Myers BD. Limitations of creatinine as a filtration marker in glomerulopathic patients. Kidney Int. 1985;28(5):830–8.
9. Delanaye P, Mariat C, Maillard N, Krzesinski JM, Cavalier E. Are the creatinine-based equations accurate to estimate glomerular filtration rate in African American populations? Clin J Am Soc Nephrol. 2011;6(4):906–12.
10. van Deventer HE, George JA, Paiker JE, Becker PJ, Katz IJ. Estimating glomerular filtration rate in black south Africans by use of the modification of diet in renal disease and Cockcroft-Gault equations. Clin Chem. 2008;54(7):1197–202.
11. Wyatt CM, Schwartz GJ, Owino Ong'or W, Abuya J, Abraham AG, Mboku C, M'Mene LB, Koima WJ, Hotta M, Maier P, et al. Estimating kidney function in HIV-infected adults in Kenya: comparison to a direct measure of glomerular filtration rate by iohexol clearance. PLoS One. 2013;8(8):e69601.
12. Madala ND, Nkwanyana N, Dubula T, Naiker IP. Predictive performance of eGFR equations in south Africans of African and Indian ancestry compared with (9)(9)mTc-DTPA imaging. Int Urol Nephrol. 2012;44(3):847–55.
13. Bukabau JB, Sumaili EK, Cavalier E, Pottel H, Kifakiou B, Nkodila A, Makulo JRR, Mokoli VM, Zinga CV, Longo AL, et al. Performance of glomerular filtration rate estimation equations in Congolese healthy adults: the inopportunity of the ethnic correction. PLoS One. 2018;13(3):e0193384.
14. Kaze FF, Halle MP, Mopa HT, Ashuntantang G, Fouda H, Ngogang J, Kengne AP. Prevalence and risk factors of chronic kidney disease in urban adult Cameroonians according to three common estimators of the glomerular filtration rate: a cross-sectional study. BMC Nephrol. 2015;16:96.
15. Myers GL, Miller WG, Coresh J, Fleming J, Greenberg N, Greene T, Hostetter T, Levey AS, Panteghini M, Welch M, et al. Recommendations for improving serum creatinine measurement: a report from the laboratory working Group of the National Kidney Disease Education Program. Clin Chem. 2006;52(1):5–18.
16. Stevens LA, Manzi J, Levey AS, Chen J, Deysher AE, Greene T, Poggio ED, Schmid CH, Steffes MW, Zhang YL, et al. Impact of creatinine calibration on performance of GFR estimating equations in a pooled individual patient database. Am J Kidney Dis. 2007;50(1):21–35.
17. Murthy K, Stevens LA, Stark PC, Levey AS. Variation in the serum creatinine assay calibration: a practical application to glomerular filtration rate estimation. Kidney Int. 2005;68(4):1884–7.
18. Verhave JC, Balje-Volkers CP, Hillege HL, de Zeeuw D, de Jong PE. The reliability of different formulae to predict creatinine clearance. J Intern Med. 2003;253(5):563–73.
19. Coresh J, Astor BC, McQuillan G, Kusek J, Greene T, Van Lente F, Levey AS. Calibration and random variation of the serum creatinine assay as critical elements of using equations to estimate glomerular filtration rate. Am J Kidney Dis. 2002;39(5):920–9.
20. Levey AS, Inker LA. Assessment of glomerular filtration rate in health and disease: a state of the art review. Clin Pharmacol Ther. 2017;102(3):405–19.
21. George C, Mogueo A, Okpechi I, Echouffo-Tcheugui JB, Kengne AP. Chronic kidney disease in low-income to middle-income countries: the case for increased screening. BMJ Glob Health. 2017;2(2):e000256.
22. Vega JA, Ochola PS, Marsh ML, HO VQ, Fisher T. Comparison of 24-hourour urine to estimated renal function using CKD-EPI, MDRD4 and cockcroft-gault in specific patient subsets. J Pharma Care Health Sys. 2015;2(129):2376.
23. Delanaye P, Ebert N, Melsom T, Gaspari F, Mariat C, Cavalier E, Bjork J, Christensson A, Nyman U, Porrini E, et al. Iohexol plasma clearance for measuring glomerular filtration rate in clinical practice and research: a review. Part 1: how to measure glomerular filtration rate with iohexol? Clin Kidney J. 2016;9(5):682–99.
24. Sagou Yayo E, Aye M, Konan JL, Emieme A, Attoungbre ML, Gnionsahe A, Cavalier E, Monnet D, Delanaye P. Inadequacy of the African-American ethnic factor to estimate glomerular filtration rate in an African general population: results from cote d'Ivoire. Nephrol Ther. 2016;12(6):454–9.
25. Bukabau JB, Yayo E, Gnionsahe A, Monnet D, Pottel H, Cavalier E, Nkodila A, Makulo JRR, Mokoli VM, Lepira FB, et al. Performance of creatinine- or cystatin C-based equations to estimate glomerular filtration rate in sub-Saharan African populations. Kidney Int. 2019;95(5):1181–9.
26. Asiki G, Murphy G, Nakiyingi-Miiro J, Seeley J, Nsubuga RN, Karabarinde A, Waswa L, Biraro S, Kasamba I, Pomilla C, et al. The general population cohort in rural South-Western Uganda: a platform for communicable and non-communicable disease studies. Int J Epidemiol. 2013;42(1):129–41.
27. Kalyesubula R, Hau JP, Asiki G, Ssebunya B, Kusemererwa S, Seeley J, Smeeth L, Tomlinson L, Newton R. Impaired renal function in a rural Ugandan population cohort. Wellcome Open Res. 2018;3:149.
28. Nakanga WP, Prynn JE, Banda L, Kalyesubula R, Tomlinson LA, Nyirenda M, Crampin AC. Prevalence of impaired renal function among rural and urban populations: findings of a cross-sectional study in Malawi. Wellcome Open Res. 2019;4:92.
29. Hancock EW, Deal BJ, Mirvis DM, Okin P, Kligfield P, Gettes LS, Bailey JJ, Childers R, Gorgels A, Josephson M, et al. AHA/ACCF/HRS recommendations for the standardization and interpretation of the electrocardiogram: part V: electrocardiogram changes associated with cardiac chamber hypertrophy: a scientific statement from the American Heart Association Electrocardiography and Arrhythmias Committee, Council on Clinical Cardiology; the American College of Cardiology Foundation; and the Heart Rhythm Society. Endorsed by the International Society for Computerized Electrocardiology. J Am Coll Cardiol. 2009;53(11):992–1002.
30. Niculescu-Duvaz I, D'Mello L, Maan Z, Barron JL, Newman DJ, Dockrell ME, Kwan JT. Development of an outpatient finger-prick glomerular filtration rate procedure suitable for epidemiological studies. Kidney Int. 2006;69(7):1272–5.
31. Fabian J, George JA, Etheredge HR, Deventer M, Kalyesubula R, Wade AN, Tomlinson LA, Tollman S, Naicker S, African Research in Kidney Disease ARK. Methods and reporting of kidney function: a systematic review of studies from sub-Saharan Africa. Clin Kidney J. 2019;12:778.
32. Crampin AC, Kayuni N, Amberbir A, Musicha C, Koole O, Tafatatha T, Branson K, Saul J, Mwaiyeghele E, Nkhwazi L, et al. Hypertension and diabetes in Africa: design and implementation of a large population-based study of burden and risk factors in rural and urban Malawi. Emerg Themes Epidemiol. 2016;13:3.
33. Kahn K, Collinson MA, Gomez-Olive FX, Mokoena O, Twine R, Mee P, Afolabi SA, Clark BD, Kabudula CW, Khosa A, et al. Profile: Agincourt health and socio-demographic surveillance system. Int J Epidemiol. 2012;41(4):988–1001.
34. Peters AM. Expressing glomerular filtration rate in terms of extracellular fluid volume. Nephrol Dial Transplant. 1992;7(3):205–10.
35. Delanaye P, Ebert N, Pottel H. Newer GFR estimating equations require validation in different populations. Am J Kidney Dis. 2017;70(4):586.
36. Bjork J, Nyman U, Berg U, Delanaye P, Dubourg L, Goffin K, Grubb A, Hansson M, Littmann K, Asling-Monemi K, et al. Validation of standardized creatinine and cystatin C GFR estimating equations in a large multicentre European cohort of children. Pediatr Nephrol. 2019;34(6):1087–98.
37. Seegmiller JC, Eckfeldt JH, Lieske JC. Challenges in measuring glomerular filtration rate: a clinical laboratory perspective. Adv Chronic Kidney Dis. 2018;25(1):84–92.
38. Seronie-Vivien S, Galteau MM, Carlier MC, Hadj-Aissa A, Hanser AM, Hym B, Marchal A, Michotey O, Pouteil-Noble C, Sternberg M, et al. Improving the interlaboratory variation for creatinine serum assay. Ann Biol Clin. 2004;62(2):165–75.
39. Werner K, Pihlsgard M, Elmstahl S, Legrand H, Nyman U, Christensson A. Combining cystatin C and creatinine yields a reliable glomerular filtration rate estimation in older adults in contrast to beta-trace protein and beta2-microglobulin. Nephron. 2017;137(1):29–37.
40. Bjork J, Grubb A, Larsson A, Hansson LO, Flodin M, Sterner G, Lindstrom V, Nyman U. Accuracy of GFR estimating equations combining standardized cystatin C and creatinine assays: a cross-sectional study in Sweden. Clin Chem Lab Med. 2015;53(3):403–14.
41. Ozmen S, Kaplan MA, Kaya H, Akin D, Danis R, Kizilkan B, Yazanel O. Role of lean body mass for estimation of glomerular filtration rate in patients with chronic kidney disease with various body mass indices. Scand J Urol Nephrol. 2009;43(2):171–6.
42. Frank M, Guarino-Gubler S, Burnier M, Maillard M, Keller F, Gabutti L. Estimation of glomerular filtration rate in hospitalised patients: are we overestimating renal function? Swiss Med Wkly. 2012;142:w13708.
43. Ulasi II, Tzur S, Wasser WG, Shemer R, Kruzel E, Feigin E, Ijoma CK, Onodugo OD, Okoye JU, Arodiwe EB, et al. High population frequencies of APOL1 risk variants are associated with increased prevalence of non-diabetic chronic kidney disease in the Igbo people from South-Eastern Nigeria. Nephron Clin Pract. 2013;123(1–2):123–8.

44. Stanifer JW, Karia F, Maro V, Kilonzo K, Qin X, Patel UD, Hauser ER. APOL1 risk alleles among individuals with CKD in northern Tanzania: a pilot study. PLoS One. 2017;12(7):e0181811.
45. Osafo C, Raji YR, Olanrewaju T, Mamven M, Arogundade F, Ajayi S, Ulasi I, Salako B, Plange-Rhule J, Mengistu Y, et al. Genomic approaches to the burden of kidney disease in sub-Saharan Africa: the human heredity and health in Africa (H3Africa) kidney disease research network. Kidney Int. 2016; 90(1):2–5.
46. Kasembeli AN, Duarte R, Ramsay M, Mosiane P, Dickens C, Dix-Peek T, Limou S, Sezgin E, Nelson GW, Fogo AB, et al. APOL1 risk variants are strongly associated with HIV-associated nephropathy in black south Africans. J Am Soc Nephrol. 2015;26(11):2882–90.
47. Okunola O, Akinsola A, Ayodele O. Kidney diseases in Africa: aetiological considerations, peculiarities and burden. Afr J Med Med Sci. 2012;41(2):119–33.
48. Arogundade FA, Hassan MO, Omotoso BA, Oguntola SO, Okunola OO, Sanusi AA, Akinsola A. Spectrum of kidney diseases in Africa: malaria, schistosomiasis, sickle cell disease, and toxins. Clin Nephrol. 2016;86(13):53–60.
49. Okpechi IG, Ameh OI, Bello AK, Ronco P, Swanepoel CR, Kengne AP. Epidemiology of histologically proven glomerulonephritis in Africa: a systematic review and meta-analysis. PLoS One. 2016;11(3):e0152203.

Publisher's Note

The NEW ENGLAND JOURNAL of MEDICINE

Please review the Supplemental Files folder to review documents not compiled in the PDF.

Measurement of kidney function in Sub-Saharan Africa

Journal:	*New England Journal of Medicine*
Manuscript ID	Draft
Article Type:	Original Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Fabian, June; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt); University of the Witwatersrand Faculty of Health Sciences, Wits Donald Gordon Medical Centre Kalyesubula, Robert; MRC/UVRI and LSHTM Uganda Research Unit; London School of Hygiene and Tropical Medicine Department of Non-communicable Disease Epidemiology Mkandawire, Joseph; Malawi Epidemiology and Intervention Research Unit; London School of Hygiene and Tropical Medicine Department of Non-communicable Disease Epidemiology Holm Hansen, Christian; London School of Hygiene & Tropical Medicine, MRC International Statistics and Epidemiology Group Nitsch, Dorothea; London School of Hygiene and Tropical Medicine Department of Non-communicable Disease Epidemiology Musenge, Eustasius; University of the Witwatersrand School of Public Health Nakanga, Wisdom ; Malawi Epidemiology and Intervention Research Unit; London School of Hygiene and Tropical Medicine Department of Non-communicable Disease Epidemiology Prynn, Josephine ; Malawi Epidemiology and Intervention Research Unit; University College London Institute of Cardiovascular Science Dreyer, Gavin; Barts Health NHS Trust, Department of Nephrology Snyman, Tracy; National Health Laboratory Service, Department of Chemical Pathology Ssebunya, Billy; MRC/UVRI and LSHTM Uganda Research Unit Ramsay, Michele ; University of the Witwatersrand, Sydney Brenner Institute for Molecular Bioscience; National Health Laboratory Service, Division of Human Genetics, School of Pathology Smeeth, Liam; London School of Hygiene and Tropical Medicine Department of Non-communicable Disease Epidemiology Tollman, Stephen; University of the Witwatersrand School of Public Health, Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt) Naicker, Saraladevi; University of the Witwatersrand, Department of Internal Medicine Crampin, Amelia; Malawi Epidemiology and Intervention Research Unit Newton, Robert; MRC/UVRI and LSHTM Uganda Research Unit; University of York Department of Health Sciences George, Jaya; National Health Laboratory Service, Department of Chemical Pathology

	Tomlinson, Laurie; London School of Hygiene and Tropical Medicine Department of Non-communicable Disease Epidemiology
Abstract:	Background The burden of kidney disease in Sub-Saharan Africa (SSA) is unknown. Equations used to estimate kidney function from serum creatinine have limited regional validation. We sought to determine the most accurate way to measure kidney function and thus estimate the prevalence of impaired kidney function in SSA. Methods We measured serum creatinine, cystatin C, and glomerular filtration rate using the slope-intercept method for iohexol plasma clearance (mGFR). We compared performance of creatinine and cystatin C-based estimating equations to mGFR, modelled and validated a new equation, and used multiple imputation based on mGFR to predict the population prevalence of impaired kidney function. Results Among 2578 included participants, creatinine-based equations overestimated kidney function compared to mGFR. The greatest bias occurred at low kidney function so that the proportion with impaired kidney function (GFR <60ml/min/1.73m2) by mGFR or estimated by cystatin C was more than double that estimated from creatinine. A new creatinine-based equation did not outperform existing equations, and no equation estimated GFR within ±30% of mGFR for 75% or more participants. Using a model to impute kidney function based on mGFR, the estimated prevalence of impaired kidney function was two- to three-fold higher than creatinine-based estimates in populations across six SSA countries. Conclusion Estimating GFR using serum creatinine substantially underestimates the individual and population-level burden of impaired kidney function in SSA with implications for understanding disease progression and complications, clinical care and service provision. Scalable and affordable ways to accurately identify impaired kidney function in SSA are urgently needed.

Measurement of kidney function in Sub-Saharan Africa

Authors

June Fabian MD[1,2]*, Robert Kalyesubula MD[3,4]*, Joseph Mkandawire MD[4-6], Christian Holm Hansen PhD[7], Dorothea Nitsch MRCP[4], Eustasius Musenge PhD[8], Wisdom P Nakanga MD[3,6], Josephine E Prynn MD[6,9], Gavin Dreyer MD[10], Tracy Snyman MSc[11], Billy Ssebunnya MD[3], Michele Ramsay PhD[12,13], Liam Smeeth FRCGP[4], Stephen Tollman PhD[1,14], Saraladevi Naicker PhD[15], Amelia Crampin[6], Robert Newton PhD[3,16], Jaya A George PhD[11], Laurie Tomlinson PhD[4] on behalf of the African Research on Kidney Disease (ARK) Consortium

*Joint first authors

Author affiliations

[1] Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit (Agincourt), School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2]Wits Donald Gordon Medical Centre, School of Clinical Medicine, Faculty of Health Sciences, University of Witwatersrand, Johannesburg, South Africa

[3]MRC/UVRI & London School of Hygiene and Tropical Medicine Research Unit, Entebbe, Uganda

[4]Department of non-communicable disease epidemiology, Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

[5]Department of Surgery, Pan-African Academy of Christian Surgeons, Malamulo, Thyolo, Malawi

[6]Malawi Epidemiology and Intervention Research Unit, PO Box 148, Lilongwe, Malawi

[7]MRC International Statistics and Epidemiology Group, London School of Hygiene and Tropical Medicine, London, WC1E 7HT, UK

[8]School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[9]Institute of Cardiovascular Science, UCL, Gower Street, London WC1E 6BT, UK

[10]Department of Nephrology, Barts Health NHS Trust, London E1 1BB.

[11]Department of Chemical Pathology, National Health Laboratory Services and University of Witwatersrand, Johannesburg, South Africa

[12] Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[13] Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, South Africa

[14] International Network for the Demographic Evaluation of Populations and their Health (INDEPTH) Network, Accra, Ghana

[15]Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[16]Department of Health Sciences, University of York, UK

corresponding author: June Fabian, primary email: june.fabian@mweb.co.za; secondary email: june.fabian@wits.ac.za; mobile: +2782 785 3579; landline: +2711 356 6395

Abstract

Background

The burden of kidney disease in Sub-Saharan Africa (SSA) is unknown. Equations used to estimate kidney function from serum creatinine have limited regional validation. We sought to determine the most accurate way to measure kidney function and thus estimate the prevalence of impaired kidney function in SSA.

Methods

We measured serum creatinine, cystatin C, and glomerular filtration rate using the slope-intercept method for iohexol plasma clearance (mGFR). We compared performance of creatinine and cystatin C-based estimating equations to mGFR, modelled and validated a new equation, and used multiple imputation based on mGFR to predict the population prevalence of impaired kidney function.

Results

Among 2578 included participants, creatinine-based equations overestimated kidney function compared to mGFR. The greatest bias occurred at low kidney function so that the proportion with impaired kidney function (GFR <60ml/min/1.73m^2) by mGFR or estimated by cystatin C was more than double that estimated from creatinine. A new creatinine-based equation did not outperform existing equations, and no equation estimated GFR within ±30% of mGFR for 75% or more participants. Using a model to impute kidney function based on mGFR, the estimated prevalence of impaired kidney function was two- to three-fold higher than creatinine-based estimates in populations across six SSA countries.

Conclusion

Estimating GFR using serum creatinine substantially underestimates the individual and population-level burden of impaired kidney function in SSA with implications for understanding disease progression and complications, clinical care and service provision. Scalable and affordable ways to accurately identify impaired kidney function in SSA are urgently needed.

Introduction

The prevalence of chronic kidney disease (CKD) in Sub-Saharan Africa (SSA) is unknown. Current estimates are 11-16% in those at high risk and 3-6% in population-representative studies. There is concern that these figures are not capturing the true burden of kidney disease for several reasons.[1] Creatinine-based equations to estimate glomerular filtration rate (eGFR) were developed in high-income countries with limited validation in African populations. Direct measurement of glomerular filtration rate (mGFR) using exogenous biomarkers such as iohexol or iothalamate needed to assess the accuracy of these equations is expensive and inaccessible in many SSA countries. The most commonly used Modification of Diet in Renal Disease (MDRD) Study and Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) equations include coefficients adjusting eGFR for individuals with African ancestry (henceforward 'ethnicity coefficient'). These adjustments are based on studies showing participants with African ancestry had relatively higher mGFR for a given creatinine than other US population groups.[2] Studies in Africa suggest that ethnicity coefficients overestimate mGFR in continental Africans, possibly from smaller body surface area (BSA) and lower muscle mass compared to the African diaspora.[3-6] Additional sources of inaccuracy in determining CKD prevalence are wide biologic and analytic variation in creatinine measurement, leading to difficulties comparing results within and between countries and over time.[7,8]

Limited data on the accuracy of kidney function measurement in SSA impacts individuals with CKD, and more broadly, African populations. For an individual, missed opportunities to implement targeted interventions may result in progression to severe kidney disease with consequent premature disability or death.[9] The absence of accurate population data for prevalent CKD limits implementation of public health measures and integration of CKD into chronic disease management platforms.[10]

To address these knowledge gaps, we formed the African Research on Kidney Disease (ARK) Consortium, comprising centers of excellence from three longitudinal population studies in Malawi (urban and rural)[11] and Uganda (rural),[12] and a Health and Demographic Surveillance System site in South Africa (rural).[13] Our primary aim was to measure GFR using plasma clearance of iohexol in large, population-based samples from each country, and to compare the performance of available eGFR equations to iohexol mGFR. Our secondary aim was to model and externally validate an eGFR equation that better predicted kidney function in SSA, enabling accurate determination of the regional burden of impaired kidney function.

Methods

Study setting and sampling strategy

Our study methods are published.[14] In summary, all partners obtained ethics approval for their respective studies and all participants provided written informed consent. First, each partner measured kidney function in a large, population-representative sample..[16,17] From this, subsamples of potential adult participants for the iohexol GFR (mGFR) study were identified, with the estimated sample size (N=730) calculated for 90% power to detect whether eGFR was within 5% of mGFR at an eGFR of 60ml/min/1.73m^2, assuming a standard deviation of 25ml/min and two-sided α=0.05.[14] Samples were increased to 1000 per country to accommodate non-participation, technical and screening failures, and stratified by sex and eGFR stage to overrepresent lower levels of kidney function (**Section S1**).

Study procedures

All study protocols were harmonised prior to starting the study. We administered 5milliliters of Omnipaque™ (350mg iodine/ml, GE Healthcare, USA) as an intravenous bolus and calculated the dose of iohexol from pre- and post-administration syringe weights, measured in milligrams to two decimal points. We drew venous samples from the contralateral arm at 5,120, 180, and 240 minutes after iohexol administration, recording exact times for iohexol administration and sampling (**Section S2**).[18] We used the slope-intercept method to calculate mGFR for three time points in the slow exponential phase of iohexol elimination.

Laboratory methods and testing

Iohexol plasma samples were processed at each partner laboratory, stored at -112F, and shipped to a national reference laboratory in Johannesburg. Iohexol plasma concentrations were assayed using high-performance liquid chromatography-tandem mass spectrometry.[19] Coefficients of variation for internal quality control with the certified reference material for iohexol at 100 and 1000mg/L were 4.1% and 4.2% respectively. The laboratory complied with Equalis external quality assurance requirements for iohexol (Uppsala, Sweden) (**Section S3.1; Table S3.1**).[20]

In each country, laboratories performed serum creatinine measurements using an IDMS-traceable assay calibrated to a standard reference material for creatinine. The modified Jaffe method was used in Malawi and South Africa, and the enzymatic method in Uganda. For cystatin C, Malawian and Ugandan samples were analyzed in Uganda, while South African samples were measured in South Africa (**Section S3.2; Table S3.2**). Intersite analytic bias was assessed with a split sample recalibration study for creatinine and cystatin C (**Section S3.3; Table S3.3**).[21,22] Creatinine values for Malawi and South Africa

were recalibrated by a constant factor of +0.11mg/dL to adjust for method-related bias and align with the Uganda enzymatic method (**Figures S3.1-2**). For cystatin C we used Passing-Bablok regression analysis to determine constants for recalibration and Uganda and Malawi cystatin C measurements were adjusted with South Africa as the reference (**Figures S3.3-4**). Recalibrated values for both analytes were used for all analyses.

GFR estimating equations

We evaluated the following eGFR equations: Cockroft Gault (adjusted for BSA);[23] Four variable MDRD re-expressed for IDMS-traceable assays;[24] CKD-EPI creatinine,[25] cystatin C and creatinine-cystatin C;[26] Revised Lund-Malmö Study;[27] and Full Age Spectrum (FAS) creatinine (**Section S4**).[28] For the FAS equation we derived country-specific healthy population creatinine values (Q) (**Table S4.1**).

Data management and statistical analysis

We included participants who met all the following criteria: complete recordings of age, sex, height, weight; exact times for administering intravenous iohexol (T0) and subsequent sampling; iohexol plasma concentrations at each timepoint and demonstrating a monotonic decline; pre- and post-administration syringe weights; serum creatinine concentration ≥0.34mg/dL. We examined distribution of data overall and by country, and evaluated agreement between mGFR and creatinine and cystatin C-based eGFR equations using Bland Altman plots. For MDRD and CKD-EPI equations, we evaluated performance with and without ethnicity coefficients. Using mGFR as the reference, for each eGFR equation, we compared performance using bias, precision, and accuracy and compared the proportion of participants correctly classified by mGFR stage. The parameters for performance expressed as GFR (mL/min/1.73m^2) were absolute bias: median difference of each individual's eGFR and their respective mGFR (eGFR-mGFR); relative bias: median percentage difference of (eGFR-mGFR), expressed as a percent; precision: interquartile range (IQR) of (eGFR-mGFR); precision: root mean square error (RMSE), representing standard deviation of (eGFR-mGFR); accuracy: median of the absolute difference of (eGFR-mGFR) expressed as a percent of mGFR for estimates within 10% (P_{10}) and 30% (P_{30}) of mGFR.Based on the British Nuclear Medicine Society Guidelines,[18] to evaluate whether our findings were influenced by systematic error we performed sensitivity analyses with three restricted datasets (i) r>0.985 for the slope-intercept measured GFR derivation; (ii) normal sex-specific calculated volumes of distribution (11-17 liters for women; 13-20 liters for men) (iii) a double-restricted dataset combining (i) and (ii).

Development of a new creatinine-based GFR estimating equation

We chose creatinine over cystatin C to ensure clinical utility in SSA. Using the whole mGFR sample we sought to develop a new eGFR model by regressing log mGFR on log creatinine as well as age and sex, with and without body mass index (BMI),

mirroring the functional form of the CKD-EPI and Lund-Malmö eGFR equations. Regression coefficients were allowed to vary by sex, and the model included a spline knot with the position chosen using lowess plots. We used likelihood ratio tests and adjusted R^2 to decide between candidate models and assessed model performance as previously defined, compared to existing creatinine-based eGFR equations. For external validation, we compared the performance of eGFR from our model and existing equations to mGFR from radionuclide studies of 651 individuals (along with their creatinine measurements, age, and sex) obtained during cohort and clinical studies in South Africa.

Estimating true population prevalence of impaired kidney function in SSA

Based on population-based studies in SSA we sought to estimate the prevalence of impaired kidney function defined as eGFR<60mls/min/1.73m^2 (analogous to CKD stages G3a-5), using the most accurate eGFR equation. However, since the predictive power of all estimating equations was limited, we used a multiple imputation model trained on the measured mGFR sample, to predict individual GFR based on creatinine, age, and sex with predictions were based on 100 imputed datasets. The imputation model builds a predictive distribution around each unobserved GFR with variance and mean determined through a linear regression on the chosen predictor variables. We tested the model on the datasets from which it was trained and then used this approach within six distinct, large population-representative datasets: (i) the baseline prevalence studies in the ARK Consortium of adults 20-80 years;[16,17] (ii) the Africa Wits-International Network for the Demographic Evaluation of Populations and their Health Partnership for Genomic Studies (AWI-Gen) in Ghana, Kenya, and Burkina Faso and South Africa, which recruited adults 40-60 years.[29,30]

Results

Characteristics of ARK participants in Malawi, South Africa, and Uganda

Across four study sites in three countries, we measured GFR using iohexol plasma excretion in 3,025 adults and included 2,578 participants in the final analysis (**Table 1**). The predominant reason for exclusion (n=266) was missing syringe weight during the pilot phase: a flowchart of derivation of the final study sample is shown in **Figure S4** and comparison of characteristics with the whole sample in **Table S5**. Mean age was 50±15 years, mean height and weight were 162±9 cm and 67±17 kg, respectively, and mean mGFR was 82±26 ml/min/1.73m^2. South African participants were taller and heavier with disproportionate obesity in women, contrasting with Ugandan participants who had the smallest body mass indices in the

study. The distribution of mGFR, volume of distribution (Vd) and correlation coefficient (r) was similar across countries (**Figures S5-7**).

Performance of eGFR equations against mGFR

Across all three countries, creatinine-based equations overestimated GFR compared to mGFR (**Figure 1; Figures S8-9**). The greatest bias, least precision, and least accuracy were observed with mGFR <45ml/min/1.73m^2 (**Figure 2; Table 2; Figures S10-11; Table S6**). When classifying by mGFR within categories analogous to CKD stages, eGFR equations overestimated G1 and underestimated lower stages of kidney function. All creatinine-based equations underestimated GFR stages 3a-5 by at least 50% (**Table S7**). Overestimation of individual GFR and misclassification by GFR stage was exacerbated when using ethnicity coefficients for the MDRD and CKD-EPI equations (**Figure 1; Table 2; Figure S12; Table S7**). The Revised Lund-Malmö Study equation correctly classified the highest proportion (54%) of participants by mGFR stage (**Table S7**).

Performance of estimating equations using cystatin C alone, or in combination with creatinine, was higher than with creatinine alone (**Figure 1; Table 2; Tables S6-7**). Although cystatin C underestimated stage G1, the estimates of prevalence of impaired kidney function were almost two-fold higher than creatinine-based estimates, mostly in stage G3a (**Table S7**). This translated to 76% of individuals being correctly classified as having a kidney function stage as or more severe than their corresponding mGFR stage, compared to 72% with the Lund-Malmö (revised) equation and 59% with the CKD-EPI equation (**Table S7**). The performance of equations at estimating GFR compared to mGFR was similar between countries (**Tables S8-10**).

Sensitivity analyses

We compared mGFR distribution (**Figure S13a-c**), agreement between eGFR and mGFR for each equation both overall (**Figure S14a-c**) and by GFR stage (**Table S11a-c**) for single and double-restricted datasets. Accuracy, as $P_{30,}$ for CKD-EPI compared to mGFR was 65% in the unrestricted dataset, 69% and 67% in the single restricted datasets, and 71% in the dataset restricted for both Vd and r (**Table 2; Table S12a-c**). The best performing equation, Lund-Malmö, had a P_{30} of 87% in the double-restricted dataset. The P_{30} for CKD-EPI cystatin increased from 70% in the main dataset to 78% in the double-restricted dataset. However, there was also a disproportionate loss of people with low mGFR, from 6% to 2% in the double-restricted dataset (**Figure S15, Table S13a-c**). Measurements of low GFR have lower r without necessarily indicating substantial error in GFR measurement.[31] In addition, people with impaired kidney function may have abnormal fluid balance, while normal volumes of distribution have not been validated in SSA. Given the risk of creating systematic bias through data restriction and

the only modest increase in performance for the equations in sensitivity analyses we used the complete dataset for further analyses.

Development of a new creatinine-based GFR estimating equation

We plotted the relationship between creatinine and mGFR and inspection of lowess curves (**Figures S16**) supported a piecewise linear model with one knot at 0.82mg/dL for men with no evidence to support fitting separate coefficients for creatinine among men and women before the knot, nor to support separate age coefficients for men and women. Including BMI in the model increased the adjusted R^2 but only from 0.225 to 0.234. We examined agreement between mGFR and the ARK models (**Figure S17**) and all equations had limited predictive power with chosen predictors accounting for 20-25% of the total variation in the data irrespective of the model form (**Table S14**). The final model was:

If male: $eGFR = 124 \times \min(1, SCr/0.82)^{-0.339} \times \max(1, SCr/0.82)^{-0.574} \times 0.993^{age}$

If female: $eGFR = 103 \times (SCr/0.82)^{-0.339} \times 0.993^{age}$

We used this equation to estimate GFR measurements among the 2,578 people in our primary dataset. Although model predictions were less biased than those of the CKD-EPI equation, our model showed limited discrimination and was not able to better predict kidney function than existing equations, categorizing 54% of people correctly for all GFR stages (**Tables S15-17**). In the external validation dataset (**Table S18**) our model did not have better performance than the CKD-EPI or Revised Lund-Malmö equations (**Tables S19-20**).

Estimating true population prevalence of impaired kidney function in SSA

The multiple imputation model predicted the proportion of people with impaired kidney function in the primary ARK datasets in which it was developed to be between 3.0% higher and 4.2% lower than mGFR (Uganda: observed 13.8% versus predicted 10.8%; Malawi: observed 20.4% versus predicted 24.6%; South Africa: observed 22.9% versus predicted 21.0%). In the external validation dataset from South Africa the same model predicted the proportion of people with impaired kidney function to be 20.9% versus observed 17.2%. With this method, we estimated the prevalence of impaired kidney function to be between 4.6% and 15.1% higher than when estimated using CKD-EPI creatinine in population-based countries in West, East, and southern Africa (**Figure 3**, **Tables S21-22**).

Discussion

Our results show that across three Sub-Saharan African countries, in diverse urban and rural populations, creatinine-based eGFR equations substantially overestimate kidney function compared to GFR measured with iohexol (mGFR). The

overestimation worsened at lower levels of mGFR and was further exacerbated by inclusion of ethnicity coefficients. Cystatin C-based equations performed better than creatinine-based equations but none of the eGFR equations achieved an accuracy (P_{30}) considered appropriate for individual clinical decision-making.[32] At population level, using creatinine-based eGFR equations substantially underestimated the prevalence of impaired kidney function. A new creatinine-based equation to better estimate GFR in SSA was not possible due to wide age- and sex-independent variability in the relationship between creatinine and mGFR, even with weight or BMI adjustment. Therefore, we used a multiple imputation method to estimate population prevalence of impaired kidney function in countries across SSA and found the prevalence was substantially higher than that measured using creatinine-based estimates of GFR.

Our large, prospective study used population-representative samples with oversampling of those with poor kidney function, and used both creatinine and cystatin to estimate GFR. Aside from US-derived eGFR equations, we included the FAS equation[28] and Revised Lund-Malmö Study equation[27] not previously evaluated in SSA. We used a gold-standard multi-sample method for iohexol plasma clearance with centralized measurements in a reference laboratory with extensive experience using mass-spectrometry assays, compliant with the Equalis quality assurance program for iohexol. We addressed intersite analytic bias for creatinine and cystatin C and recalibrated measurements accordingly. However, there are also limitations to this work. Because we measured creatinine and mGFR at a single time point, we did not fulfil the temporal requirements for the definition of CKD[32] although our approach is in keeping with other large cohort studies of GFR.[34] In addition, participants were well, and therefore (aside from seasonal changes in eGFR),[33] we would not anticipate substantial variation in kidney function. Our sampling frame was population-based with high mean kidney function compared to CKD cohorts used to develop GFR equations.[2,26,35] Measurement of low creatinine values is more prone to biological and analytic variation than high level.s[36] It is possible that variability in direct GFR measurement inherent to a study largely conducted in rural SSA, despite rigorous quality control procedures, may have contributed to the reduced performance of GFR estimating equations. We have quantified this with transparently reported sensitivity analyses and in these estimating equations (compared to mGFR) did not attain the level of performance reported in studies in different settings.[6,26,37] These restricted analyses resulted in disproportionate exclusion of people with low GFR, where methods of determining quality of GFR measurement may not be appropriate,[31] and so we undertook further analysis in the whole dataset. In addition, prevalence of impaired kidney function measured with iohexol GFR was very similar to that estimated from cystatin, reinforcing that it is variability in the association between creatinine and GFR that leads to poor performance of eGFR in this population rather than measurement error of iohexol GFR. Finally, while our study is the largest of its kind from SSA and includes measured

GFR from three countries, the diversity of African populations means that there is likely to be greater variability in the relationship between creatinine and GFR than we have measured. This could affect the accuracy of the imputed GFR estimates in the AWI-GEN datasets but in internal validation the model estimates prevalence of impaired kidney function close to that measured in the ARK datasets, and with differences substantially less than the variation we see between creatinine-based and imputed estimates.

Our finding that creatinine-based eGFR equations overestimate well-preserved kidney function (stages G1-2) but underestimate declining kidney function (stages G3-5) has been replicated in studies from Scandinavia and West Africa, [37,38] but differs from those in the Japan and the USA.[24,26,39] Greater inaccuracy in GFR estimation from creatinine-based equations has been described in Asia,[40] and there are reasons why the relationship between creatinine and GFR may be more variable in SSA compared to US or European populations. Our study and others report lower BMI, with likely lower muscle mass and therefore lower baseline creatinine, in continental Africans when compared to the African diaspora.[4,5,16,17] Perinatal and early childhood factors resulting in undernutrition and growth stunting, which predisposes to low lean muscle mass and short stature in adulthood even in the presence of adult obesity, are common in SSA.[13,41-43] Wasting from chronic infection or inflammation, such as tuberculosis and HIV, low dietary protein ingestion, and undiagnosed liver disease also impact muscle mass and creatinine generation.[44-47] Finally, variants in genes affecting pathways of creatinine production and tubular secretion may impact the association between serum creatinine and true GFR in SSA.[48,49]

The first implication of our results is to confirm conclusively that ethnicity coefficients exaggerate the overestimation of GFR by creatinine and should not be used in SSA. Secondly, overestimation of kidney function at lower levels of GFR impacts clinical care and public health. For an individual, inaccurate estimation of GFR risks a missed diagnosis of CKD and the opportunity to address modifiable risk factors to slow progression of kidney disease, and possible harm from unadjusted doses of renally cleared drugs. At population level underestimation of the burden of CKD in SSA limits prioritization of country-specific and regional strategies to minimize CKD progression and manage end-stage kidney disease. Limited access to specialist renal care in SSA compounds the individual risk for progression, premature disability, and death. While cystatin C identified a greater proportion of people with CKD and might be a valuable confirmatory test, assays are more expensive than creatinine and not widely available in SSA. There is therefore urgent need for further research to develop accurate and low-cost ways to measure kidney function in SSA.

Conclusion

In this large collaborative study from three Sub-Saharan African countries, we prospectively measured kidney function using consistent and robust techniques. We showed that creatinine-based GFR estimating equations overestimate kidney function compared to iohexol and cystatin C measures. Our results suggest the burden of kidney disease is markedly underestimated in SSA, with substantial implications for individual health care and public health interventions to address the challenge of kidney disease in resource-limited settings.

Table 1: Characteristics of ARK participants by sex, for each country, and overall (N=2,578)

	MALAWI		SOUTH AFRICA		UGANDA		OVERALL
Characteristic[1]	**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	
	N=474	**N=424**	**N=636**	**N=311**	**N=413**	**N=320**	**N=2,578[2]**
Age - year	53.1 (13.5)	52.8 (16.3)	45.6 (14.4)	44.5 (16.3)	50.6(14.2)	51.5 (15.1)	49.5 (15.2)
Age category							
<40 year	70 (14.8%)	87 (20.5%)	244 (38.4%)	138 (44.4%)	92 (22.3%)	74 (23.1%)	705 (27.4%)
40-60 year	269 (56.8%)	209 (49.3%)	280 (44.0%)	109 (35.1%)	232 (56.2%)	158 (49.4%)	1257 (48.8%)
>60 year	135 (28.5%)	128 (30.2%)	112 (17.6%)	64 (20.6%)	89 (21.6%)	88 (27.5%)	616 (23.9%)
BMI[3]	26.76 (5.86)	23.11 (3.86)	30.14 (6.41)	25.02 (5.08)	23.92 (4.47)	21.03 (3.12)	25.61 (5.98)
BMI category[4]							
<18.5 (underweight)	14 (3.0%)	25 (5.9%)	9 (1.4%)	14 (4.5%)	21 (5.1%)	63 (19.7%)	146 (5.7%)
18.5-24.9 (normal)	194 (40.9%)	293 (69.1%)	140 (22.0%)	162 (52.1%)	259 (62.7%)	222 (69.4%)	1270 (49.3%)
25.0-29.9 (overweight)	148 (31.2%)	86 (20.3%)	182 (28.6%)	84 (27.0%)	91 (22.0%)	32 (10.0%)	623 (24.2%)
≥30.0 (obese)	118 (24.9%)	20 (4.7%)	305 (48.0%)	51 (16.4%)	42 (10.2%)	3 (0.9%)	539 (20.9%)
Weight (kg)	65.00 (15.48)	63.24 (11.82)	78.70 (17.37)	74.64 (16.71)	57.56 (11.89)	56.95 (9.64)	67.06 (16.72)
Height (cm)	155.67 (5.99)	165.28 (6.83)	161.56 (5.91)	172.60 (7.14)	154.98 (6.83)	164.42 (6.36)	161.72 (8.53)
BSA(m^2)[5]	1.68 (0.22)	1.70 (0.18)	1.90 (0.23)	1.89 (0.24)	1.58 (0.19)	1.61 (0.16)	1.74 (0.24)
Serum creatinine[6]	0.73 (0.20)	0.92 (0.30)	0.61 (0.16)	0.83 (0.23)	0.74 (0.20)	0.89 (0.31)	0.77 (0.25)
Serum cystatin C[7]	0.97 (0.34)	1.01 (0.30)	0.99 (0.27)	1.01 (0.28)	0.88 (0.23)	0.91 (0.31)	0.97 (0.29)
Iohexol GFR[8]	74.8 (20.1)	79.1 (20.4)	78.6 (25.0)	82.0 (26.4)	86.1 (26.3)	97.1 (31.2)	81.9 (25.6)
Iohexol GFR category[9]							
G1: ≥90	100 (21.1%)	124 (29.3%)	205 (32.2%)	130 (41.8%)	160 (38.7%)	190 (59.4%)	909 (35.3%)

G2: 60-89	271 (57.2%)	220 (51.9%)	282 (44.3%)	113 (36.3%)	195 (47.2%)	87 (27.2%)	1168 (45.3%)
G3a: 45-59	76 (16.0%)	61 (14.4%)	92 (14.5%)	40 (12.9%)	42 (10.2%)	29 (9.1%)	340 (13.2%)
G3b: 30-44	21 (4.4%)	18 (4.3%)	49 (7.7%)	21 (6.8%)	10 (2.4%)	10 (3.1%)	129 (5.0%)
G4-5: <30	6 (1.3%)	1 (0.2%)	8 (1.3%)	7 (2.3%)	6 (1.45%)	4 (1.3%)	32 (1.2%)
Estimated GFR[9]	82.6 (19.4)	86.2 (20.1)	99.7 (19.1)	98.5 (21.3)	93.7 (19.9)	97.8 (20.5)	93.2 (21.0)
Estimated GFR category[10]							
G1: ≥90	174 (36.7%)	194 (45.8%)	468 (73.6%)	220 (70.7%)	241 (58.4%)	237 (74.1%)	1534 (59.5%)
G2: 60-89	245 (51.7%)	189 (44.6%)	146 (23.0%)	71 (22.8%)	154 (37.3%)	69 (21.6%)	874 (33.9%)
G3a: 45-59	41 (8.7%)	26 (6.1%)	17 (2.7%)	14 (4.5%)	13 (3.2%)	5 (1.6%)	116 (4.5%)
G3b: 30-44	12 (2.5%)	10 (2.4%)	4 (0.6%)	5 (1.6%)	3 (0.7%)	4 (1.3%)	38 (1.5%)
G4-5: <30	2 (0.4%)	5 (1.2%)	1 (0.2%)	1 (0.3%)	2 (0.5%)	5 (1.6%)	16 (0.6%)

[1]All data reported as mean (standard deviation); categories reported as number (%); percentages may sum to +/-100 due to rounding; [2]N=2,578 for creatinine samples; N=2433 for cystatin-C samples; [3]Body mass index (BMI) calculated by dividing weight (kilograms) by height squared (meters); [4]BMI category: WHO classification for obesity[46]; [5]Body surface area calculated using the Haycock formula[20]; [6]Serum creatinine in mg/dL (conventional units): to convert to µmol/L, multiply by 88.4; creatinine data unadjusted for intersite calibration study; creatinine data unadjusted for intersite calibration study; [7]Serum cystatin C in mg/dL: cystatin C data unadjusted for intersite calibration study; [8]Iohexol GFR: ml/min/1.73m^2; [9]Iohexol GFR category: GFR stage defined by KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease[32]; [10]Estimated GFR: ml/min/1.73m^2 using CKD-EPI creatinine equation (no ethnicity co-efficient)

Table 2: Overall performance of GFR estimating equations compared to iohexol GFR (N=2,578)

Variable	Estimated GFR (ml/min/1.73m² of body surface area)				
Iohexol GFR (ml/min/1.73m² of BSA)	>=90	60 – 89	45 – 59	<45	overall
GFR Stage	G1	G2	G3a	G3b-5	
[1]Absolute bias (95% CI)					
Cockroft Gault	-1.6 (-3.9 - 0.3)	13.5 (11.8 - 15.1)	21.9 (18.0 - 25.7)	36.5 (31.3 - 44.5)	11.3 (10.1 - 12.7)
MDRD	-8.1 (-9.8 - -6.5)	8.6 (7.6 - 9.8)	20.3 (17.8 - 23.6)	34.5 (28.2 - 40.9)	6.6 (5.6 - 7.7)
MDRD + ethnicity coefficient	12.2 (10.9 - 14.5)	26.2 (25.0 - 27.7)	35.8 (33.2 - 40.2)	48.8 (42.0 - 57.0)	24.7 (23.4 - 25.7)
Full Age Spectrum$_{(creatinine)}$	-0.9 (-3.6 - 0.7)	14.1 (12.4 - 15.9)	23.2 (19.1 - 26.7)	36.0 (30.0 - 41.1)	11.2 (10.1 - 12.3)
Revised Lund-Malmö Study	-13.0 (-14.5- -11.6)	6.5 (5.7 - 7.4)	19.3 (16.7 - 22.3)	33.8 (27.5 - 38.8)	2.3 (1.3- 3.3)
CKD-EPI$_{(creatinine)}$	-1.1 (-3.4 - 0.35)	16.1 (14.9 - 17.2)	26.1 (23.5 - 29.9)	40.7 (32.7- 47.2)	11.5 (10.6 - 12.7)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	15.2 (13.0 - 17.4)	30.9 (29.3 - 32.0)	39.0 (35.6 - 43.4)	53.9 (44.2 - 59.7)	26.7 (25.4 - 27.6)
CKD-EPI$_{(creatinine + cystatin C)}$	-7.6 (-9.5 - -6.2)	6.9 (5.7 - 8.5)	16.5 (13.4 - 21.0)	27.2 (22.9 - 36.4)	4.1 (3.1 - 5.1)
CKD-EPI$_{(cystatin C)}$	-15.0 (-17.2- -13.0)	0.5 (-1.4 - 2.1)	9.6 (6.8 - 12.4)	20.6 (13.9 - 27.7)	-1.7 (-2.9 - -0.8)
[2]Relative bias (95% CI)					
Cockroft Gault	0.99 (0.96 - 1.00)	1.18 (1.16 - 1.20)	1.40 (1.35 - 1.47)	1.97 (1.83 - 2.27)	1.15 (1.13 - 1.17)
MDRD	0.92 (0.91 - 0.94)	1.12 (1.10 - 1.13)	1.39 (1.34 - 1.45)	1.99 (1.75 - 2.11)	1.08 (1.07 - 1.10)
MDRD + ethnicity coefficient	1.12 (1.10 - 1.14)	1.35 (1.34 - 1.37)	1.68 (1.62 - 1.76)	2.41 (2.12 - 2.56)	1.31 (1.30 - 1.33)
Full Age Spectrum$_{(creatinine)}$	0.99 (0.97 - 1.01)	1.19 (1.17 - 1.21)	1.43 (1.37 - 1.49)	2.00 (1.81 - 2.15)	1.14 (1.13 - 1.16)
Revised Lund-Malmö Study	0.88 (0.87 – 0.89)	1.08 (1.07 - 1.10)	1.36 (1.30 - 1.42)	1.94 (1.68 - 2.06)	1.03 (1.02 - 1.04)
CKD-EPI$_{(creatinine)}$	0.99 (0.97- 1.00)	1.21 (1.19 - 1.23)	1.50 (1.43-1.56) l	2.12 (1.93 - 2.26)	1.15 (1.13 - 1.16)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	1.15 (1.12 - 1.16)	1.41 (1.38 - 1.43)	1.74 (1.66 - 1.81)	2.46 (2.23 - 2.62)	1.33 (1.31 - 1.35)
CKD-EPI$_{(creatinine + cystatin C)}$	0.93 (0.91 - 0.94)	1.09 (1.08 - 1.11)	1.31 (1.24 - 1.39)	1.72 (1.58 - 1.96)	1.05 (1.04 - 1.07)
CKD-EPI$_{(cystatin C)}$	0.86 (0.84 - 0.88)	1.01 (0.98 - 1.03)	1.17 (1.13 - 1.24)	1.55 (1.37 - 1.73)	0.98 (0.96 - 0.99)

[3]Precision (RMSE) (95% CI)					
Cockroft Gault	31.5 (30.1 - 33.0)	26.7 (25.7 - 27.9)	29.0 (27.0 - 31.4)	34.9 (31.4 - 39.2)	31.7 (30.8 - 32.6)
MDRD	26.5 (25.4 - 27.8)	20.9 (20.1 - 21.8)	21.4 (19.9 - 23.2)	28.3 (25.5 - 31.8)	27.1 (26.4 - 27.9)
MDRD + ethnicity coefficient	30.5 (29.2 - 32.0)	25.0 (24.0 - 26.1)	25.9 (24.1 - 28.0)	33.8 (30.5 - 37.9)	30.1 (29.3 - 30.9)
Full Age Spectrum$_{(creatinine)}$	28.0 (26.8 - 29.3)	21.4 (20.6 - 22.3)	23.5 (21.9 - 25.5)	29.9 (27.0 - 33.6)	27.8 (27.0 - 28.6)
Revised Lund-Malmö Study	21.2 (20.3 - 22.3)	15.3 (14.7 - 16.0)	18.1 (16.9 - 19.6)	25.0 (22.6 - 28.1)	24.0 (23.3 - 24.6)
CKD-EPI$_{(creatinine)}$	22.9 (21.9 - 24.0)	17.6 (16.9 - 18.4)	21.2 (19.7 - 22.9)	28.5 (25.7 - 32.0)	24.9 (24.3 - 25.6)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	24.9 (23.8 - 26.1)	20.2 (19.4 - 21.0)	24.4 (22.7 - 26.4)	32.6 (29.4 - 36.7)	26.4 (25.7 - 27.2)
CKD-EPI$_{(creatinine + cystatin C)}$	21.9 (20.9 - 23.0)	17.5 (16.8 - 18.2)	20.3 (18.9 - 22.0)	27.4 (24.6 - 30.8)	23.8 (23.1 - 24.5)
CKD-EPI$_{(cystatin C)}$	24.1 (23.0 - 25.3)	20.6 (19.8 - 21.5)	21.9 (20.3 - 23.8)	27.3 (24.6 - 30.8)	25.9 (25.2 - 26.6)
[4]Precision (IQR) (95% CI)					
Cockroft Gault	40.4 (37.6 - 43.2)	34.2 (31.6 - 36.7)	42.0 (35.6 - 48.5)	51.0 (42.6 - 59.5)	37.6 (35.7 - 39.5)
MDRD	32.2 (29.3 - 35.1)	26.0 (24.1 - 27.8)	28.2 (23.7 - 32.7)	35.9 (30.5 - 41.4)	31.1 (29.6 - 32.6)
MDRD + ethnicity coefficient	37.8 (34.9 - 40.7)	30.5 (28.5 - 32.6)	33.7 (28.0 - 39.3)	44.2 (37.3 - 51.2)	35.7 (34.0 - 37.3)
Full Age Spectrum$_{(creatinine)}$	35.6 (32.5 - 38.6)	26.1 (24.3 - 30.0)	34.2 (30.5 - 37.8)	44.7 (35.5 - 54.0)	31.4 (29.8 - 33.0)
Revised Lund-Malmö Study	24.4 (22.6 - 26.1)	19.9 (18.5 - 21.2)	25.6 (22.3 - 28.9)	35.7 (29.2 - 42.3)	26.9 (25.6 - 28.2)
CKD-EPI$_{(creatinine)}$	27.3 (25.2 - 29.4)	24.1 (22.3 - 25.9)	31.3 (27.8 - 34.8)	40.5 (33.8 - 47.2)	28.5 (27.1 - 30.0)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	30.5 (27.8 - 33.2)	27.2 (25.3 - 29.2)	36.5 (32.3 - 40.7)	46.9 (39.0 - 54.8)	31.3 (29.7 - 32.7)
CKD-EPI$_{(creatinine + cystatin C)}$	27.1 (24.8 - 29.4)	24.4 (22.9 - 25.9)	28.8 (26.0 - 31.7)	39.7 (32.6 - 46.8)	27.4 (26.0 - 28.8)
CKD-EPI$_{(cystatin C)}$	31.2 (28.7 - 33.6)	27.9 (26.0 - 29.9)	28.9 (23.1 - 34.7)	35.2 (27.0 - 43.3)	30.7 (28.9 - 32.5)
[5]Accuracy % P_{10} (95% CI)					
Cockroft Gault	0.27 (0.24 - 0.30)	0.24 (0.22 - 0.27)	0.17 (0.13 - 0.21)	0.08 (0.04 - 0.13)	0.23 (0.22 - 0.25)
MDRD	0.32 (0.29 - 0.35)	0.29 (0.27 - 0.32)	0.14 (0.10 - 0.18)	0.05 (0.02 - 0.10)	0.27 (0.25 - 0.28)
MDRD + ethnicity coefficient	0.27 (0.25 - 0.30)	0.15 (0.13 - 0.17)	0.04 (0.02 - 0.07)	0.04 (0.02 - 0.09)	0.17 (0.16 - 0.19)

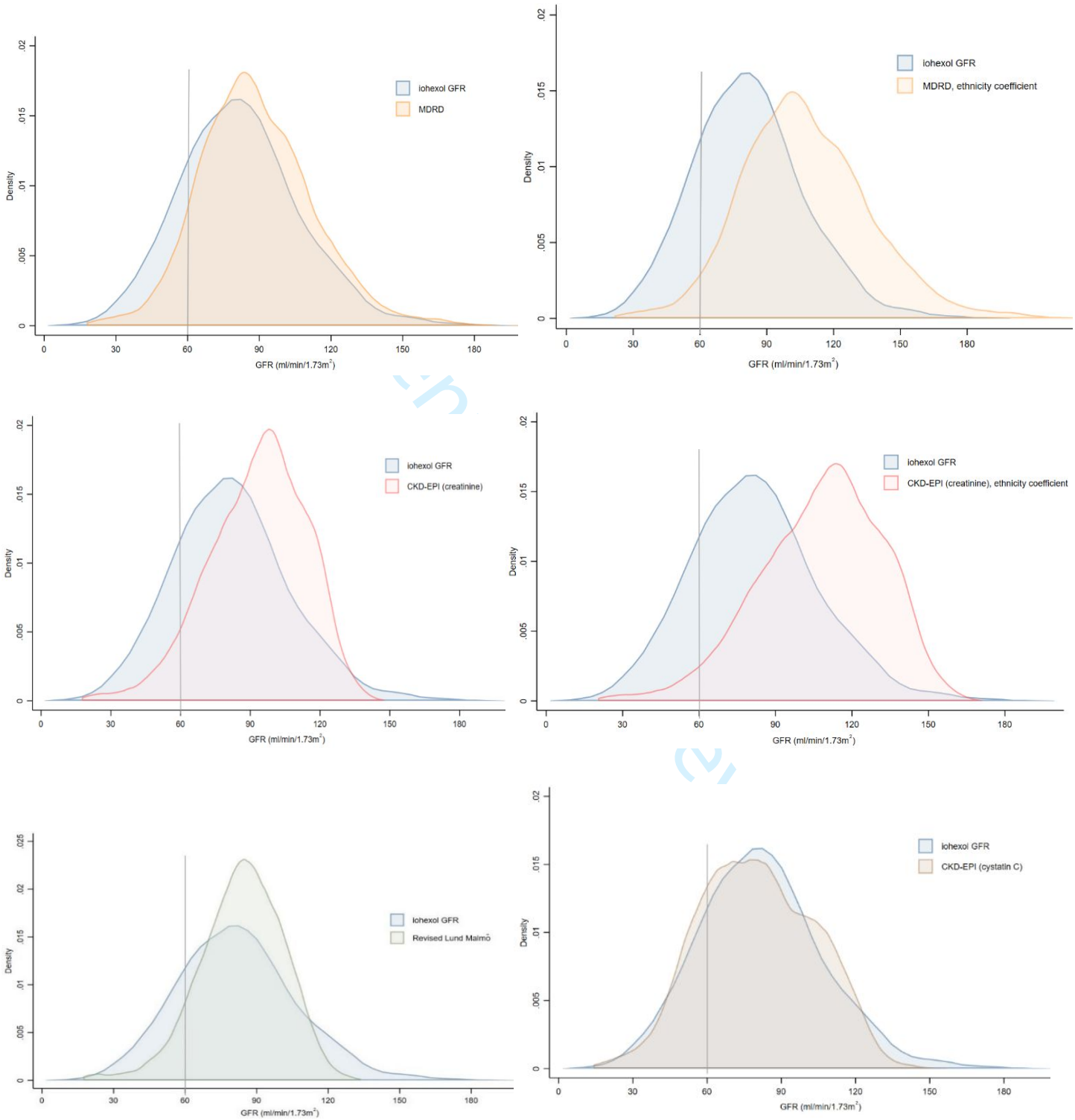
Full Age Spectrum$_{(creatinine)}$	0.32 (0.29 - 0.35)	0.25 (0.22 - 0.28)	0.14 (0.10 - 0.18)	0.06 (0.03 - 0.11)	0.25 (0.23 - 0.26)
Revised Lund-Malmö Study	0.36 (0.33 - 0.40)	0.36 (0.34 - 0.39)	0.16 (0.12 - 0.20)	0.08 (0.04 - 0.13)	0.32 (0.30 - 0.34)
CKD-EPI$_{(creatinine)}$	0.42 (0.39 - 0.45)	0.23 (0.21 - 0.26)	0.12 (0.09 - 0.16)	0.06 (0.03 - 0.11)	0.27 (0.26 - 0.29)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	0.26 (0.23 - 0.29)	0.11 (0.09 - 0.13)	0.04 (0.02 - 0.07)	0.04 (0.01 - 0.08)	0.15 (0.14 - 0.16)
CKD-EPI$_{(creatinine + cystatin C)}$	0.36 (0.33 - 0.40)	0.32 (0.29 - 0.34)	0.18 (0.14 - 0.22)	0.07 (0.04 - 0.12)	0.30 (0.28 - 0.32)
CKD-EPI$_{(cystatin C)}$	0.29 (0.26 - 0.32)	0.27 (0.25 - 0.30)	0.25 (0.20 - 0.30)	0.08 (0.04 - 0.13)	0.26 (0.25 - 0.28)
[6]Accuracy % (1 - P_{10}) (95% CI)					
Cockroft Gault	0.73 (0.70 - 0.76)	0.76 (0.73 - 0.78)	0.83 (0.79 - 0.87)	0.92 (0.87 - 0.96)	0.77 (0.75 - 0.78)
MDRD	0.68 (0.65 - 0.71)	0.71 (0.68 - 0.74)	0.86 (0.82 - 0.90)	0.95 (0.90 - 0.98)	0.73 (0.72 - 0.75)
MDRD + ethnicity coefficient	0.73 (0.70 - 0.76)	0.85 (0.83 - 0.88)	0.96 (0.94 - 0.98)	0.96 (0.91 - 0.98)	0.83 (0.82 - 0.84)
Full Age Spectrum$_{(creatinine)}$	0.68 (0.65 - 0.71)	0.75 (0.73 - 0.78)	0.86 (0.82 - 0.90)	0.94 (0.89 - 0.97)	0.75 (0.74 - 0.77)
Revised Lund-Malmö Study	0.64 (0.60 - 0.67)	0.64 (0.61 - 0.67)	0.84 (0.80 - 0.88)	0.92 (0.87 - 0.96)	0.68 (0.66 - 0.70)
CKD-EPI$_{(creatinine)}$	0.58 (0.55 - 0.61)	0.77 (0.74 - 0.79)	0.88 (0.84 - 0.92)	0.94 (0.89 - 0.97)	0.73 (0.71 - 0.75)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	0.74 (0.71 - 0.77)	0.89 (0.87 - 0.91)	0.96 (0.93 - 0.98)	0.96 (0.92 - 0.99)	0.85 (0.84 - 0.86)
CKD-EPI$_{(creatinine + cystatin C)}$	0.64 (0.60 - 0.67)	0.68 (0.66 - 0.71)	0.82 (0.78 - 0.86)	0.93 (0.88 - 0.97)	0.70 (0.68 - 0.72)
CKD-EPI$_{(cystatin C)}$	0.71 (0.68 - 0.74)	0.73 (0.70 - 0.75)	0.75 (0.70 - 0.80)	0.92 (0.87 - 0.96)	0.74 (0.72 - 0.76)
[5]Accuracy % P_{30} (95% CI)					
Cockroft Gault	0.73 (0.70 - 0.76)	0.61 (0.58 - 0.64)	0.39 (0.34 - 0.44)	0.18 (0.12 - 0.25)	0.60 (0.58 - 0.62)
MDRD	0.78 (0.75 - 0.81)	0.72 (0.69 - 0.75)	0.40 (0.35 - 0.45)	0.16 (0.10 - 0.22)	0.66 (0.65 - 0.68)
MDRD + ethnicity coefficient	0.68 (0.65 - 0.71)	0.42 (0.39 - 0.45)	0.17 (0.13 - 0.22)	0.10 (0.06 - 0.16)	0.46 (0.44 - 0.48)
Full Age Spectrum$_{(creatinine)}$	0.78 (0.75 - 0.81)	0.64 (0.61 - 0.66)	0.38 (0.33 - 0.44)	0.19 (0.14 - 0.26)	0.63 (0.61 - 0.64)
Revised Lund-Malmö Study	0.83 (0.81 - 0.86)	0.83 (0.81 - 0.85)	0.43 (0.38 - 0.48)	0.20 (0.14 - 0.27)	0.74 (0.72 - 0.76)
CKD-EPI$_{(creatinine)}$	0.88 (0.85 - 0.90)	0.63 (0.61 - 0.66)	0.32 (0.27 - 0.37)	0.15 (0.10 - 0.21)	0.65 (0.63 - 0.67)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	0.75 (0.72 - 0.78)	0.35 (0.32 - 0.38)	0.16 (0.13 - 0.21)	0.11 (0.07 - 0.17)	0.45 (0.43 - 0.47)

CKD-EPI$_{(creatinine + cystatin C)}$	0.80 (0.77 - 0.82)	0.73 (0.71 - 0.76)	0.45 (0.39 - 0.50)	0.20 (0.14 - 0.27)	0.69 (0.67 - 0.70)
CKD-EPI$_{(cystatin C)}$	0.71 (0.68 - 0.74)	0.71 (0.69 - 0.74)	0.54 (0.49 - 0.60)	0.32 (0.25 - 0.40)	0.67 (0.65 - 0.68)
[6]Accuracy % (1 - P_{30}) (95% CI)					
Cockroft Gault	0.27 (0.24 - 0.30)	0.39 (0.36 - 0.42)	0.61 (0.55 - 0.66)	0.82 (0.75 - 0.88)	0.40 (0.38 - 0.42)
MDRD	0.22 (0.19 - 0.25)	0.28 (0.25 - 0.31)	0.60 (0.55 - 0.66)	0.84 (0.78 - 0.90)	0.34 (0.32 - 0.35)
MDRD + ethnicity coefficient	0.32 (0.29 - 0.35)	0.58 (0.55 - 0.61)	0.83 (0.78 - 0.87)	0.90 (0.84 - 0.94)	0.54 (0.52 - 0.56)
Full Age Spectrum$_{(creatinine)}$	0.22 (0.20 - 0.25)	0.36 (0.34 - 0.39)	0.62 (0.56 - 0.67)	0.81 (0.74 - 0.87)	0.37 (0.36 - 0.39)
Revised Lund-Malmö Study	0.17 (0.14 - 0.19)	0.17 (0.15 - 0.19)	0.57 (0.52 - 0.62)	0.80 (0.73 - 0.86)	0.26 (0.25 - 0.28)
CKD-EPI$_{(creatinine)}$	0.12 (0.10 - 0.15)	0.37 (0.34 - 0.39)	0.68 (0.63 - 0.73)	0.85 (0.79 - 0.90)	0.35 (0.33 - 0.37)
CKD-EPI$_{(creatinine)}$ + ethnicity coefficient	0.25 (0.22 - 0.28)	0.65 (0.62 - 0.68)	0.84 (0.79 - 0.87)	0.89 (0.83 - 0.93)	0.55 (0.53 - 0.57)
CKD-EPI$_{(creatinine + cystatin C)}$	0.20 (0.18 - 0.23)	0.27 (0.24 - 0.29)	0.55 (0.50 - 0.61)	0.80 (0.73 - 0.86)	0.31 (0.30 - 0.33)
CKD-EPI$_{(cystatin C)}$	0.29 (0.26 - 0.32)	0.29 (0.26 - 0.31)	0.46 (0.41 - 0.51)	0.68 (0.60 - 0.75)	0.33 (0.31 - 0.35)

N=2,578 for creatinine-based equations; N=2433 for cystatin C-based equations; all values reported with 95% confidence intervals in parentheses;

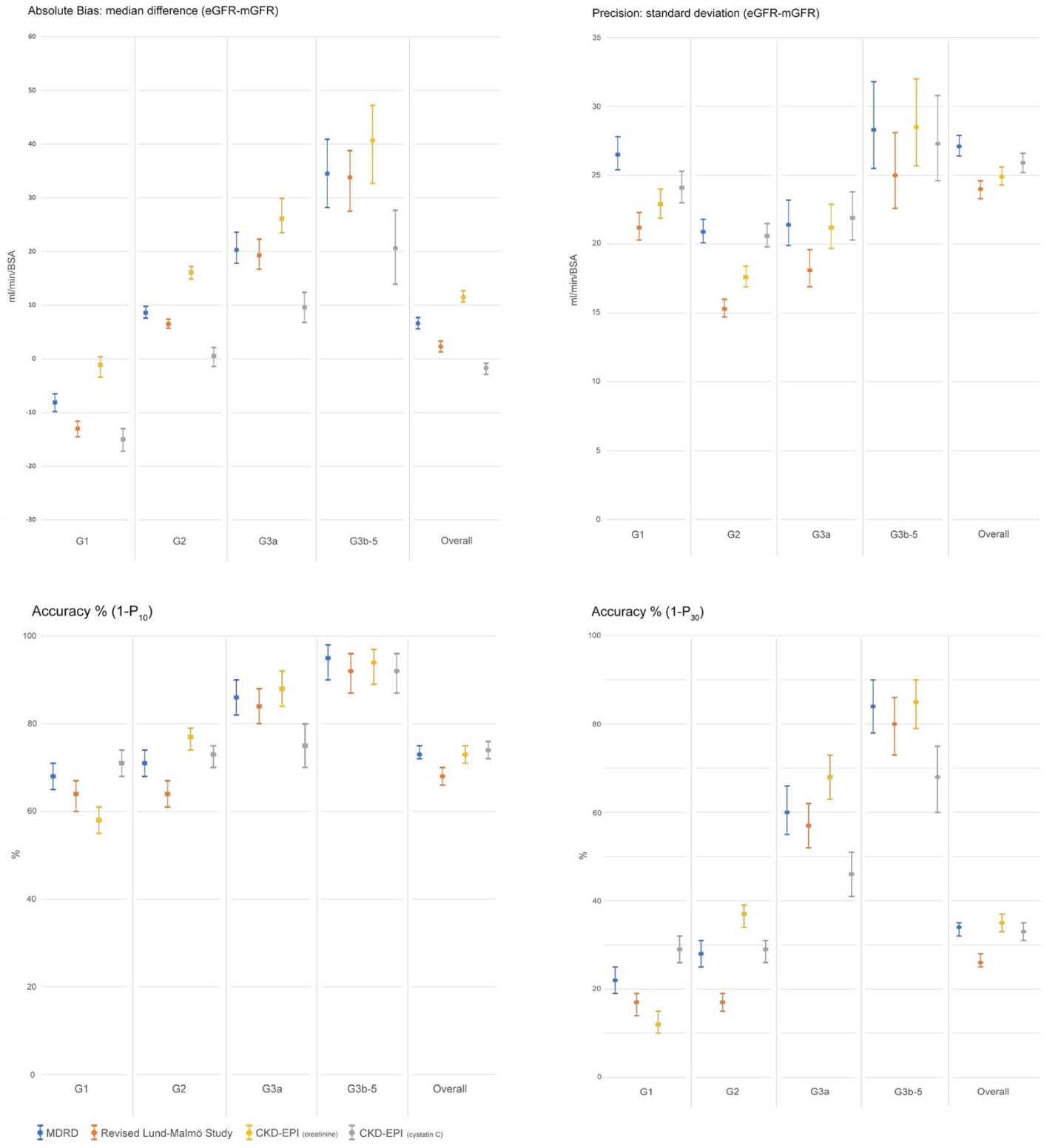
[1]Absolute bias: median (estimated GFR - iohexol GFR); [2]Relative bias: median ([estimated GFR - iohexol GFR]/iohexol GFR); [3] Precision RMSE (Root Mean Square Error): standard deviation (estimated GFR - iohexol GFR);[4]Precision IQR (Interquartile Range): IQR (estimated GFR - iohexol GFR); [5]Accuracy: median (estimated GFR - iohexol GFR) expressed as a percent of iohexol GFR for estimates within 10% (P10) and within 30% of iohexol GFR; [6] Accuracy: median (estimated GFR - iohexol GFR) expressed as a percent of iohexol GFR for estimates that differed by more than 10% (1-P10) and more than 30% (1-P30) of iohexol GFR.

Figure 1: Comparison of GFR distributions between GFR estimating equations and iohexol GFR for the ARK Study



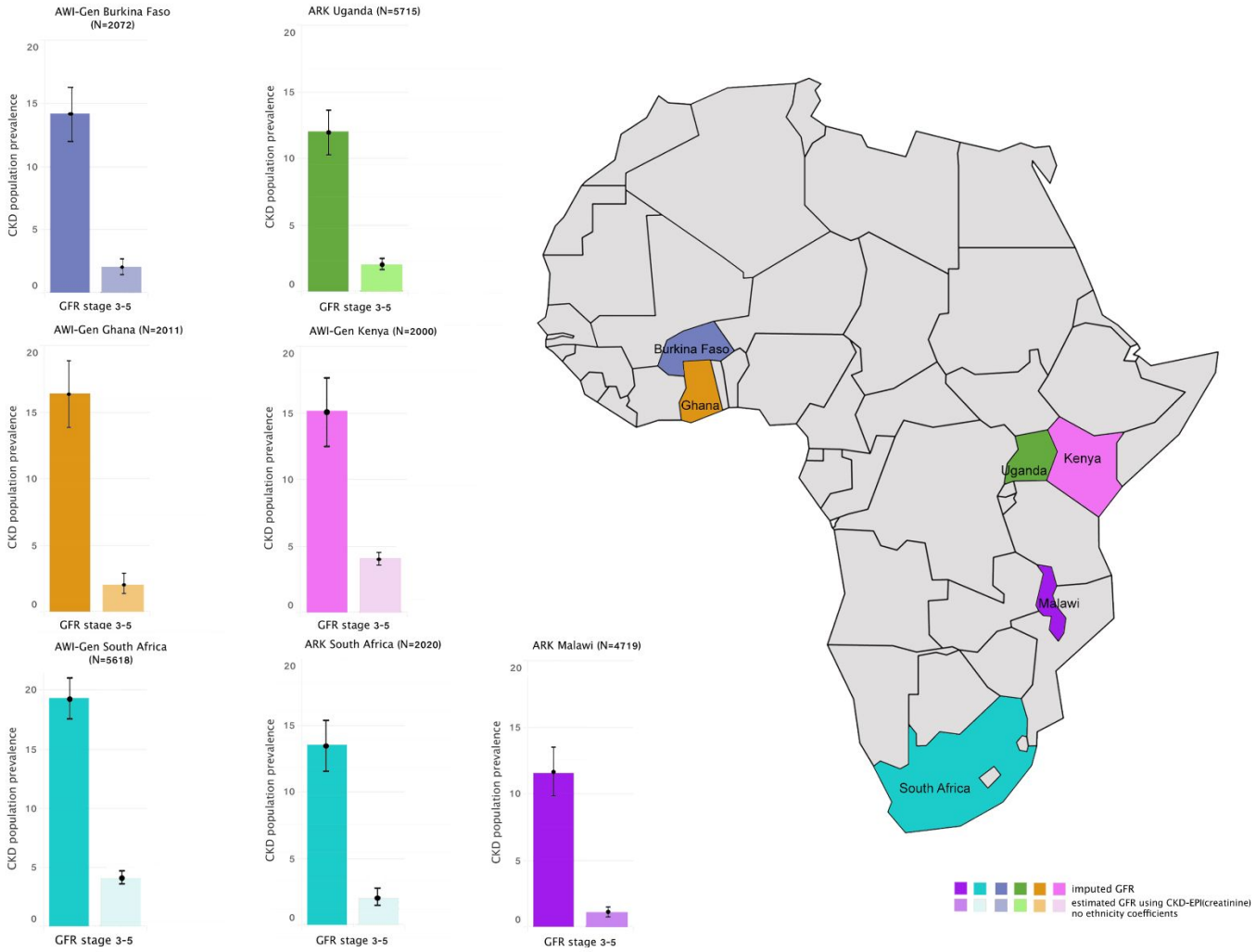
Data shown are from the pooled ARK Iohexol Study including Malawi, South Africa, and Uganda; N=2578 for creatinine-based equations; N=2433 for cystatin C-based equations; vertical line represents GFR cut-off of 60ml/min/1.73m^2

Figure 2: Bias, precision, accuracy of GFR estimating equations compared to iohexol GFR for the ARK Study



Accuracy depicted by the percentage of estimates that were greater than 10% (1-P_{10}), and 30% (1-P_{30}) of iohexol GFR. Vertical bars for data points represent the 95% confidence intervals

Figure 3: Predicted prevalence of impaired kidney function in Sub-Saharan Africa using imputed GFR compared to estimated GFR (CKD-EPI creatinine) across West, East, and Southern Africa*



Population-based datasets were derived from ARK Malawi, South Africa and Uganda for ages 20 - 80 years; and AWI-GEN countries including Burkina Faso, Ghana, Kenya, and South Africa for ages 40 - 60 years; datasets for ARK and AWI-Gen South Africa did not overlap; prevalence is not adjusted for age distributions

Declaration of Funding

The Uganda/Malawi study was funded by a project grant from the GSK Africa Non- Communicable Disease Open Lab (Project number: 8111). The SA study was jointly funded by the South African MRC, MRC UK (via the Newton Fund), and GSK Africa Non-Communicable Disease Open Lab (via a supporting grant Project Number: 074). Authors retained control of the final content of the publication.

Additional sources of funding for the SA project were obtained from the International Society for Nephrology (ISN) Clinical Research Program: 15-2-015_Validation of eGFR equations in South Africans (South Africa); the National Health Laboratory Services (NHLS); Faculty of Health Sciences Research Incentive Grant; Grant number: 00128384342035121105000000000000000004550; University of the Witwatersrand; FHS (Wits).

Acknowledgment of Support

For the South African study site, study data were collected and managed using REDCap electronic data capture tools hosted at the University of the Witwatersrand.[44,45] We would like to acknowledge Professor Mboyo-Di-Tamba Vangu, Head of the Department of Nuclear Medicine, Charlotte Maxeke Johannesburg Academic Hospital, Emeritus Professor Geoffrey Candy, University of the Witwatersrand, and Professor Jean F. Botha, Head of the Wits Transplant Unit, Wits Donald Gordon Medical Centre, University of the Witwatersrand for granting the ARK Consortium access to clinical records and radionuclide GFR test results that comprised the external validation dataset. We would like to acknowledge Marike Mapham for the design and layout of the graphics.

References

1. Kaze AD, Ilori T, Jaar BG, Echouffo-Tcheugui JB. Burden of chronic kidney disease on the African continent: a systematic review and meta-analysis. BMC nephrology 2018;19:1-11.

2. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Annals of internal medicine 1999;130:461-70.

3. van Deventer HE, George JA, Paiker JE, Becker PJ, Katz IJ. Estimating glomerular filtration rate in black South Africans by use of the modification of diet in renal disease and Cockcroft-Gault equations. Clinical chemistry 2008;54:1197-202.

4. Eastwood JB, Kerry SM, Plange-Rhule J, et al. Assessment of GFR by four methods in adults in Ashanti, Ghana: the need for an eGFR equation for lean African populations. Nephrology Dialysis Transplantation 2010;25:2178-87.

5. Wyatt CM, Schwartz GJ, Ong'or WO, et al. Estimating kidney function in HIV-infected adults in Kenya: comparison to a direct measure of glomerular filtration rate by iohexol clearance. PloS one 2013;8:e69601.

6. Bukabau JB, Yayo E, Gnionsahé A, et al. Performance of creatinine-or cystatin C–based equations to estimate glomerular filtration rate in sub-Saharan African populations. Kidney international 2019;95:1181-9.

7. Coresh J, Astor BC, McQuillan G, et al. Calibration and random variation of the serum creatinine assay as critical elements of using equations to estimate glomerular filtration rate. American Journal of Kidney Diseases 2002;39:920-9.

8. Rowe C, Sitch AJ, Barratt J, et al. Biological variation of measured and estimated glomerular filtration rate in patients with chronic kidney disease. Kidney international 2019;96:429-35.

9. Ashuntantang G, Osafo C, Olowu WA, et al. Outcomes in adults and children with end-stage kidney disease requiring dialysis in sub-Saharan Africa: a systematic review. The Lancet Global Health 2017;5:e408-e17.

10. Bradshaw C, Kondal D, Montez-Rath ME, et al. Early detection of chronic kidney disease in low-income and middle-income countries: development and validation of a point-of-care screening strategy for India. BMJ global health 2019;4:e001644.

11. Crampin AC, Kayuni N, Amberbir A, et al. Hypertension and diabetes in Africa: design and implementation of a large population-based study of burden and risk factors in rural and urban Malawi. Emerging themes in epidemiology 2016;13:1-14.

12. Asiki G, Murphy G, Nakiyingi-Miiro J, et al. The general population cohort in rural south-western Uganda: a platform for communicable and non-communicable disease studies. International journal of epidemiology 2013;42:129-41.

13. Kahn K, Collinson MA, Gómez-Olivé FX, et al. Profile: Agincourt health and socio-demographic surveillance system. International journal of epidemiology 2012;41:988-1001.

14. Kalyesubula R, Fabian J, Nakanga W, et al. How to estimate glomerular filtration rate in sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study. BMC Nephrology 2020;21:1-12.

15. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. The Journal of the American College of Dentists 2014;81:14.

16. Nakanga WP, Prynn JE, Banda L, et al. Prevalence of impaired renal function among rural and urban populations: findings of a cross-sectional study in Malawi. Wellcome Open Res 2019;4:92.

17. Kalyesubula R, Hau JP, Asiki G, et al. Impaired renal function in a rural Ugandan population cohort. Wellcome Open Res 2018;3:149.

18. Fleming JS, Zivanovic MA, Blake GM, Burniston M, Cosgriff PS. Guidelines for the measurement of glomerular filtration rate using plasma sampling. Nuclear medicine communications 2004;25:759-69.

19. Annesley TM, Clayton LT. Ultraperformance liquid chromatography–tandem mass spectrometry assay for iohexol in human serum. Clinical chemistry 2009;55:1196-202.

20. Nordin G, Ekvall S, Kristoffersson C, et al. Accuracy of determination of the glomerular filtration marker iohexol by European laboratories as monitored by external quality assessment. Clinical Chemistry and Laboratory Medicine (CCLM) 2019;57:1006-11.

21. Peake M, Whiting M. Measurement of serum creatinine–current status and future goals. Clinical biochemist reviews 2006;27:173.

22. Averina M, Wilsgaard T, Watkins H, et al. Why does Russia have such high cardiovascular mortality rates? Comparisons of blood-based biomarkers with Norway implicate non-ischaemic cardiac damage. J Epidemiol Community Health 2020.

23. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16:31-41.

24. Levey AS, Coresh J, Greene T, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Annals of internal medicine 2006;145:247-54.

25. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Annals of internal medicine 2009;150:604-12.

26. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. New England Journal of Medicine 2012;367:20-9.

27. Björk J, Grubb A, Sterner G, Nyman U. Revised equations for estimating glomerular filtration rate based on the Lund-Malmö Study cohort. Scandinavian journal of clinical and laboratory investigation 2011;71:232-9.

28. Pottel H, Hoste L, Dubourg L, et al. An estimated glomerular filtration rate equation for the full age spectrum. Nephrology Dialysis Transplantation 2016;31:798-806.

29. Ramsay M, Crowther N, Tambo E, et al. H3Africa AWI-Gen Collaborative Centre: a resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. Global health, epidemiology and genomics 2016;1.

30. George JA, Brandenburg J-T, Fabian J, et al. Kidney damage and associated risk factors in rural and urban sub-Saharan Africa (AWI-Gen): a cross-sectional population study. The Lancet Global Health 2019;7:e1632-e43.

31. White CA, Akbari A, Allen C, et al. Simultaneous glomerular filtration rate determination using inulin, iohexol, and 99mTc-DTPA demonstrates the need for customized measurement protocols. Kidney International 2020.

32. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int Suppl 2013;3:1-150.

33. Ephraim RKD, Asamoah CA, Abaka-Yawson A, Kwadzokpui PK, Adusei S. Climate change causes changes in biochemical markers of kidney disease. BMC nephrology 2020;21:1-8.

34. Brück K, Stel VS, Gambaro G, et al. CKD prevalence varies across the European general population. Journal of the American Society of Nephrology 2016;27:2135-47.

35. Biörk J, Bäck SE, Sterner G, et al. Prediction of relative glomerular filtration rate in adults: New improved equations based on Swedish Caucasians and standardized plasma-creatinine assays. Scandinavian journal of clinical and laboratory investigation 2007;67:678-95.

36. Poggio ED, Wang X, Greene T, Van Lente F, Hall PM. Performance of the modification of diet in renal disease and Cockcroft-Gault equations in the estimation of GFR in health and in chronic kidney disease. Journal of the American Society of Nephrology 2005;16:459-66.

37. Nyman U, Grubb A, Larsson A, et al. The revised Lund-Malmö GFR estimating equation outperforms MDRD and CKD-EPI across GFR, age and BMI intervals in a large Swedish population. Clinical Chemistry and Laboratory Medicine (CCLM) 2014;52:815-24.

38. Bukabau JB, Sumaili EK, Cavalier E, et al. Performance of glomerular filtration rate estimation equations in Congolese healthy adults: the inopportunity of the ethnic correction. PloS one 2018;13:e0193384.

39. Horio M, Imai E, Yasuda Y, Watanabe T, Matsuo S. Modification of the CKD epidemiology collaboration (CKD-EPI) equation for Japanese: accuracy and use for population estimates. American Journal of Kidney Diseases 2010;56:32-8.

40. Teo BW, Zhang L, Guh JY, et al. Glomerular Filtration Rates in Asians. Adv Chronic Kidney Dis 2018;25:41-8.

41. Wells JC. Body composition of children with moderate and severe undernutrition and after treatment: a narrative review. BMC medicine 2019;17:1-9.

42. Ntenda PAM. Association of low birth weight with undernutrition in preschool-aged children in Malawi. Nutrition journal 2019;18:1-15.

43. Adebisi YA, Ibrahim K, Lucero-Prisno III DE, et al. Prevalence and socio-economic impacts of malnutrition among children in Uganda. Nutrition and metabolic insights 2019;12:1178638819887398.

44. Wasserman P, Segal-Maurer S, Rubin DS. High prevalence of low skeletal muscle mass associated with male gender in midlife and older HIV-infected persons despite CD4 cell reconstitution and viral suppression. Journal of the International Association of Providers of AIDS Care (JIAPAC) 2014;13:145-52.

45. Villamor E, Saathoff E, Mugusi F, Bosch R, Urassa W, Fawzi W. Wasting and body composition of adults with pulmonary tuberculosis in relation to HIV-1 coinfection, socioeconomic status, and severity of tuberculosis. European journal of clinical nutrition 2006;60:163-71.

46. Spearman CW, Sonderup MW. Health disparities in liver disease in sub-Saharan Africa. Liver International 2015;35:2063-71.

47. Paruk IM, Pirie FJ, Motala AA. Non-alcoholic fatty liver disease in Africa: a hidden danger. Global health, epidemiology and genomics 2019;4.

48. Zanetti D, Rao A, Gustafsson S, Assimes TL, Montgomery SB, Ingelsson E. Identification of 22 novel loci associated with urinary biomarkers of albumin, sodium, and potassium excretion. Kidney international 2019;95:1197-208.

49. Fatumo S, Chikowore T, Kalyesubula R, et al. Discovery and fine-mapping of kidney function loci in first genome-wide association study in Africans. bioRxiv 2020.

Measurement of kidney function in Sub-Saharan Africa

Supplementary Appendix

Table of Contents

Because of the density of the data in these tables, the only way we could ensure readability, was to insert them in PDF format:

SECTION S1 – STUDY SETTING AND SAMPLING STRATEGY

The African Research on Kidney Disease (ARK) Study is a multicentre study nested in three longitudinal population studies in Malawi (urban and rural) and Uganda (rural), and a Health and Demographic Surveillance System (HDSS) site in South Africa (rural). The Malawi Epidemiology and Intervention Research Unit (MEIRU) has an urban site in Lilongwe and a rural site in the northern Karonga district; in Uganda the General Population Cohort is located in the Kalunga district, south western Uganda; and in South Africa, the Medical Research Council/Wits University Rural Public Health and Health Transitions Research Unit is located in Bushbuckridge, a rural subdistrict of the Mpumalanga province, in the north eastern part of South Africa.[1] Each partner country obtained prior ethics approval for their respective studies and all participants signed written informed consent. For Uganda, ethics approvals were obtained from the Uganda Virus Research Institute (UVRI-REC-#HS 1978) and the Uganda National Council for Science and Technology (UNCST-#SS 4283). For Malawi, permission was granted from the Malawi National Health Sciences Research Committee (#1072). For South Africa, permission was obtained from the University of the Witwatersrand Human Research Ethics Committee (#160938).

Prior to this study, each country determined their population prevalence of Impaired kidney function based on estimating glomerular filtration rate (eGFR) using the CKD-EPI (creatinine) equation without ethnicity coefficient.[2,3] Then, we classified participants by eGFR stage (G1-5), and selected a subsample of 1000 adults from each country stratified by sex and eGFR stage. We intended to recruit 3000 participants in total, based on a power calculation for the number of study participants needed to examine the accuracy of eGFR equations for predicting iohexol (measured) GFR in each country. We specified this as

having 90% power to detect whether eGFR is within 5% of iohexol GFR at an eGFR of 60mls/min, assuming a standard deviation of 25mls/mins and a two-sided α=0.05. This gives an estimated required sample size of 730 participants, however, to allow for participants who do not wish to participate and technical and screening failures, we oversampled to 1000 participants in each country.

SECTION S2 – STUDY PROCEDURES

Potential participants, 18 years and older, were screened for contraindications to intravenous iohexol administration, namely pregnancy or breastfeeding in women, uncontrolled epilepsy, severe, uncontrolled hypertension, or any acute pyrexial illness.
If eligible, a trained fieldworker obtained written, informed consent in the participant's first language. On the day of iohexol measurement, we recorded blood pressure, temperature, height and weight, and trained nurses established two intravenous (IV) access points in the antecubital fossa (in most cases) of the dominant and contralateral arm. From the dominant arm, we drew a baseline serum sample for creatinine and cystatin C and flushed the cannula with 10ml normal saline. A single bolus of 5ml Omnipaque^{TM} (350mg iodine/ml) was administered intravenously over 30 seconds, followed by a 10ml normal saline flush the cannula was removed. Time zero (T0) was the start time of administration of the intravenous bolus of iohexol. In the contralateral arm, at 5, 120, 180, and 240 minutes from T0, 1ml of aspirated venous blood was discarded, followed by collection of 4mls of venous blood in heparinized tubes. For each sampling time, we recorded the exact time of sample collection. Using a calibrated scale, each syringe was weighed pre- and post-administration of iohexol in grams, to two decimal points. These weights were used to calculate the dose of iohexol administered.

To calculate iohexol GFR adjusted for body surface area (BSA)(mL/min/1.73m^2) from plasma clearance of iohexol, we used the slope-intercept method for three time points (using the exact time of measurement of the samples intended to be taken at 120, 180 and 240 minutes) in the second (slow) exponential phase of iohexol elimination.[4] The extrapolated y-intercept is the calculated concentration of iohexol at time zero, and the coefficient, k, is the slope of the semilog plot of plasma iohexol concentration against time [$P(t)=P_0 \exp(-kt)$]. We adjusted for body surface area using the Haycock formula [$BSA = 0.024265*Wt^{0.5378} * Ht^{0.3964}$] followed by the Brochner-Mortensen (BM) correction for adults, to account for iohexol plasma clearance in the first (rapid) exponential phase [$BM = (0.9908*GFR) – (0.001218*GFR^2)$].[5,6]

SECTION S3 – LABORATORY METHODS AND TESTING

In this section we detail laboratory methods and quality control measures for (i) iohexol in South Africa; (ii) creatinine in Malawi, South Africa, and Uganda; (iii) cystatin C testing in South Africa and Uganda. To test for systematic differences in laboratory procedures between partner countries, we conducted a recalibration study with split sample testing, which was particularly relevant as we knew that Malawi and South Africa were using the modified Jaffe method for creatinine and Uganda was using an enzymatic method. For this recalibration, 20 serum samples from Malawi and 50 serum samples from Uganda were split, and re-assayed in South Africa. Agreement between paired measurements was analysed using Bland Altman plots and, where appropriate, Passing-Bablok regression was used to derive calibration equations to correct for systematic between-country differences.

S3.1 Iohexol

In each partner country, iohexol plasma samples were processed on the same day of collection and stored at -112°F. Iohexol measurement was centralized and each country shipped samples to the National Health Laboratory Services (NHLS) in Johannesburg, South Africa. The NHLS is accredited to the ISO 15189 standard and participates in the Equalis External Quality Assurance Programme for iohexol (Uppsala, Sweden).[7] Using a published method, all iohexol samples were analysed using ultra high performance liquid chromatography-tandem mass spectrometry ((SCIEX (Redwood, CA, USA) 5500 QTRAP (LC-MS/MS)).[8] Certified reference materials for iohexol (CRM: USP H0J211), and ioversol (CRM: USP 34510F) were purchased from Industrial Analytical (Kyalami, South Africa) for the calibration curve and internal standard, respectively. Both compounds were lyophilised and weighed to make a stock standard of 10g/L in deionised water for iohexol, and methanol for ioversol. Working standards were then prepared with further dilutions to create a three-point calibration curve of 50mg/L, 500mg/L, and 1500mg/L for iohexol by spiking the stock standard into drug-free serum. The working solution for ioversol was prepared by diluting the stock standard in methanol to a final concentration of 20mg/L. Internal quality control (IQC) samples were prepared by spiking the certified reference material for iohexol (standard) to create final concentrations of 100 and 1000 mg/L for respective low and high internal quality control (IQC). Coefficients of variation for internal quality control with the iohexol standard at 100 and 1000mg/L were 4.1% and 4.2% respectively. Equalis samples were included in every run as an additional quality check. Iohexol and ioversol were eluted on a gradient profile using a 2.7μm Halo C18 (0.3 x 50mm) column purchased from SCIEX (Redwood, CA, USA). Mass spectrometry was carried out using electrospray ionisation in

positive mode (ESI+) and multiple reaction monitoring was used for identification of iohexol and ioversol with transitions of 821.7>803.7m/z and 807.9>589m/z, respectively. Equalis samples run as quality controls within each batch were within the accepted calculated z-score of < 2.0.

Table S3.1: NHLS External Quality Assurance Compliance (Equalis AB)

Year	Assigned value for iohexol concentration (mg/L)	Measured mean (SD) for iohexol concentration (mg/L)	CV (ratio SD/mean)%
2018			
1a	54.4	55.0 (1.6)	2.9
1b	38.0	37.5 (1.1)	2.9
2a	98.7	99.3 (6.0)	6.0
2b	54.3	52.5 (2.3)	4.4
2019			
1a	38.0	38.7 (2.8)	7.3
1b	19.2	18.5 (0.6)	3.5
2a	56.8	56.7 (2.7)	4.7
2b	19.5	19.4 (1.0)	5.0

S3.2 Creatinine and Cystatin C

Malawi, South Africa, and Uganda each performed their own creatinine measurements. All partner countries standardised their creatinine assays using an isotope-dilution mass spectrometry (IDMS) assay traceable to a standard reference material (967) for creatinine. The modified Jaffe method was used in Malawi and South Africa, and the enzymatic method in Uganda. For cystatin C measurements, Uganda and Malawi samples were processed by the Uganda laboratory, with South African samples processed by the NHLS in Johannesburg. The South African and Ugandan laboratories each procured the Tina-quant® Cystatin C Gen.2 test kits from the same batch and ran all samples after completion of the study. The

assay is an immunoturbidimetric assay standardised to the ERM-DA471/IFCC reference material (Roche Diagnostics, USA).

Table S3.2: ARK laboratory analytic methods for serum creatinine and cystatin C

Laboratory	Laboratory instrument	Laboratory assay
Malawi	Beckman Coulter BD AU480 chemistry analyser	Creatinine: modified Jaffe method, standardised to an isotope-dilution mass spectrometry (IDMS) assay
South Africa	Roche Cobas C501 (6000) analyser	Creatinine: modified Jaffe method, standardised to an isotope-dilution mass spectrometry (IDMS) assay
Uganda	Roche Cobas C501 (6000) analyser	Creatinine: enzymatic method, standardised to an isotope-dilution mass spectrometry (IDMS) assay
South Africa	Roche Cobas E602	cystatin C: immunoturbidimetric method, standardised using ERM-DA471/IFCC reference material
Uganda	Roche Cobas C501 (6000) analyser	cystatin C: immunoturbidimetric method, standardised using ERM-DA471/IFCC reference material

S3.3 Recalibration and split sample testing study for creatinine and cystatin C

Systematic bias may arise between study site laboratories because of different pre-analytic and analytic methods for biomarker measurement. As far as possible, bias can be eliminated in the pre-analytic stages through standardising study protocols prior to initiation of the study. Bias during the analytic stage can be controlled by a recalibration study, where measurements from one study site are recalibrated to the measurements in a reference study site. Recalibration allows for correction of inter-site laboratory differences in time or space, which relate to the types of assay, the manufacturer, and the analytic platform.[9] Despite the inherent quality assurance of manufacturer assays and the availability of standardised reference materials to aid calibration, evidence suggests that some assays do not meet optimal bias limits and calibration differences persist, explaining the need for

rigorous internal and external quality control procedures in each study laboratory. All these potential sources of analytic bias may impact research data. For epidemiological studies involving population-level data, these small systematic differences may result in a shift of the distribution of a biomarker potentially biasing estimates of mean values and the prevalence of dichotomously defined variables, for example, the presence or absence of kidney disease.[9] In this recalibration and split sample testing study, we assessed interlaboratory bias for serum creatinine and cystatin C biomarkers measured in the ARK study partner laboratories. If significant systematic differences were observed, we determined the recalibration equation to correct for the analytic bias. We assessed between-partner laboratory creatinine and cystatin C measurement variability by shipping stored, randomly selected split serum samples from Uganda (n=50) and Malawi (n=20) for repeat testing in South Africa as the reference laboratory.

Quality control procedures

Daily internal quality control was performed as per standard laboratory practice and inter-assay coefficients of variation for each laboratory biomarker were calculated based on analyses of commercial controls. For external quality assurance, standard serum is distributed to the participating laboratory for testing of common analytes. For creatinine, the laboratory in Malawi complied with the requirements of the Thistle QA Laboratory Services (South Africa) requirements, and likewise, the Uganda and South Africa laboratories met the requirements of the College of American Physicians (CAP) Quality Assurance Program. For cystatin C, the South Africa laboratory complied with the requirements of the Equalis External Quality Assurance Program (Uppsala, Sweden). The Uganda and South Africa laboratories are accredited to the ISO 15189 standard.

Recalibration of creatinine

For Uganda and Malawi respectively, we compared the split sample results from the reference laboratory in South Africa using scatter and differential plots. Outliers were flagged for review by the study team and defined as creatinine values more than three standard deviations from the mean difference between paired values.[9] A single outlier for creatinine measurement from Uganda was excluded after confirming it was transcriptional error during data entry, and there were no further outliers. Agreement analysis was performed using Bland-Altman plots and Lin's concordance correlation coefficient.

Recalibration of cystatin C

We used a similar approach to the recalibration of creatinine. Cystatin C measurements for Malawi and Uganda were performed in the Uganda laboratory, so for this component of the study we compared split samples from Uganda to repeat samples in South Africa reference laboratory. Initially, scatter and differential plots were examined, no outliers were identified, and the agreement analysis proceeded. A cumulative sum (Cusum) test was performed to assess the linearity and regression analysis (Passing-Bablok) was used to determine the calibration function for the relationship between the paired cystatin C values - assuming the regression equation Y= A(intercept) +B(slope)*X. Statistical calculations were performed in Stata/SE, version 16.1.

Calibration for serum creatinine

Compared to split samples for creatinine in the South Africa laboratory, the correlation was good, and the relationship linear for Malawi and Uganda creatinine values (Table S2, Figure S1). With South Africa as reference, the bias was -2.70µmol/L (95% CI -6.70 – 1.59) and +9.29 µmol/L (95% CI 7.25 – 11.33 µmol/L) for Malawi and Uganda creatinine samples,

respectively. Since laboratories in Malawi and South Africa used the modified Jaffe method, and Uganda used the enzymatic method, the systematic bias was ascribed to these methodological differences. Ideally, KDIGO recommends the enzymatic method in preference to Jaffe, as the former is less biased and less prone to interference.[10] On this basis, we recalibrated creatinine values for Malawi and South Africa by a constant factor of +9.29µmol/L to align with the Uganda laboratory.

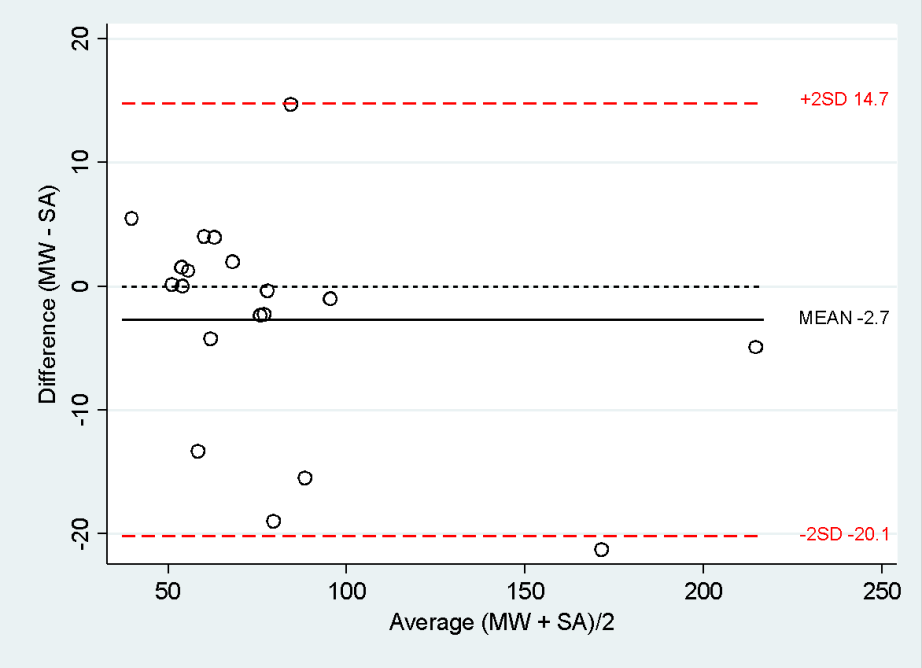
Calibration for cystatin C

When comparing Uganda cystatin C values to South Africa, the correlation was good, and the relationship linear with a bias of 0.10mg/dL (95% CI 0.36 – 0.17). With South Africa as the reference laboratory for the split sample, Passing-Bablok regression analysis was performed to determine the constants for the recalibration equation. (Table S2; Figure S2). The resulting regression coefficients (slope and intercept) were used to recalibrate the Uganda and Malawi cystatin C measurements for comparability to the South Africa reference laboratory values.

Table S3.3: Agreement and recalibration coefficients for creatinine and cystatin C

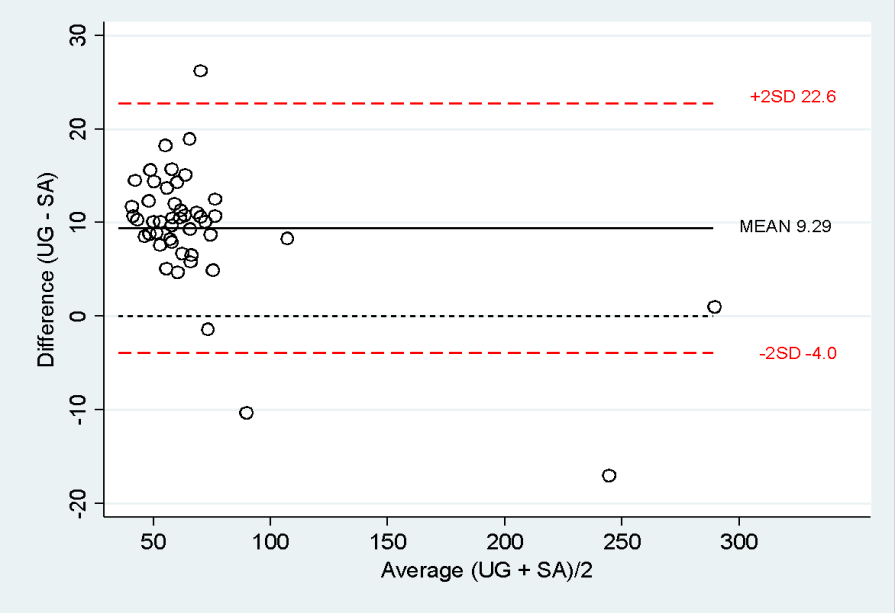
Country	Analyte	No of excluded outliers	R^2 Lin's Concordance Correlation Coefficient	Intercept	Slope
Malawi	creatinine(µmol/L)	0	0.977	n/a	n/a
Uganda	creatinine(µmol/L)	1	0.967	n/a	n/a
Uganda	Cystatin C (mg/dl)	0	0.942	0.178 (95% CI 0.090 – 0.349)	0.922 (95% CI 0.776 – 1)

Figure S3.1: Agreement: serum creatinine measurements for Malawi and South Africa



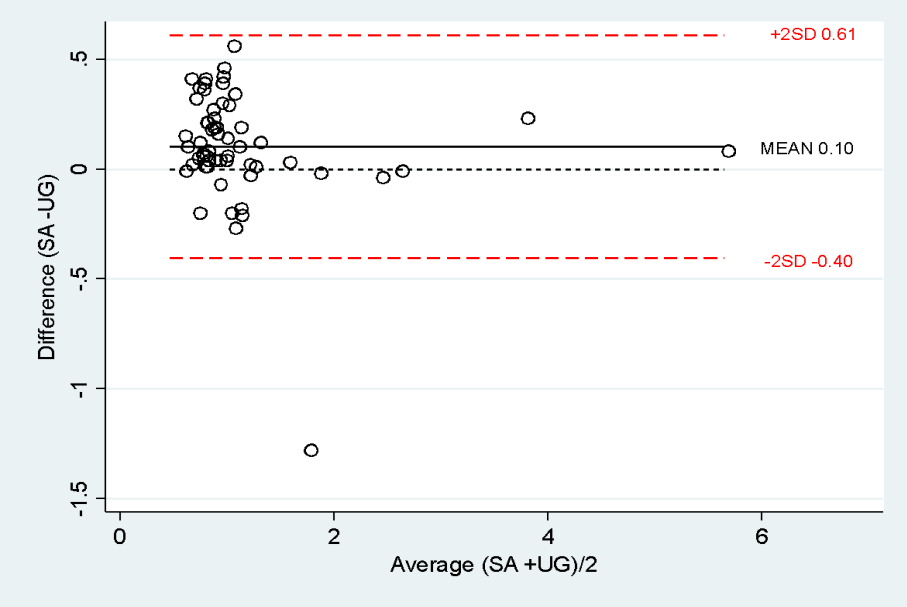
*Dashed red lines represent the 95% confidence intervals for creatinine in the calibration samples.

Figure S3.2: Agreement: serum creatinine measurements for Uganda and South Africa



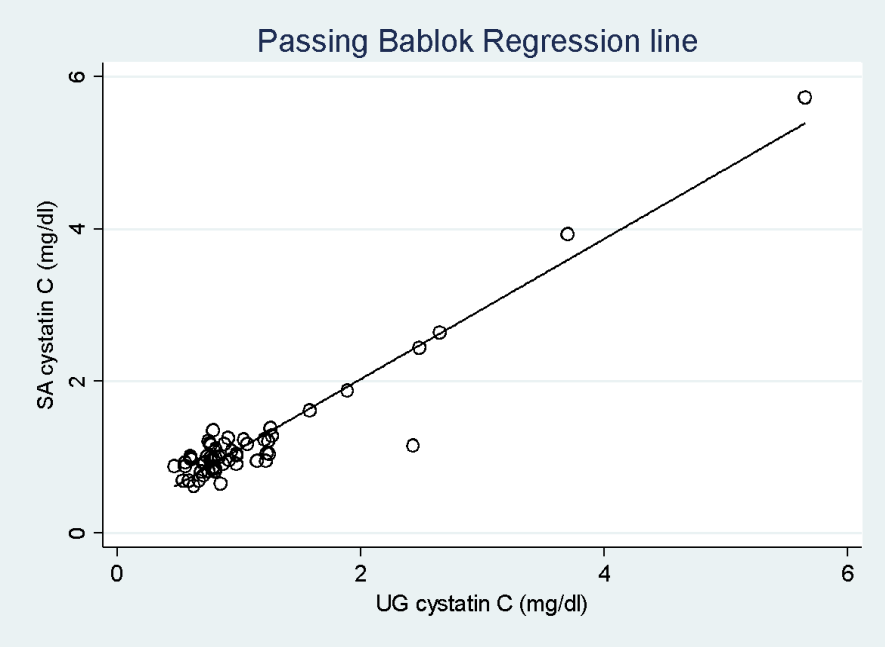
*Dashed red lines represent the 95% confidence intervals for creatinine in the calibration samples.

Figure S3.3: Agreement: serum cystatin C measurements in Uganda and South Africa



*Dashed red lines represent the 95% confidence intervals for cystatin C in the calibration samples.

Figure S3.4: Recalibration of cystatin C in South Africa and Uganda



SECTION S4 – GFR PREDICTION EQUATIONS

For all equations, we used standard conventional units (mg/dL) for serum creatinine (sCr) values rounded to the nearest 100th of a whole number. When sCr was expressed as standard international (SI) units (µmol/L), we rounded to the nearest whole number. The formula to convert sCr from conventional to SI units = [sCr (conventional units) x 88.4]. The units for serum cystatin C (scysC) are mg/l, and all equations were adjusted for body surface area (BSA) with units for GFR as mL/min/1.73m^2. We evaluated the performance of the following serum creatinine and serum cystatin C-based GFR prediction equations:

Cockroft Gault equation adjusted for BSA

GFR = [140-age (years) x weight (kg) x (0.85 if female) x 1.73m^2)] / [sCr x BSA^5 (m^2)]

In its original form, the Cockroft Gault equation did not adjust for body surface area (BSA).[11] However, this adjustment is required when comparing performance to other eGFR equations (all of which are adjusted for BSA). For our analyses we adjusted Cockroft Gault for BSA using the Haycock formula.[5]

4-variable Modification of Diet in Renal Disease (4-vMDRD) equation[12*]

GFR = 175 x sCr-1.154 x age-0.203(years) (x 0.742 if female) (x 1.1212 if African American)

*re-expressed for IDMS assays traceable to a standard reference material for creatinine

2009 Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) creatinine equation[13]

GFR = 141 × min(sCr/κ, 1)$^{\alpha}$ × max(sCr/κ, 1)$^{-1.209}$ × 0.993^{age} (× 1.018 if female) (× 1.159 if African American)

κ is 0.7mg/dl (62 µmol/L) for females and 0.9mg/dl (80 µmol/L) for males
α is -0.329 for females and -0.411 for males

min indicates the minimum of sCr/κ or 1
max indicates the maximum of sCr/κ or 1

Equations expressed for specified sex and serum creatinine level[10]

Female ≤0.7 mg/dl (≤62 mmol/L) = 144 x (sCr/62)$^{-0.329}$ x 0.993^{age} [x 1.159 if African American]
Female >0.7 mg/dl (>62 mmol/L) = 144 x (sCr/62)$^{-1.209}$ x 0.993^{age} [x 1.159 if African American]
Male ≤0.9 mg/dl (≤80 mmol/L) = 141 x (sCr/80)$^{-0.411}$ x 0.993^{age} [x 1.159 if African American]
Male >0.9 mg/dl (>80 mmol/L) = 141 x (sCr/80)$^{-1.209}$ x 0.993^{age} [x 1.159 if African American]

2012 CKD-EPI cystatin C equation[14]

GFR = 133 x min(scysC/0.8, 1)$^{-0.499}$ x max(scysC/0.8, 1)$^{-1.328}$ x 0.996^{age} (x 0.932 if female)

min indicates the minimum of scysC/0.8 or 1
max indicates the maximum of scysC/0.8 or 1

Equations expressed for serum cystatin C level[10]

scysC ≤0.8: 133 x (scysC/0.8)$^{-0.499}$ x 0.996^{age} [x 0.932 if female]
scysC >0.8: 133 x (scysC/0.8)$^{-1.328}$ x 0.996^{age} [x 0.932 if female]

2012 CKD-EPI creatinine-cystatin C equation[14]

GFR = 135 x $\min(sCr/\kappa, 1)^{\alpha}$ x $\max(SCr/\kappa, 1)^{-0.601}$ x $\min(scysC/0.8, 1)^{-0.375}$ x $\max(scysC/0.8, 1)^{-0.711}$ x 0.995^{age} (x 0.969 if female) (x 1.08 if African American)

κ is 0.7mg/dl (62 µmol/L) for females and 0.9mg/dl (80 µmol/L) for males
α is -0.248 for females and -0.207 for males

min indicates the minimum of sCr/κ or 1
max indicates the maximum of sCr/κ or 1
min(scysC/0.8, 1) indicates the minimum of scysC/0.8 or 1
max(scysC/0.8,1) indicates the maximum of scysC/0.8 or 1

Equations expressed for specified sex, serum creatinine, and serum cystatin C level[10,15]

Female sCr≤62 cysC≤0.8 = 130 x $(sCr/62)^{-0.248}$ x $(scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]
Female sCr≤62 cysC>0.8 = 130 x $(sCr/62)^{-0.248}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]
Female sCr>62 cysC≤0.8 = 130 x $(sCr/62)^{-0.601}$ x $(scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]
Female sCr>62 cysC>0.8 = 130 x $(sCr/62)^{-0.601}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]
Male sCr≤80 cysC≤0.8 = 135 x $(sCr/80)^{-0.207}$ x $(/scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]
Male sCr≤80 cysC>0.8 = 135 x $(sCr/80)^{-0.207}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]
Male sCr≤80 cysC≤0.8 = 135 x $(sCr/80)^{-0.601}$ x $(scysC/0.8)^{-0.375}$ x 0.995^{age} [x 1.08 if African American]
Male sCr≤80 cysC>0.8 = 135 x $(sCr/80)^{-0.601}$ x $(scysC/0.8)^{-0.711}$ x 0.995^{age} [x 1.08 if African American]

Revised Lund-Malmö Study equation[15]

GFR = $eX–0.0158\times Age+0.438\times \ln(Age)$

Female sCr<150: X = $2.50+0.0121 \times (150–sCr)$
Female sCr≥150: X = $2.50–0.926 \times \ln(sCr/150)$
Male sCr<180: X = $2.56+0.00968 \times (180–sCr)$
Male sCr≥180: X = $2.56–0.926 \times \ln(sCr/180)$

Full Age Spectrum (FAS) creatinine equation[16]

$$GFR = \frac{107.3}{(sCr/Q)} \text{ (if ≤40 years)}$$

$$GFR = \frac{107.3}{(sCr/Q)} \times 0.988^{(Age\ -40)} \text{ (if >40 years)}$$

[sCr/Q (female) = 0.70mg/dL or 61.88µmol/L; sCr/*Q (male) = 0.90mg/dL or 79.56µmol/L.[17]]

Population reference creatinine for the FAS (creatinine) equation in the ARK Study

One of the variables needed to estimate GFR using the FAS equation (creatinine) is a population-specific reference serum creatinine measurement.[16,17] We used the population prevalence data from the baseline ARK studies to establish country specific population reference creatinine measures. To do this, we included all serum creatinine measures ≥30µmol/L from apparently healthy adults, defined as those without hypertension, HIV infection, diabetes, obesity (BMI >30kg/m^2), and we excluded all participants recruited for this iohexol measured GFR study to ensure independent datasets. The median creatinine for men and women, by country, was used to calculate the Q-value for the FAS equation.

Table S4.1: Population reference serum creatinine measurements for the FAS equation, by country and sex

Country	Sex	Sample size (N)	Serum Creatinine (median, µmol/L)	Serum Creatinine (median, mg/dL)
Malawi	Female	1693	67	0.76
	Male	1542	84	0.95
South Africa	Female	269	64	0.73
	Male	337	81	0.92
Uganda	Female	2111	59	0.67
	Male	1585	72	0.81

FIGURES AND TABLES

Figure S4: Derivation of the ARK iohexol GFR (mGFR) study participants

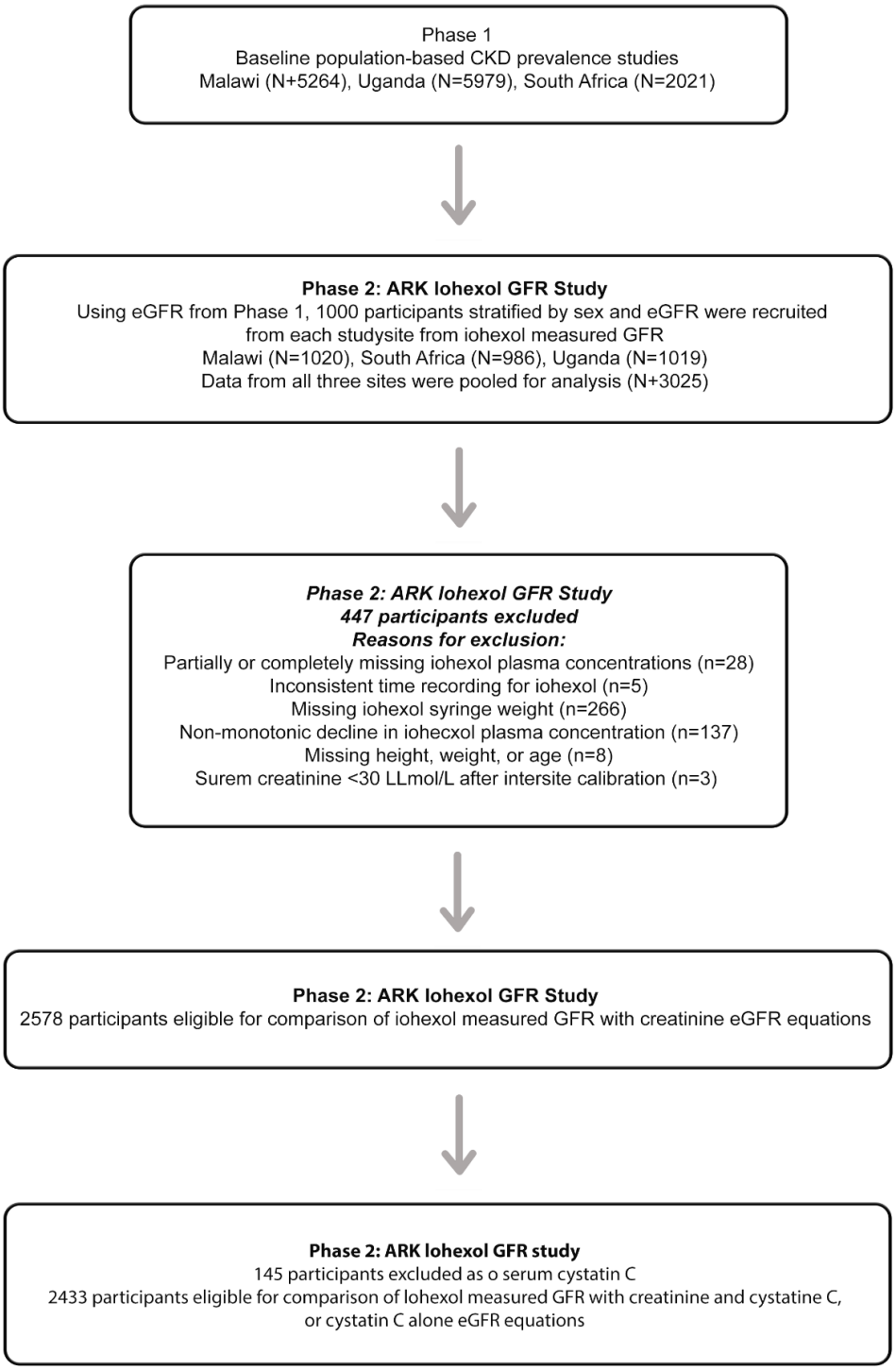


Table S5: Characteristics of the ARK iohexol GFR study participants before (N=3025) and after exclusions (N=2578), by sex, by country, and overall

	Pooled Sample (N=3025)							Pooled Sample after exclusions (N=2578)						
Characteristic[1]	**Malawi**		**South Africa**		**Uganda**		**Overall**	**Malawi**		**South Africa**		**Uganda**		**Overall**
	Females	**Males**	**Females**	**Males**	**Females**	**Males**		**Females**	**Males**	**Females**	**Males**	**Females**	**Males**	
	N=542	**N=477**	**N=664**	**N=322**	**N=561**	**N=458**	**N=3025**	**N=474**	**N=424**	**N=636**	**N=311**	**N=413**	**N=320**	**N=2578**
Age - yr	53.2 (13.4)	53.0 (16.3)	45.7 (14.4)	44.4 (16.3)	50.9 (14.4)	51.5 (15.2)	49.9 (15.2)	53.1 (13.5)	52.8 (16.3)	45.6 (14.4)	44.5 (16.3)	50.6(14.2)	51.5 (15.1)	49.5 (15.2)
Age group category														
<40 yr	77 (14.3%)	95 (19.9%)	251 (37.8%)	143 (44.4%)	126 (22.5%)	106 (23.1%)	798 (26.4%)	70 (14.8%)	87 (20.5%)	244 (38.4%)	138 (44.4%)	92 (22.3%)	74 (23.1%)	705 (27.4%)
40-60 yr	304 (56.6%)	235 (49.3%)	293 (44.1%)	113 (35.1%)	311 (55.4%)	219 (47.8%)	1475 (48.8%)	269 (56.8%)	209 (49.3%)	280 (44.0%)	109 (35.1%)	232 (56.2%)	158 (49.4%)	1257 (48.8%)
>60 yr	156 (29.1%)	147 (30.8%)	120(18.1%)	66 (20.5%)	124 (22.1%)	133 (29.0%)	752 (24.9%)	135 (28.5%)	128 (30.2%)	112 (17.6%)	64 (20.6%)	89 (21.6%)	88 (27.5%)	616 (23.9%)
Body mass index[2]	26.63 (5.78)	23.25 (3.98)	30.11 (6.36)	25.04 (5.10)	23.91 (4.49)	21.05 (3.09)	25.35 (5.87)	26.76 (5.86)	23.11 (3.86)	30.14 (6.41)	25.02 (5.08)	23.92 (4.47)	21.03 (3.12)	25.61 (5.98)
Body mass index category[3]														
<18.5 (underweight)	18 (3.3%)	27 (5.7%)	9 (1.4%)	15 (4.7%)	32 (5.7%)	89 (19.5%)	190 (6.3%)	14 (3.0%)	25 (5.9%)	9 (1.4%)	14 (4.5%)	21 (5.1%)	63 (19.7%)	146 (5.7%)
18.5-24.9 (normal)	224 (41.3%)	323 (67.9%)	146 (22.0%)	166 (51.6%)	344 (61.4%)	318 (69.7%)	1521 (50.4%)	194 (40.9%)	293 (69.1%)	140 (22.0%)	162 (52.1%)	259 (62.7%)	222 (69.4%)	1270 (49.3%)
25.0-29.9 (overweight)	165 (30.4%)	99 (20.8%)	190 (28.6%)	87 (27.0%)	130 (23.2%)	44 (9.7%)	715 (23.7%)	148 (31.2%)	86 (20.3%)	182 (28.6%)	84 (27.0%)	91 (22.0%)	32 (10.0%)	623 (24.2%)
>=30.0 (obese)	135 (24.9%)	27 (5.7%)	319 (48.0%)	54 (16.8%)	54 (9.6%)	5 (1.1%)	594 (19.7%)	118 (24.9%)	20 (4.7%)	305 (48.0%)	51 (16.4%)	42 (10.2%)	3 (0.9%)	539 (20.9%)
Weight (kg)	64.78 (15.28)	63.59 (12.08)	78.66 (17.28)	74.72 (16.81)	57.59 (11.81)	57.27 (9.57)	66.24 (16.41)	65.00 (15.48)	63.24 (11.82)	78.70 (17.37)	74.64 (16.71)	57.56 (11.89)	56.95 (9.64)	67.06 (16.72)
Height (cm)	155.78 (6.03)	165.26 (6.83)	161.59 (5.96)	172.58 (7.13)	155.07 (6.70)	164.81 (6.45)	161.57 (8.49)	155.67 (5.99)	165.28 (6.83)	161.56 (5.91)	172.60 (7.14)	154.98 (6.83)	164.42 (6.36)	161.72 (8.53)
Body Surface Area (m^2)[4]	1.68 (0.22)	1.71 (0.19)	1.90 (0.23)	1.89 (0.24)	1.58 (0.18)	1.61 (0.16)	1.73 (0.24)	1.68 (0.22)	1.70 (0.18)	1.90 (0.23)	1.89 (0.24)	1.58 (0.19)	1.61 (0.16)	1.74 (0.24)
Serum creatinine (mg/dL)[5]	0.73 (0.22)	0.94 (0.36)	0.61 (0.16)	0.83 (0.24)	0.74 (0.20)	0.88 (0.27)	0.77 (0.27)	0.73 (0.20)	0.92 (0.30)	0.61 (0.16)	0.83 (0.23)	0.74 (0.20)	0.89 (0.31)	0.77 (0.25)
Serum creatinine (µmol/L)[6]	64.3 (19.1)	83.0 (31.6)	54.2 (13.9)	73.7 (20.9)	65.8 (17.6)	77.7 (24.3)	68.3 (23.6)	64.7 (17.8)	81.6 (26.1)	54.1 (13.9)	73.4 (20.2)	65.5 (17.7)	78.6 (27.3)	67.8 (22.5)
Serum cystatin C (mg/dL)[7]	0.98 (0.40)	1.03 (0.34)	0.99 (0.28)	1.01 (0.29)	0.88 (0.23)	0.90 (0.28)	0.97 (0.31)	0.97 (0.34)	1.01 (0.30)	0.99 (0.27)	1.01 (0.28)	0.88 (0.23)	0.91 (0.31)	0.97 (0.29)
Iohexol GFR[8]	72.6 (25.7)	78.2 (29.3)	78.1 (33.4)	83.1 (37.3)	82.7 (35.8)	94.4 (43.7)	80.0 (33.7)	74.8 (20.1)	79.1 (20.4)	78.6 (25.0)	82.0 (26.4)	86.1 (26.3)	97.1 (31.2)	81.9 (25.6)
Iohexol GFR category[9]														
≥90 ml/min/1.73m²	104 (20.3%)	127 (28.2%)	211 (31.8%)	132 (41.0%)	164 (37.9%)	197 (56.6%)	935 (34.3%)	100 (21.1%)	124 (29.3%)	205 (32.2%)	130 (41.8%)	160 (38.7%)	190 (59.4%)	909 (35.3%)
60-89 ml/min/1.73m²	287 (56.1%)	225 (50.0%)	290 (43.7%)	117 (36.3%)	199 (46.0%)	94 (27.0%)	1212 (44.4%)	271 (57.2%)	220 (51.9%)	282 (44.3%)	113 (36.3%)	195 (47.2%)	87 (27.2%)	1168 (45.3%)
45-59 ml/min/1.73m²	86 (16.8%)	62 (13.8%)	95 (14.3%)	41 (12.7%)	45 (10.4%)	33 (9.5%)	362 (13.3%)	76 (16.0%)	61 (14.4%)	92 (14.5%)	40 (12.9%)	42 (10.2%)	29 (9.1%)	340 (13.2%)
30-44 ml/min/1.73m²	22 (4.3%)	24 (5.3%)	51 (7.7%)	22 (6.8%)	11 (2.5%)	12 (3.5%)	142 (5.2%)	21 (4.4%)	18 (4.3%)	49 (7.7%)	21 (6.8%)	10 (2.4%)	10 (3.1%)	129 (5.0%)
<30 ml/min/1.73m²	13 (2.5%)	12 (2.7%)	17 (2.6%)	10 (3.1%)	14 (3.2%)	12 (3.5%)	78 (2.9%)	6 (1.3%)	1 (0.2%)	8 (1.3%)	7 (2.3%)	6 (1.45%)	4 (1.3%)	32 (1.2%)
Estimated GFR[10]	83.3 (19.9)	85.7 (20.8)	99.4 (19.2)	98.4 (21.4)	93.3 (19.8)	98.2 (19.7)	92.9 (21.0)	82.6 (19.4)	86.2 (20.1)	99.7 (19.1)	98.5 (21.3)	93.7 (19.9)	97.8 (20.5)	93.2 (21.0)
Estimated GFR category[9]														
≥90 ml/min/1.73m²	205 (38.2%)	212 (44.4%)	486 (73.2%)	227 (71.0%)	331 (59.0%)	336 (73.4%)	1797 (59.5%)	174 (36.7%)	194 (45.8%)	468 (73.6%)	220 (70.7%)	241 (58.4%)	237 (74.1%)	1534 (59.5%)
60-89 ml/min/1.73m²	273 (50.8%)	216 (45.3%)	154 (23.2%)	74 (23.0%)	201 (35.8%)	107 (23.4%)	1025 (34.0%)	245 (51.7%)	189 (44.6%)	146 (23.0%)	71 (22.8%)	154 (37.3%)	69 (21.6%)	874 (33.9%)
45-59 ml/min/1.73m²	44 (8.2%)	29 (6.1%)	18 (2.7%)	14 (4.4%)	20 (3.6%)	6 (1.3%)	131 (4.3%)	41 (8.7%)	26 (6.1%)	17 (2.7%)	14 (4.5%)	13 (3.2%)	5 (1.6%)	116 (4.5%)
30-44 ml/min/1.73m²	12 (2.2%)	12 (2.5%)	5 (0.8%)	6 (1.9%)	6 (1.1%)	4 (0.9%)	45 (1.5%)	12 (2.5%)	10 (2.4%)	4 (0.6%)	5 (1.6%)	3 (0.7%)	4 (1.3%)	38 (1.5%)
<30 ml/min/1.73m²	3 (0.6%)	8 (1.7%)	1 (0.2%)	1 (0.3%)	3 (0.5%)	5 (1.1%)	21 (0.7%)	2 (0.4%)	5 (1.2%)	1 (0.2%)	1 (0.3%)	2 (0.5%)	5 (1.6%)	16 (0.6%)

[1]All data reported as mean (standard deviation) unless otherwise stated; categories reported as number (%); percentages may sum to +/-100 due to rounding; [2]Body mass index (BMI) calculated by dividing weight (kilograms) by height squared (metres); [3]BMI category: WHO classification for obesity[22]; [4]Body surface area calculated according to the Haycock formula[5]; [5] Serum creatinine in mg/dL (conventional units): to convert to µmol/L, multiply by 88.4; creatinine data unadjusted for the calibration study; [6]Serum creatinine in µmol/L (SI units): to convert to mg/dl, divide by 88.4; creatinine data unadjusted for the calibration study; [7]Serum cystatin C in mg/dL: cystatin C data unadjusted for the calibration study; [8]Iohexol GFR: ml/min/1.73m²; [9]GFR category as per KDIGO[10]; [10]Estimated GFR: ml/min/1.73m² using CKD-EPI creatinine equation without ethnicity coefficient. *For the pooled sample (N=3025)*: Overall: N=3024 for sex; N=3019 for age, age category; N=3020 for height, weight, BMI, BMI category, BSA; N=2729 for iohexol GFR, iohexol GFR category; Iohexol GFR reported as median (IQR). Malawi females: N=537 for age and age category; N=529 for serum creatinine; N=511 for serum cystatin C; N=512 for iohexol GFR, iohexol GFR category. Malawi males: N=476 for height, weight, BMI, BMI category, BSA; N=467 for serum creatinine; N=457 for serum cystatin C; N=450 for iohexol GFR, iohexol GFR category. Uganda females: N=560 for height, weight, BMI, BMI category; N=465 for serum cystatin C; N=433 for iohexol GFR, iohexol GFR category. Uganda males: N=456 for height, weight, BMI, BMI category, BSA; N=385 for serum cystatin C. South Africa females: N=659 for serum cystatin C; N=348 for iohexol GFR, iohexol GFR category. *For the pooled sample after exclusions (N=2578):* Overall N=2433 for cystatin C.

Figure S5: Distribution of iohexol GFR overall and by country

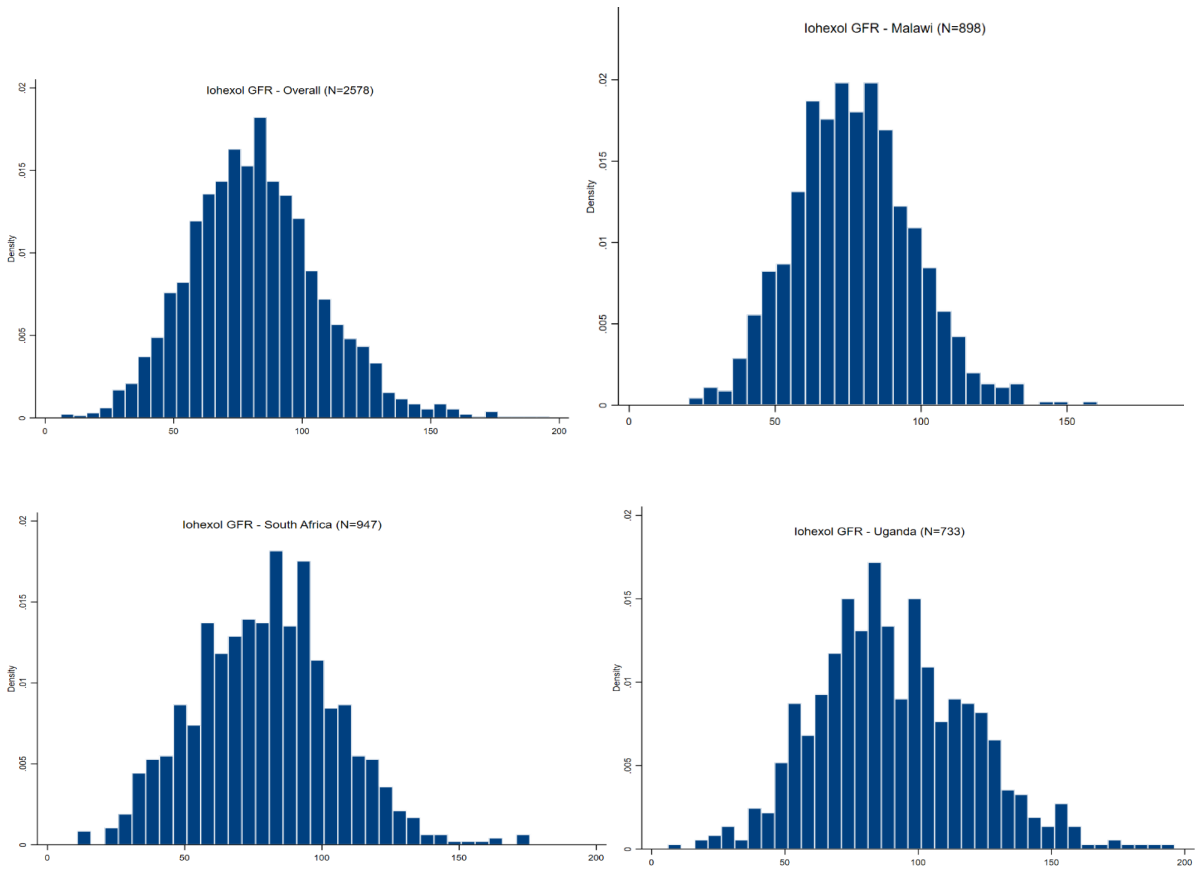


Figure S6: Volume of distribution overall and by country

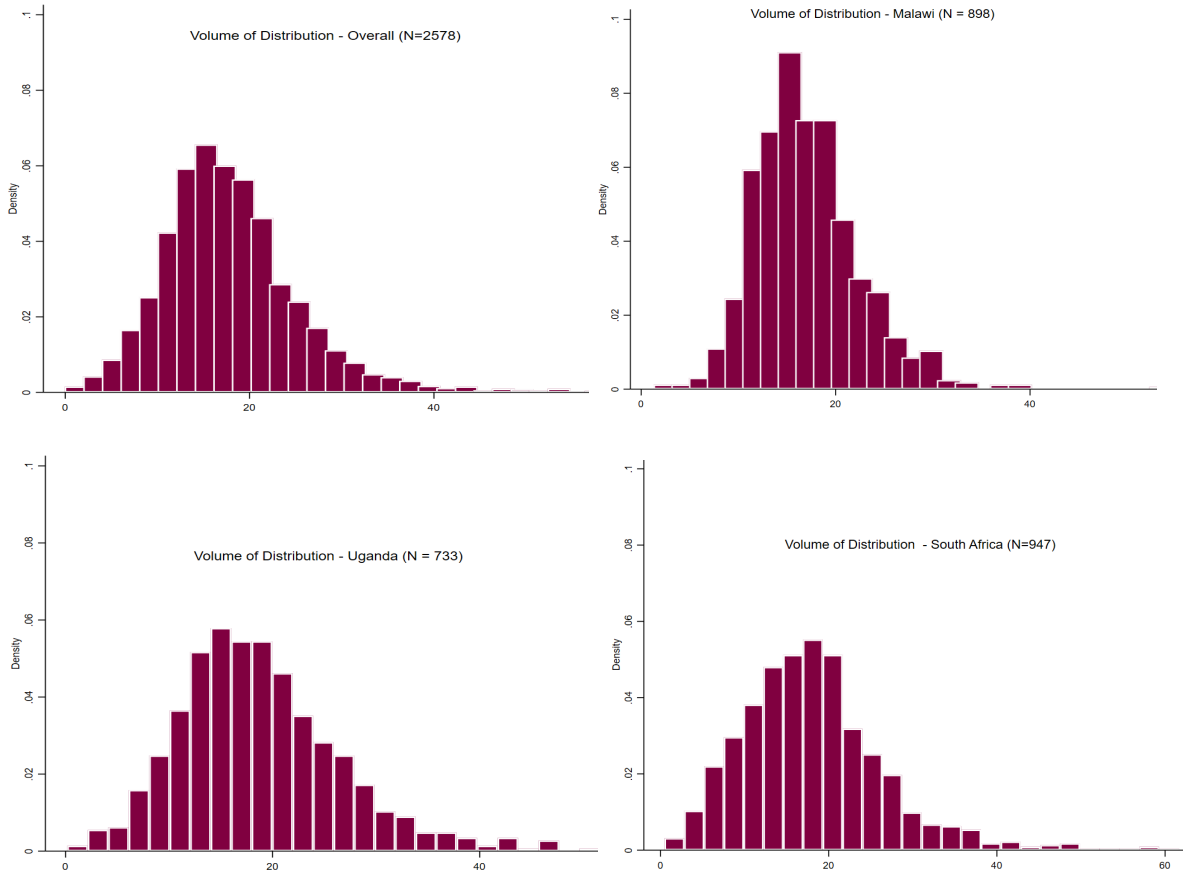


Figure S7: Cumulative distribution plot: correlation coefficient (r) for the slope-intercept iohexol GFR derivation overall and by country

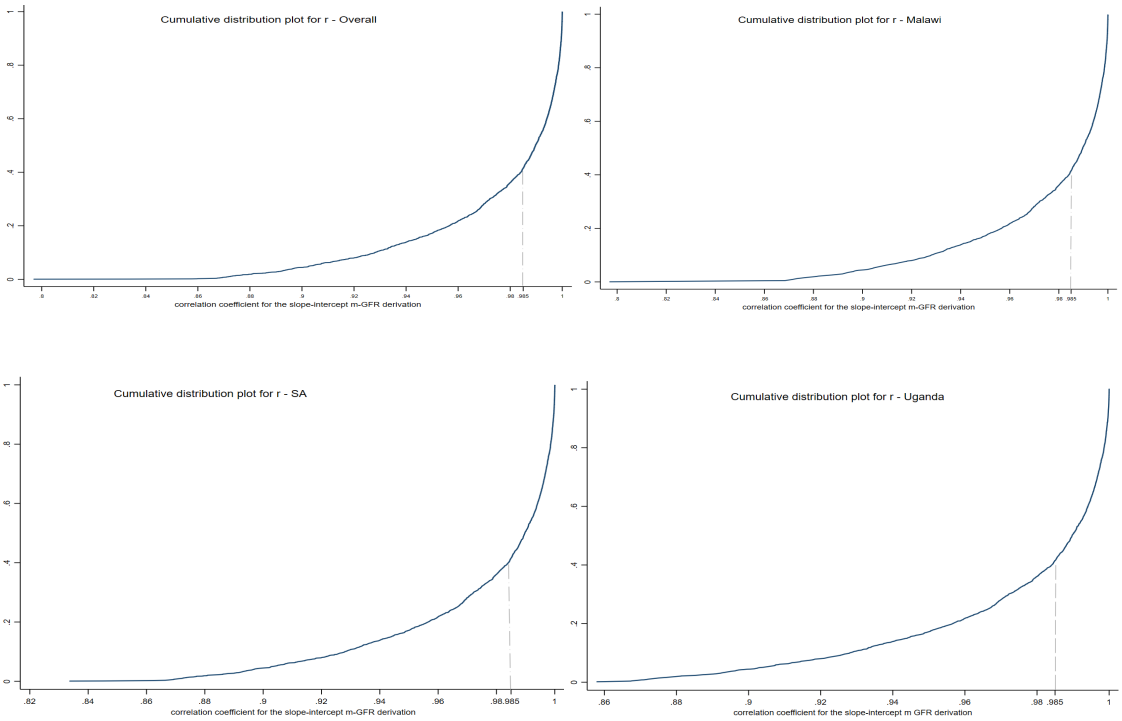
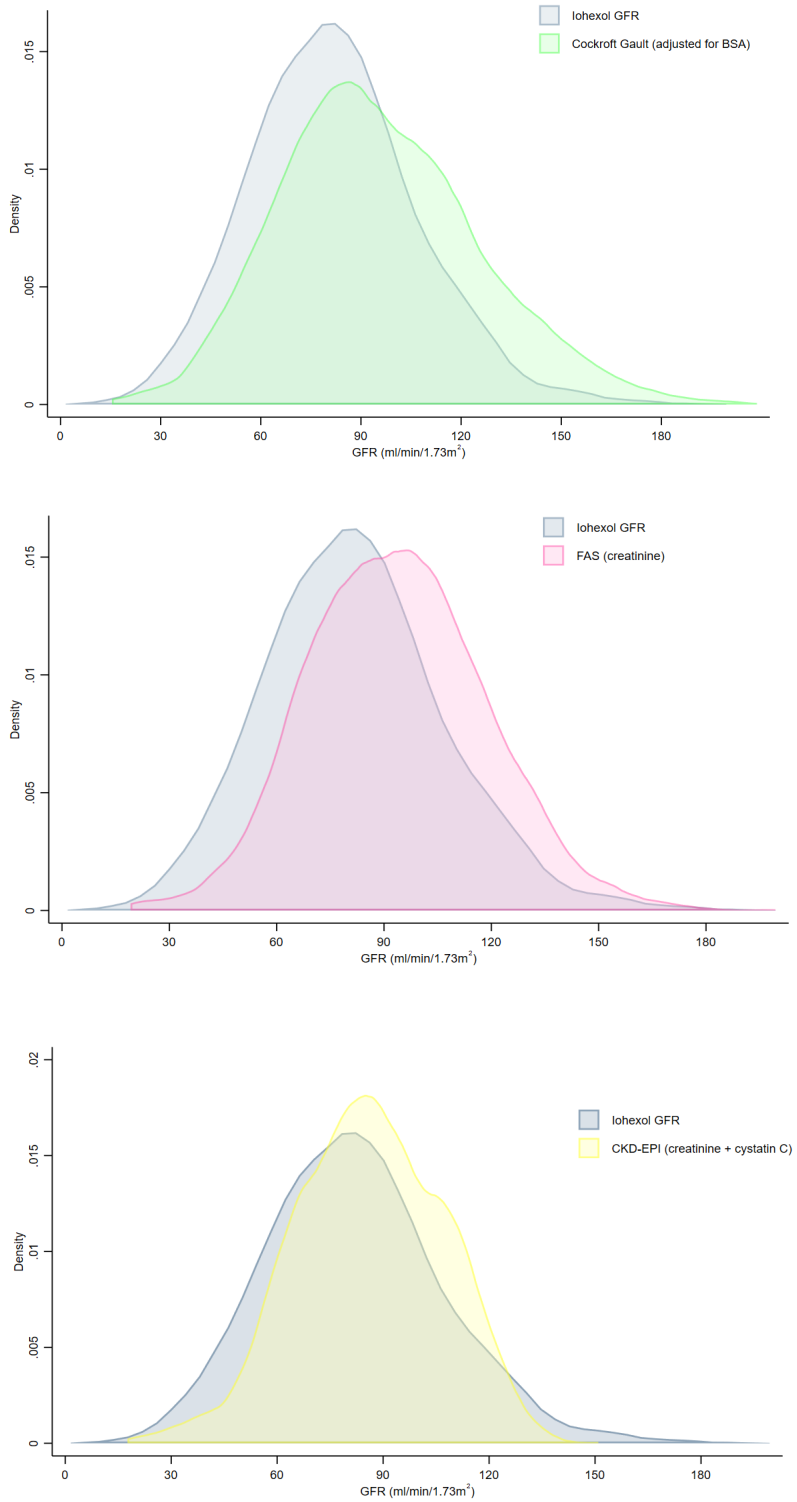
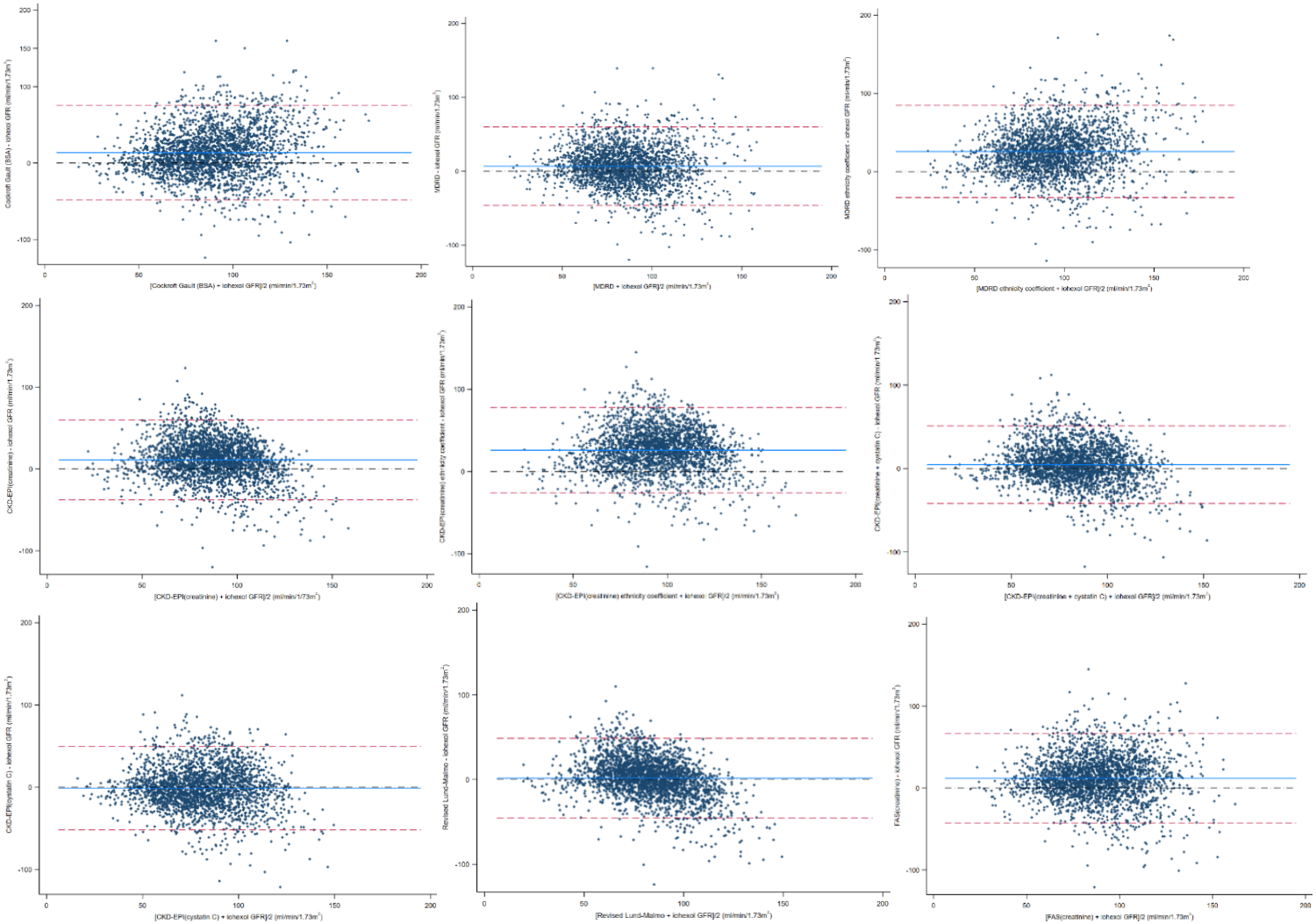


Figure S8: GFR distribution: comparing eGFR equations to iohexol GFR for the ARK Study



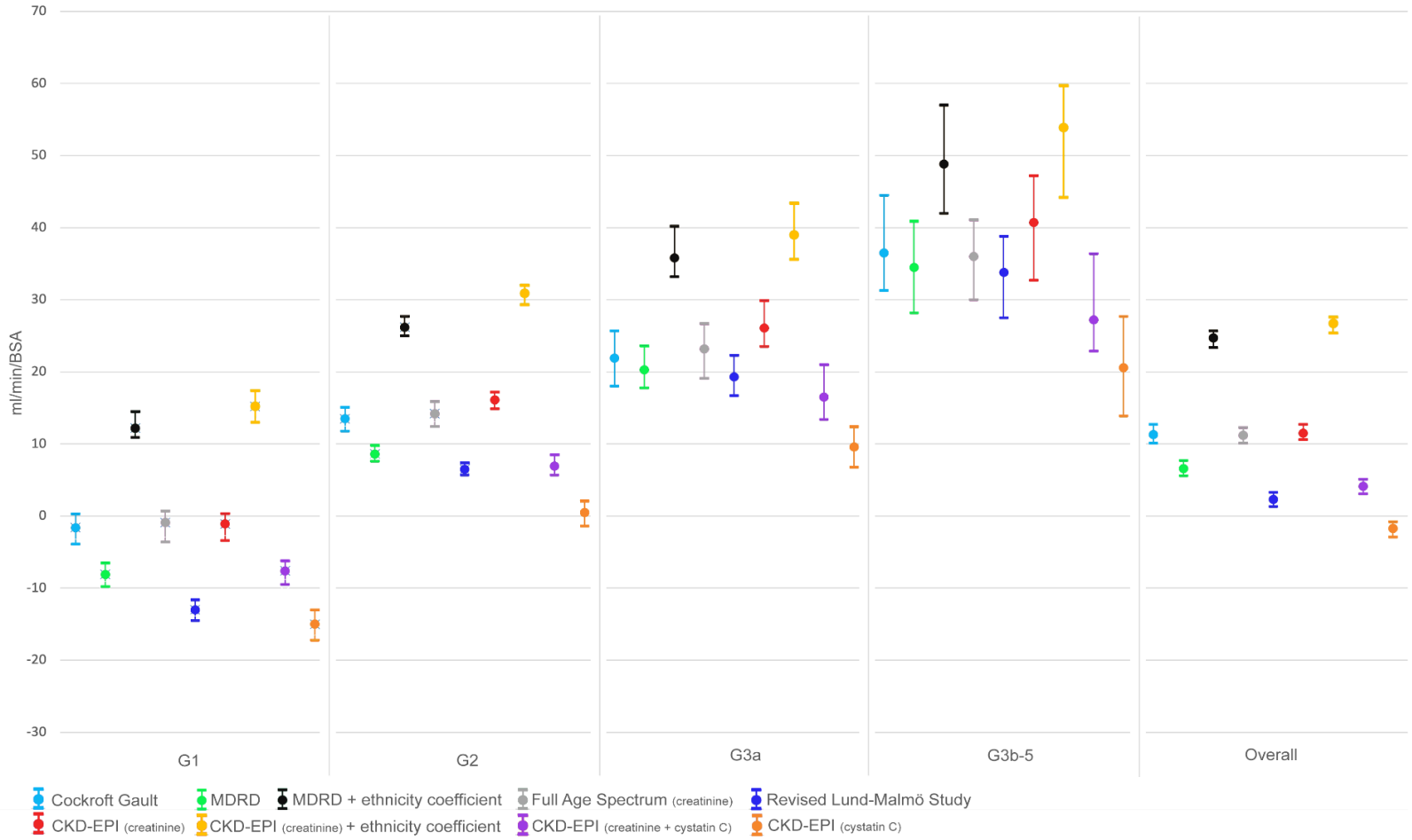
Data shown are from the pooled ARK Iohexol Study including Malawi, South Africa, and Uganda; N=2578 for creatinine-based equations; N=2433 for cystatin C-based equations

Figure S9: Agreement between eGFR equations and iohexol GFR overall



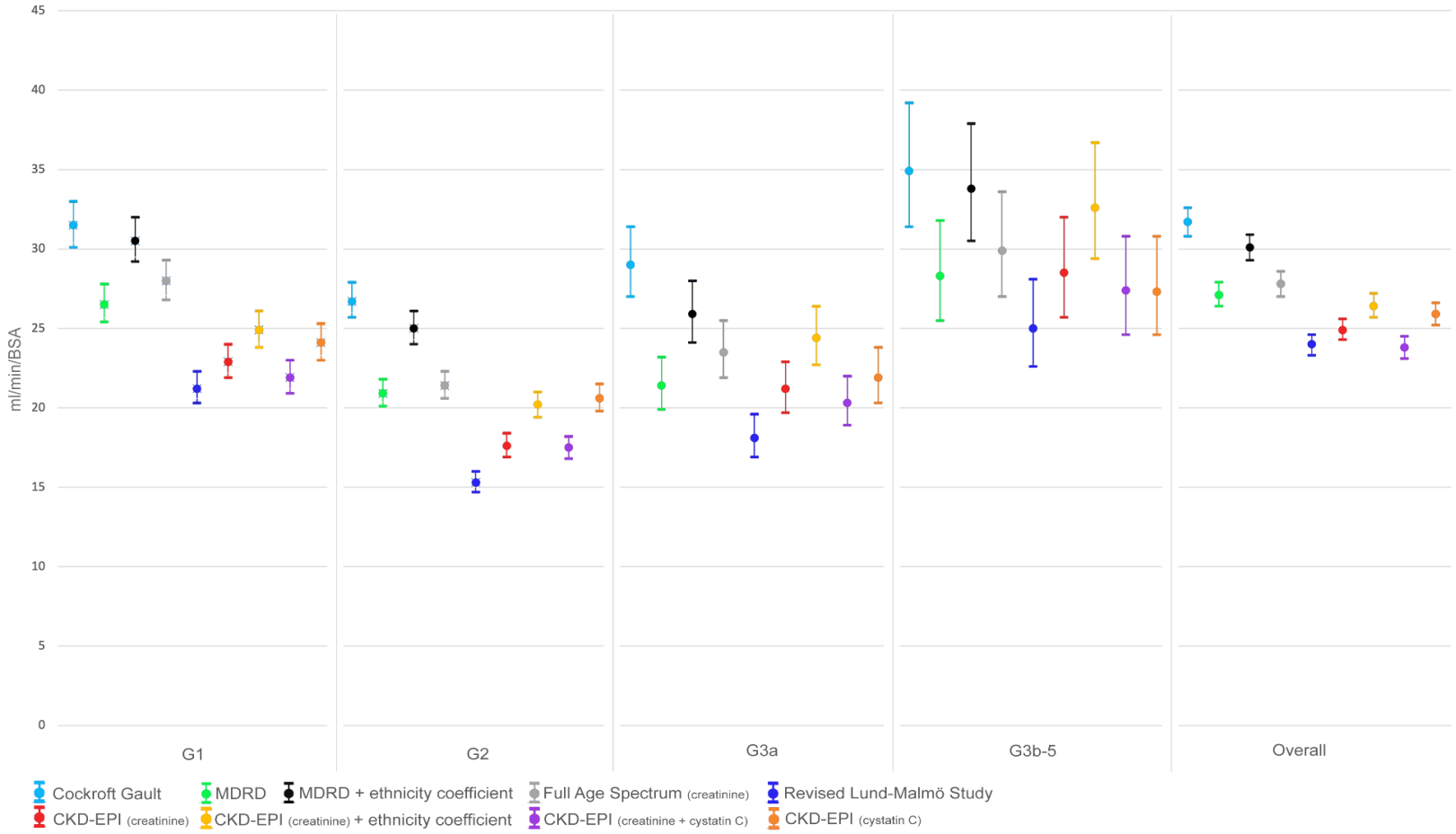
Red dashed lines indicate limits of agreement (+-2 SD); solid blue line indicates bias (ml/min/1.72m^2)

Figure S10 (a): Absolute Bias for each eGFR equation by GFR stage



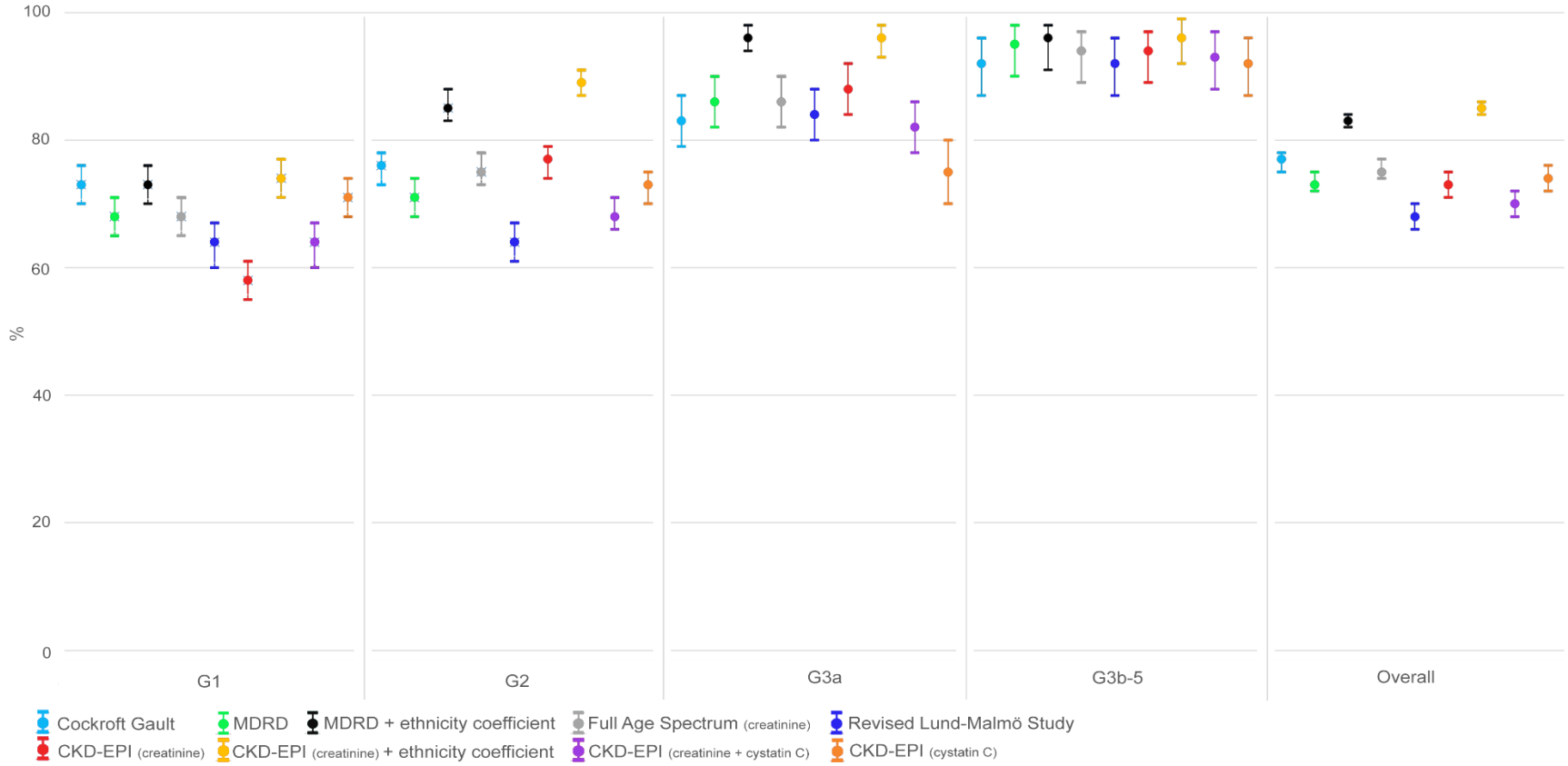
*median difference (eGFR - mGFR) (95% CI)

Figure S10 (b): Precision for each eGFR equation by GFR stage



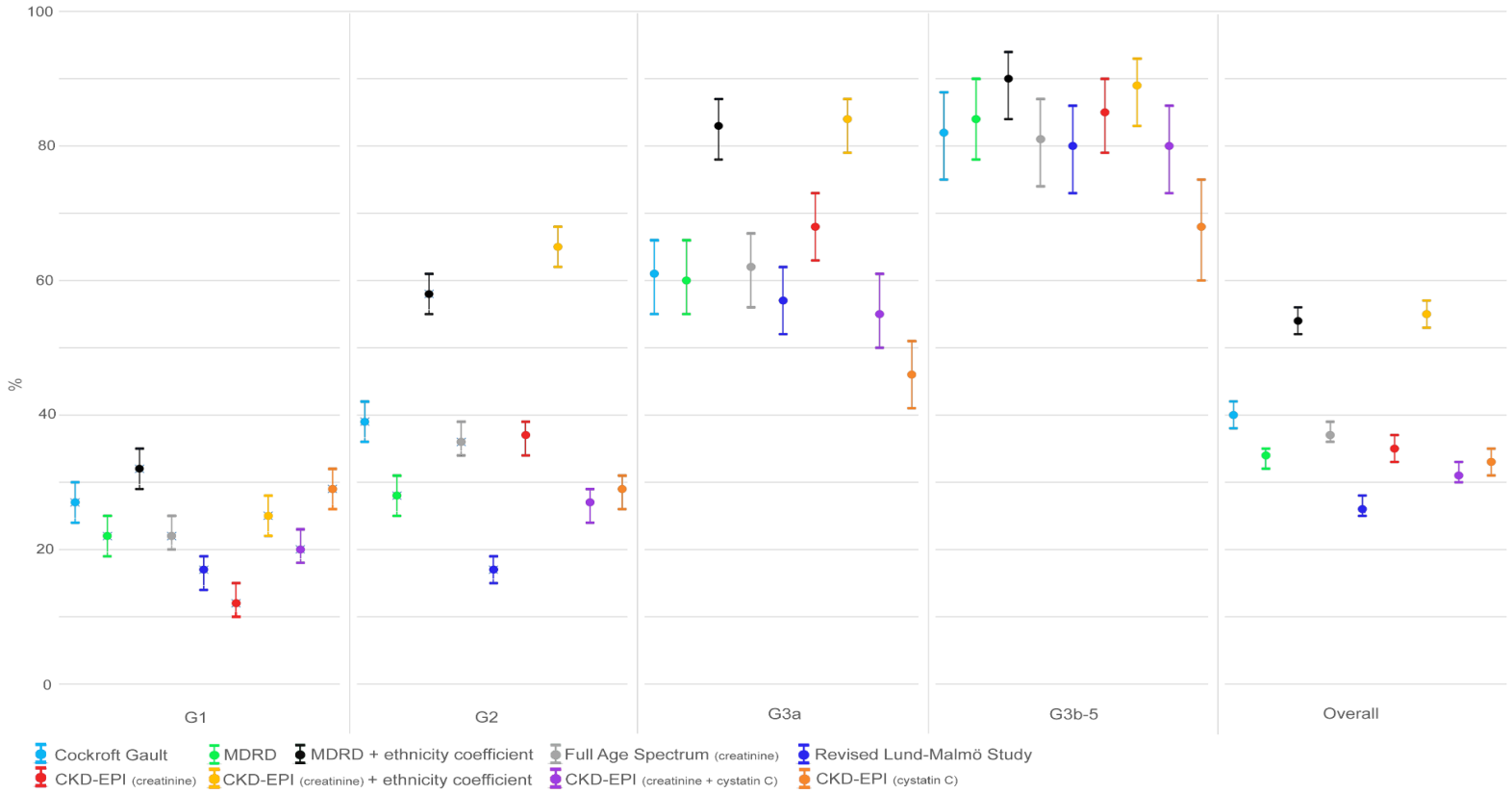
*standard deviation (eGFR - mGFR) (95% CI)

Figure S10 (c): Accuracy (1 - P_{10})* (95% CI) for each eGFR equation by GFR stage



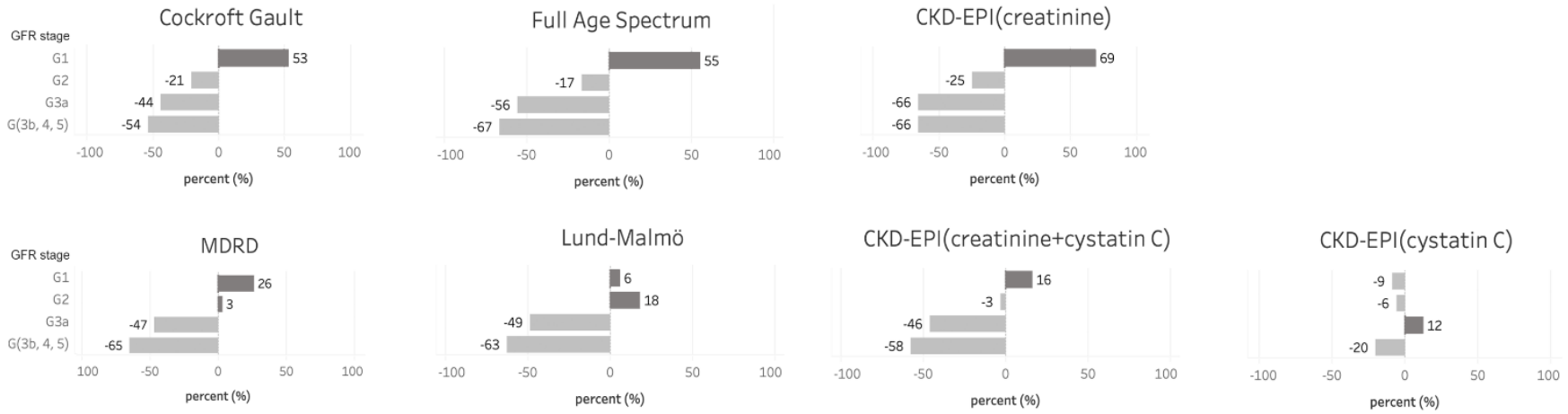
*median (eGFR - mGFR) as the percent of mGFR for estimates that differed by more than 10% (1-P10) of mGFR

Figure S10 (d): Accuracy (1 - P_{30})* (95% CI) for each eGFR equation by GFR stage



*median (eGFR - mGFR) as the percent of mGFR for estimates that differed by more than 30% (1-P30) of mGFR

Figure S11: Schematic using relative bias to depict misclassification of iohexol GFR stage by various eGFR equations



[1]Relative bias (median percentage difference) = median of individual differences between estimated and iohexol GFR, expressed as a percent relative to iohexol GFR median [eGFR - iohexol GFR]/[iohexol GFR]%; Dark grey horizontal bands depict proportional addition of participants to GFR stage, light grey bands depict the proportional subtraction of participants to CKD for each eGFR equation compared to iohexol GFR

Table S6: Overall and by iohexol GFR category: performance of creatinine and cystatin C-based eGFR equations

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	730 (80.3)	456 (54.4)	561 (66.9)	634 (54.3)	294 (26.2)	394 (35.1)	117 (34.4)	58 (18.1)	71 (22.2)	53 (32.9)	18 (11.7)	26 (16.9)	1534 (59.5)	826 (33.9)	1052 (43.2)
eGFR stage 2	168 (18.5)	320 (38.2)	256 (30.5)	474 (40.6)	598 (53.3)	642 (57.3)	172 (50.6)	124 (38.8)	160 (50.0)	60 (37.3)	56 (36.4)	73 (47.4)	874 (33.9)	1098 (45.1)	1131 (46.5)
eGFR stage 3a	6 (0.7)	47 (5.6)	19 (2.3)	53 (4.5)	194 (17.3)	75 (6.7)	35 (10.3)	101 (31.6)	67 (20.9)	22 (13.7)	39 (25.3)	21 (13.6)	116 (4.5)	381 (15.7)	182 (7.5)
eGFR stage 3b,4,5	5 (0.6)	15 (1.8)	2 (0.2)	7 (0.6)	35 (3.1)	10 (0.9)	16 (4.7)	37 (11.6)	22 (6.9)	26 (16.1)	41 (26.6)	34 (22.1)	54 (2.1)	128 (5.3)	68 (2.8)
eGFR all stages	909 [35.3]	838 [34.4]	838 [34.4]	1168 [45.3]	1121 [46.1]	1121 [46.1]	340 [13.2]	320 [13.2]	320 [13.2]	161 [6.2]	154 [6.3]	154 [6.3]	2578 [100]	2433 [100]	2433 [100]
Correctly classified*	730 (80.3)	456 (54.4)	561 (66.9)	474 (40.6)	598 (53.3)	642 (57.3)	35 (10.3)	101 (31.6)	67 (20.9)	26 (16.1)	41 (26.6)	34 (22.1)	1265 (49.1)	1196 (49.2)	1304 (53.6)
Stage as or more severe than mGFR	909 (100)	838 (100)	838 (100)	534 (45.7)	827 (73.8)	727 (64.9)	51 (15.0)	138 (43.1)	89 (27.8)	26 (16.1)	41 (26.6)	34 (22.1)	1520 (59.0)	1844 (75.8)	1688 (69.4)
Absolute bias (mean: eGFR-mGFR)	-5.3	-17	-11	15.2	2.4	8.1	27.3	14.5	19.7	39.3	25.0	31.0	11.1	-1.1	4.7
Absolute bias (median: eGFR-mGFR)	-1.1	-15	-7.6	16.1	0.5	6.9	26.1	9.6	16.5	40.7	20.6	27.2	11.5	-1.7	4.1
Relative bias (mean: eGFR/mGFR)	0.97	0.86	0.92	1.21	1.04	1.11	1.52	1.28	1.37	2.31	1.91	2.08	1.23	1.06	1.14
Relative bias (median: eGFR/mGFR)	0.99	0.86	0.93	1.21	1.01	1.09	1.50	1.17	1.31	2.12	1.55	1.72	1.15	0.98	1.05
Precision, RMSE	22.9	24.1	21.9	17.6	20.6	17.5	21.2	21.9	20.3	28.5	27.3	27.4	24.9	25.9	23.8
Precision, IQR	-17;10.2	-31; 0.1	-23; 4.4	3.3;27.4	-12;15.7	-4.5;19.9	11.1;42.4	-0.8;28.1	4.6;33.4	18.2;58.7	5.0;40.2	10.8;50.5	-3.0;25.6	-17;14.0	-8.8;18.6
P10	0.42	0.31	0.39	0.23	0.28	0.33	0.12	0.26	0.19	0.06	0.08	0.07	0.27	0.28	0.32
P30	0.88	0.77	0.86	0.63	0.74	0.76	0.32	0.58	0.48	0.15	0.34	0.21	0.65	0.70	0.73
mGFR overestimated by 2+ ml/min	396 (43.6)	188 (22.4)	238 (28.4)	900 (77.1)	531 (47.4)	690 (61.6)	304 (89.4)	212 (66.3)	256 (80.0)	143 (88.8)	123 (79.9)	128 (83.1)	1743 (67.6)	1054 (43.3)	1312 (53.9)
eGFR within +/- 2 ml/min of mGFR	73 (8.0)	51 (6.1)	69 (8.2)	71 (6.1)	78 (7.0)	95 (8.5)	10 (2.9)	42 (13.1)	21 (6.6)	5 (3.1)	7 (4.5)	8 (5.2)	159 (6.2)	178 (7.3)	193 (7.9)
mGFR underestimated by 2+ ml/min	440 (48.4)	599 (71.5)	531 (63.4)	197 (16.9)	512 (45.7)	336 (30.0)	26 (7.6)	66 (20.6)	43 (13.4)	13 (8.1)	24 (15.6)	18 (11.7)	676 (26.2)	1201 (49.4)	928 (38.1)
mGFR overestimated by 5+ ml/min	334 (36.7)	155 (18.5)	196 (23.4)	844 (72.3)	458 (40.9)	615 (54.9)	286 (84.1)	191 (59.7)	236 (73.8)	140 (87.0)	116 (75.3)	125 (81.2)	1604 (62.2)	920 (37.8)	1172 (48.2)
eGFR within +/- 5 ml/min of mGFR	179 (19.7)	128 (15.3)	169 (20.2)	173 (14.8)	207 (18.5)	238 (21.2)	38 (11.2)	78 (24.4)	58 (18.1)	11 (6.8)	17 (11.0)	14 (9.1)	401 (15.6)	430 (17.7)	479 (19.7)
mGFR underestimated by 5+ ml/min	396 (43.6)	555 (66.2)	473 (56.4)	151 (12.9)	456 (40.7)	268 (23.9)	16 (4.7)	51 (15.9)	26 (8.1)	10 (6.2)	21 (13.6)	15 (9.7)	573 (22.2)	1083 (44.5)	782 (32.1)
mGFR overestimated by 10+ ml/min	228 (25.1)	104 (12.4)	133 (15.9)	736 (63.0)	365 (32.6)	488 (43.5)	262 (77.1)	157 (49.1)	203 (63.4)	133 (82.6)	96 (62.3)	118 (76.6)	1359 (52.7)	722 (29.7)	942 (38.7)
eGFR within +/- 10 ml/min of mGFR	366 (40.3)	248 (29.6)	327 (39.0)	340 (29.1)	435 (38.8)	477 (42.6)	69 (20.3)	135 (42.2)	104 (32.5)	23 (14.3)	45 (29.2)	28 (18.2)	798 (31.0)	863 (35.5)	936 (38.5)
mGFR underestimated by 10+ ml/min	315 (34.7)	486 (58.0)	378 (45.1)	92 (7.9)	321 (28.6)	156 (13.9)	9 (2.6)	28 (8.8)	13 (4.1)	5 (3.1)	13 (8.4)	8 (5.2)	421 (16.3)	848 (34.9)	555 (22.8)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S6 (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	686 (75.5)	543 (59.7)	667 (73.4)	567 (48.5)	325 (27.8)	562 (48.1)	106 (31.2)	66 (19.4)	106 (31.2)	46 (28.6)	28 (17.4)	51 (31.7)	1405 (54.5)	962 (37.3)	1386 (53.8)
eGFR stage 2	205 (22.6)	346 (38.1)	217 (23.9)	535 (45.8)	761 (65.2)	506 (43.3)	169 (49.7)	197 (57.9)	148 (43.5)	62 (38.5)	78 (48.4)	57 (35.4)	971 (37.7)	1382 (53.6)	928 (36.0)
eGFR stage 3a	14 (1.5)	14 (1.5)	20 (2.2)	58 (5.0)	74 (6.3)	90 (7.7)	50 (14.7)	59 (17.4)	59 (17.4)	27 (16.8)	27 (16.8)	21 (13.0)	149 (5.8)	174 (6.7)	190 (7.4)
eGFR stage 3b,4,5	4 (0.4)	6 (0.7)	5 (0.6)	8 (0.7)	8 (0.7)	10 (0.9)	15 (4.4)	18 (5.3)	27 (7.9)	26 (16.1)	28 (17.4)	32 (19.9)	53 (2.1)	60 (2.3)	74 (2.9)
eGFR all stages	909 [35.3]	909 [35.3]	909 [35.3]	1168 [45.3]	1168 [45.3]	1168 [45.3]	340 [13.2]	340 [13.2]	340 [13.2]	161 [6.2]	161 [6.2]	161 [6.2]	2578 [100]	2578 [100]	2578 [100]
Correctly classified*	686 (75.5)	543 (59.7)	667 (73.4)	535 (45.8)	761 (65.2)	506 (43.3)	50 (14.7)	59 (17.4)	59 (17.4)	26 (16.1)	28 (17.4)	32 (19.9)	1297 (50.3)	1391 (54.0)	1264 (49.0)
Stage as or more severe than mGFR	909 (100)	909 (100)	909 (100)	601 (51.5)	843 (72.2)	606 (51.9)	65 (19.1)	77 (22.6)	86 (25.3)	26 (16.1)	28 (17.4)	32 (19.9)	1601 (62.1)	1857 (72.0)	1633 (63.3)
Absolute bias (mean: eGFR-mGFR)	-3.3	-16	-0.4	15.7	6.4	17.1	26.4	19.9	26.7	38.5	32.3	40.6	11.8	1.8	13.6
Absolute bias (median: eGFR-mGFR)	-0.9	-13	-1.6	14.1	6.5	13.5	23.2	19.3	21.9	36.0	33.8	36.5	11.2	2.3	11.3
Relative bias (mean: eGFR/mGFR)	0.99	0.86	1.01	1.22	1.09	1.23	1.50	1.38	1.51	2.29	2.09	2.36	1.24	1.11	1.26
Relative bias (median: eGFR/mGFR)	0.99	0.88	0.99	1.19	1.08	1.18	1.43	1.36	1.40	2.00	1.94	1.97	1.14	1.03	1.15
Precision, RMSE	28.0	21.2	31.5	21.4	15.3	26.7	23.5	18.1	29.0	29.9	25.0	34.9	27.8	24.0	31.7
Precision, IQR	-20;15.6	-27;-2.2	-20;20.6	1.6;27.7	-3.7;16.1	-1.4;32.7	8.7;42.8	7.2;32.7	4.7;46.7	14.9;59.6	14.2;50.0	13.9;64.9	-3.8;27.6	-11;15.7	-5.7;32.0
P10	0.32	0.36	0.27	0.25	0.36	0.24	0.14	0.16	0.17	0.06	0.07	0.08	0.25	0.32	0.23
P30	0.78	0.83	0.73	0.64	0.83	0.61	0.38	0.43	0.39	0.19	0.20	0.18	0.62	0.74	0.60
mGFR overestimated by 2+ ml/min	411 (45.2)	168 (18.5)	405 (44.6)	868 (74.3)	717 (61.4)	810 (69.3)	293 (86.2)	282 (82.9)	274 (80.6)	144 (89.4)	139 (86.3)	138 (85.7)	1716 (66.6)	1306 (50.7)	1627 (63.1)
eGFR within +/- 2 ml/min of mGFR	56 (6.2)	55 (6.1)	58 (6.4)	73 (6.3)	114 (9.8)	82 (7.0)	22 (6.5)	21 (6.2)	12 (3.5)	8 (5.0)	7 (4.3)	10 (6.2)	159 (6.2)	197 (7.6)	162 (6.3)
mGFR underestimated by 2+ ml/min	442 (48.6)	686 (75.5)	446 (49.1)	227 (19.4)	337 (28.9)	276 (23.6)	25 (7.4)	37 (10.9)	54 (15.9)	9 (5.6)	15 (9.3)	13 (8.1)	703 (27.3)	1075 (41.7)	789 (30.6)
mGFR overestimated by 5+ ml/min	369 (40.6)	120 (13.2)	376 (41.4)	803 (68.8)	641 (54.9)	746 (63.9)	278 (81.8)	270 (79.4)	252 (74.1)	140 (87.0)	135 (83.9)	136 (84.5)	1590 (61.7)	1166 (45.2)	1510 (58.6)
eGFR within +/- 5 ml/min of mGFR	142 (15.6)	162 (17.8)	128 (14.1)	189 (16.2)	266 (22.8)	199 (17.0)	44 (12.9)	47 (13.8)	51 (15.0)	14 (8.7)	14 (8.7)	14 (8.7)	389 (15.1)	489 (19.0)	392 (15.2)
mGFR underestimated by 5+ ml/min	398 (43.8)	627 (69.0)	405 (44.6)	176 (15.1)	261 (22.3)	223 (19.1)	18 (5.3)	23 (6.8)	37 (10.9)	7 (4.3)	12 (7.5)	11 (6.8)	599 (23.2)	923 (35.8)	676 (26.2)
mGFR overestimated by 10+ ml/min	294 (32.3)	56 (6.2)	330 (36.3)	673 (57.6)	473 (40.5)	657 (56.3)	245 (72.1)	231 (67.9)	229 (67.4)	130 (80.7)	129 (80.1)	126 (78.3)	1342 (52.1)	889 (34.5)	1342 (52.1)
eGFR within +/- 10 ml/min of mGFR	281 (30.9)	321 (35.3)	237 (26.1)	387 (33.1)	536 (45.9)	359 (30.7)	87 (25.6)	98 (28.8)	91 (26.8)	26 (16.1)	25 (15.5)	26 (16.1)	781 (30.3)	980 (38.0)	713 (27.7)
mGFR underestimated by 10+ ml/min	334 (36.7)	532 (58.5)	342 (37.6)	108 (9.2)	159 (13.6)	152 (13.0)	8 (2.4)	11 (3.2)	20 (5.9)	5 (3.1)	7 (4.3)	9 (5.6)	455 (17.6)	709 (27.5)	523 (20.3)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S6 (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	829 (91.2)	595 (65.5)	794 (87.3)	896 (76.7)	427 (36.6)	831 (71.1)	186 (54.7)	82 (24.1)	170 (50.0)	82 (50.9)	36 (22.4)	76 (47.2)	1993 (77.3)	1140 (44.2)	1871 (72.6)
eGFR stage 2	73 (8.0)	289 (31.8)	106 (11.7)	256 (21.9)	647 (55.4)	316 (27.1)	122 (35.9)	190 (55.9)	140 (41.2)	42 (26.1)	74 (46.0)	54 (33.5)	493 (19.1)	1200 (46.5)	616 (23.9)
eGFR stage 3a	4 (0.4)	19 (2.1)	5 (0.6)	11 (0.9)	87 (7.4)	16 (1.4)	23 (6.8)	50 (14.7)	24 (7.1)	19 (11.8)	25 (15.5)	15 (9.3)	57 (2.2)	181 (7.0)	60 (2.3)
eGFR stage 3b,4,5	3 (0.3)	6 (0.7)	4 (0.4)	5 (0.4)	7 (0.6)	5 (0.4)	9 (2.6)	18 (5.3)	6 (1.8)	18 (11.2)	26 (16.1)	16 (9.9)	35 (1.4)	57 (2.2)	31 (1.2)
eGFR all stages	909 [35.3]	909 [35.3]	909 [35.3]	1168 [45.3]	1168 [45.3]	1168 [45.3]	340 [13.2]	340 [13.2]	340 [13.2]	161 [6.2]	161 [6.2]	161 [6.2]	2578 [100]	2578 [100]	2578 [100]
Correctly classified*	829 (91.2)	595 (65.5)	794 (87.3)	256 (21.9)	647 (55.4)	316 (27.1)	23 (6.8)	50 (14.7)	24 (7.1)	18 (11.2)	26 (16.1)	16 (9.9)	1126 (43.7)	1318 (51.1)	1150 (44.6)
Stage as or more severe than mGFR	909 (100)	909 (100)	909 (100)	272 (23.3)	741 (63.4)	337 (28.9)	32 (9.4)	68 (20.0)	30 (8.8)	18 (11.2)	26 (16.1)	16 (9.9)	1231 (47.8)	1744 (67.6)	1292 (50.1)
Absolute bias (mean: eGFR-mGFR)	11.2	-8.9	12.3	29.6	10.5	28.7	40.1	23.5	39.8	51.2	35.8	50.9	25.8	6.9	25.8
Absolute bias (median: eGFR-mGFR)	15.2	-8.1	12.2	30.9	8.6	26.2	39.0	20.3	35.8	53.9	34.5	48.8	26.7	6.6	24.7
Relative bias (mean: eGFR/mGFR)	1.12	0.93	1.13	1.40	1.15	1.39	1.76	1.45	1.75	2.67	2.21	2.68	1.43	1.18	1.43
Relative bias (median: eGFR/mGFR)	1.15	0.92	1.12	1.41	1.12	1.35	1.74	1.39	1.68	2.46	1.99	2.41	1.33	1.08	1.31
Precision, RMSE	24.9	26.5	30.5	20.2	20.9	25.0	24.4	21.4	25.9	32.6	28.3	33.8	26.4	27.1	30.1
Precision, IQR	-2.7;27.8	-24; 7.7	-6.2;31.6	16.1;43.3	-3.5;22.4	12.0;42.6	21.8;58.3	7.9;36.1	20.8;54.5	26.5;73.4	16.0;51.9	26.4;70.6	10.4;41.6	-8.6;22.5	7.4;43.0
P10	0.26	0.32	0.27	0.11	0.29	0.14	0.04	0.14	0.04	0.04	0.05	0.04	0.15	0.27	0.17
P30	0.75	0.78	0.68	0.35	0.72	0.42	0.16	0.40	0.17	0.11	0.16	0.10	0.45	0.66	0.46
mGFR overestimated by 2+ ml/min	629 (69.2)	298 (32.8)	592 (65.1)	1064 (91.1)	751 (64.3)	1030 (88.2)	325 (95.6)	292 (85.9)	326 (95.9)	150 (93.2)	143 (88.8)	151 (93.8)	2168 (84.1)	1484 (57.6)	2099 (81.4)
eGFR within +/- 2 ml/min of mGFR	49 (5.4)	60 (6.6)	45 (5.0)	34 (2.9)	88 (7.5)	39 (3.3)	6 (1.8)	19 (5.6)	4 (1.2)	3 (1.9)	5 (3.1)	5 (3.1)	92 (3.6)	172 (6.7)	93 (3.6)
mGFR underestimated by 2+ ml/min	231 (25.4)	551 (60.6)	272 (29.9)	70 (6.0)	329 (28.2)	99 (8.5)	9 (2.6)	29 (8.5)	10 (2.9)	8 (5.0)	13 (8.1)	5 (3.1)	318 (12.3)	922 (35.8)	386 (15.0)
mGFR overestimated by 5+ ml/min	598 (65.8)	263 (28.9)	554 (60.9)	1034 (88.5)	677 (58.0)	994 (85.1)	318 (93.5)	275 (80.9)	318 (93.5)	146 (90.7)	141 (87.6)	148 (91.9)	2096 (81.3)	1356 (52.6)	2014 (78.1)
eGFR within +/- 5 ml/min of mGFR	115 (12.7)	139 (15.3)	116 (12.8)	80 (6.8)	229 (19.6)	105 (9.0)	13 (3.8)	43 (12.6)	14 (4.1)	10 (6.2)	10 (6.2)	9 (5.6)	218 (8.5)	421 (16.3)	244 (9.5)
mGFR underestimated by 5+ ml/min	196 (21.6)	507 (55.8)	239 (26.3)	54 (4.6)	262 (22.4)	69 (5.9)	9 (2.6)	22 (6.5)	8 (2.4)	5 (3.1)	10 (6.2)	4 (2.5)	264 (10.2)	801 (31.1)	320 (12.4)
mGFR overestimated by 10+ ml/min	528 (58.1)	210 (23.1)	491 (54.0)	969 (83.0)	545 (46.7)	915 (78.3)	303 (89.1)	241 (70.9)	306 (90.0)	142 (88.2)	130 (80.7)	142 (88.2)	1942 (75.3)	1126 (43.7)	1854 (71.9)
eGFR within +/- 10 ml/min of mGFR	217 (23.9)	276 (30.4)	233 (25.6)	167 (14.3)	463 (39.6)	216 (18.5)	32 (9.4)	88 (25.9)	30 (8.8)	16 (9.9)	27 (16.8)	17 (10.6)	432 (16.8)	854 (33.1)	496 (19.2)
mGFR underestimated by 10+ ml/min	164 (18.0)	423 (46.5)	185 (20.4)	32 (2.7)	160 (13.7)	37 (3.2)	5 (1.5)	11 (3.2)	4 (1.2)	3 (1.9)	4 (2.5)	2 (1.2)	204 (7.9)	598 (23.2)	228 (8.8)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

Table S7: Overall – GFR stage comparing eGFR equations to iohexol GFR

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	909 (**35**)	1386(**54**)	1140 (**44**)	1871 (**73**)	1405 (**55**)	962 (**37**)	1534 (**60**)	1993 (**77**)	1052 (**43**)	826 (**34**)
G2	1168 (**45**)	928 (**36**)	1200 (**47**)	616 (**24**)	971 (**38**)	1382 (**54**)	874 (**34**)	493 (**19**)	1131(**47**)	1098 (**45**)
G3a	340 (**13**)	190 (**7**)	181 (**7**)	60 (**2**)	149 (**6**)	174 (**7**)	116 (**5**)	57 (**2**)	182 (**8**)	381(**16**)
G3b, G4, G5	161 (**6**)	74 (**3**)	57 (**2**)	31(**1**)	53 (**2**)	60 (**2**)	54 (**2**)	35 (**1**)	68 (**3**)	128 (**5**)
% classified correctly	reference category	1264 (**49**)	1318 (**51**)	1150 (**45**)	1297 (**50**)	1391 (**54**)	1265 (**49**)	1126 (**44**)	1304 (**54**)	1196 (**49**)
% classified as or more severe than mGFR stage	reference category	1633 (**63**)	1744 (**68**)	1292 (**50**)	1601(**62**)	1857 (**72**)	1520 (**59**)	1231 (**48**)	1688 (**69**)	1844 (**76**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15
[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;
[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;
[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.
N=2578 for creatinine-based equations; N=2433 for cystatin C-based equations

Figure S12: Iohexol GFR stage compared to GFR stage estimated by the CKD-EPI (creatinine) equation without ethnicity coefficient, overall and by country

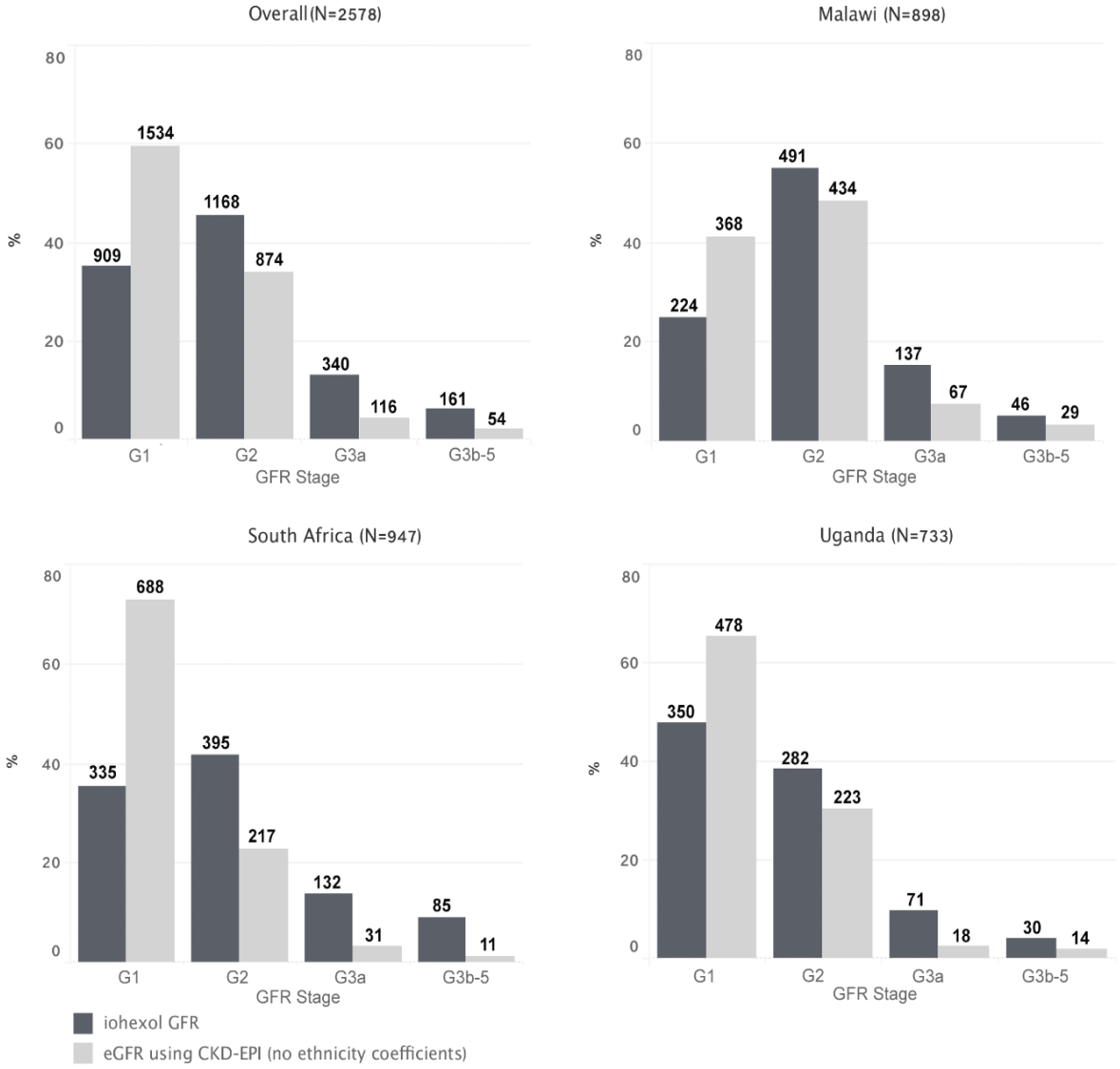


Table S8 (a): Malawi – Performance of eGFR equations compared to iohexol GFR

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[4]P_{10}	[5]P_{30}
Cockroft Gault (adjusted for BSA)	3.8	1.05	23.5	0.31	0.73
MDRD	1.8	1.02	22.3	0.30	0.75
MDRD ethnicity coefficient	18.0	1.24	25.2	0.21	0.56
FAS (creatinine)	10.2	1.14	22.0	0.28	0.70
Lund-Malmö (revised)	-0.5	0.99	18.4	0.36	0.81
CKD-EPI (creatinine)	7.6	1.09	19.4	0.32	0.74
CKD-EPI (creatinine) ethnicity coefficient	21.3	1.27	21.0	0.18	0.52
CKD-EPI (creatinine + cystatin C)	1.4	1.02	18.1	0.36	0.81
CKD-EPI (cystatin C)	-4.3	0.94	20.8	0.29	0.76

[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3] Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=898 for creatinine-based equations; N=871 for cystatin C-based equations

Table S8 (b): Malawi – GFR stage comparing eGFR equations to iohexol GFR

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	224 (**25**)	311 (**35**)	243 (**27**)	511 (**57**)	403 (**45**)	177 (**20**)	368 (**41**)	572 (**64**)	245 (**28**)	196 (**23**)
G2	491 (**55**)	438 (**49**)	518 (**58**)	336(**37**)	405 (**45**)	588 (**66**)	434 (**48**)	280 (**31**)	493 (**57**)	430 (**49**)
G3a	137 (**15**)	108 (**12**)	107 (**12**)	37 (**4**)	68 (**8**)	101 (**11**)	67 (**8**)	29 (**3**)	96 (**11**)	178 (**20**)
G3b, G4, G5	46 (**5**)	41 (**5**)	30 (**3**)	14 (**2**)	22 (**2**)	32 (**4**)	29 (**3**)	17 (**2**)	37 (**4**)	67 (**8**)
% classified correctly	reference category	476 (**53**)	478 (**53**)	398 (**44**)	467 (**52**)	517 (**58**)	455 (**51**)	386 (**43**)	498 (**57**)	434 (**50**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15
[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;
[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;
[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.
N=898 for creatinine-based equations; N=871 for cystatin C-based equations

Table S18 (c): Malawi – overall and by iohexol GFR category: performance of creatinine and cystatin C-based eGFR equations

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	154 (68.8)	93 (43.1)	122 (56.5)	190 (38.7)	92 (19.2)	111 (23.2)	19 (13.9)	11 (8.4)	11 (8.4)	5 (10.9)	0 (0.0)	1 (2.2)	368 (41.0)	196 (22.5)	245 (28.1)
eGFR stage 2	66 (29.5)	103 (47.7)	89 (41.2)	264 (53.8)	269 (56.2)	319 (66.6)	87 (63.5)	50 (38.2)	70 (53.4)	17 (37.0)	8 (17.8)	15 (33.3)	434 (48.3)	430 (49.4)	493 (56.6)
eGFR stage 3a	3 (1.3)	17 (7.9)	5 (2.3)	33 (6.7)	98 (20.5)	43 (9.0)	22 (16.1)	49 (37.4)	38 (29.0)	9 (19.6)	14 (31.1)	10 (22.2)	67 (7.5)	178 (20.4)	96 (11.0)
eGFR stage 3b,4,5	1 (0.4)	3 (1.4)	0 (0.0)	4 (0.8)	20 (4.2)	6 (1.3)	9 (6.6)	21 (16.0)	12 (9.2)	15 (32.6)	23 (51.1)	19 (42.2)	29 (3.2)	67 (7.7)	37 (4.2)
eGFR all stages	224 [24.9]	216 [24.8]	216 [24.8]	491 [54.7]	479 [55.0]	479 [55.0]	137 [15.3]	131 [15.0]	131 [15.0]	46 [5.1]	45 [5.2]	45 [5.2]	898 [100]	871 [100]	871 [100]
Correctly classified*	154 (68.8)	93 (43.1)	122 (56.5)	264 (53.8)	269 (56.2)	319 (66.6)	22 (16.1)	49 (37.4)	38 (29.0)	15 (32.6)	23 (51.1)	19 (42.2)	455 (50.7)	434 (49.8)	498 (57.2)
Stage as or more severe than mGFR	224 (100)	216 (100)	216 (100)	301 (61.3)	387 (80.8)	368 (76.8)	31 (22.6)	70 (53.4)	50 (38.2)	15 (32.6)	23 (51.1)	19 (42.2)	571 (63.6)	696 (79.9)	653 (75.0)
Absolute bias (mean: eGFR-mGFR)	-5.6	-16	-12	9.2	-1.1	3.3	17.9	8.1	12.0	20.8	7.1	12.2	7.4	-2.9	1.4
Absolute bias (median: eGFR-mGFR)	-3.8	-17	-11	9.7	-3.7	2.3	15.7	4.0	10.0	17.6	5.8	11.2	7.6	-4.3	1.4
Relative bias (mean: eGFR/mGFR)	0.95	0.85	0.89	1.13	0.99	1.05	1.34	1.15	1.23	1.60	1.21	1.34	1.14	0.99	1.05
Relative bias (median: eGFR/mGFR)	0.96	0.83	0.89	1.13	0.95	1.03	1.30	1.07	1.19	1.42	1.15	1.32	1.09	0.94	1.02
Precision, RMSE	19.0	20.6	17.5	16.1	19.0	15.5	16.5	17.7	15.2	25.2	17.4	19.8	19.4	20.8	18.1
Precision, IQR	-17; 8.1	-29;-1.7	-23;-0.1	-2.0;20.3	-14; 9.6	-7.4;13.0	5.4;29.6	-2.8;17.2	1.9;22.1	1.0;31.9	-7.5;13.3	-4.0;19.7	-4.2;19.9	-16; 9.3	-9.4;12.5
P10	0.43	0.27	0.39	0.32	0.31	0.40	0.20	0.33	0.24	0.11	0.11	0.11	0.32	0.29	0.36
P30	0.92	0.81	0.90	0.77	0.77	0.85	0.50	0.69	0.63	0.33	0.58	0.42	0.74	0.76	0.81
mGFR overestimated by 2+ ml/min	93 (41.5)	44 (20.4)	47 (21.8)	321 (65.4)	189 (39.5)	246 (51.4)	115 (83.9)	73 (55.7)	98 (74.8)	34 (73.9)	28 (62.2)	29 (64.4)	563 (62.7)	334 (38.3)	420 (48.2)
eGFR within +/- 2 ml/min of mGFR	12 (5.4)	13 (6.0)	13 (6.0)	48 (9.8)	35 (7.3)	45 (9.4)	8 (5.8)	23 (17.6)	11 (8.4)	2 (4.3)	3 (6.7)	4 (8.9)	70 (7.8)	74 (8.5)	73 (8.4)
mGFR underestimated by 2+ ml/min	119 (53.1)	159 (73.6)	156 (72.2)	122 (24.8)	255 (53.2)	188 (39.2)	14 (10.2)	35 (26.7)	22 (16.8)	10 (21.7)	14 (31.1)	12 (26.7)	265 (29.5)	463 (53.2)	378 (43.4)
mGFR overestimated by 5+ ml/min	75 (33.5)	37 (17.1)	37 (17.1)	290 (59.1)	151 (31.5)	203 (42.4)	103 (75.2)	62 (47.3)	86 (65.6)	33 (71.7)	24 (53.3)	27 (60.0)	501 (55.8)	274 (31.5)	353 (40.5)
eGFR within +/- 5 ml/min of mGFR	43 (19.2)	26 (12.0)	42 (19.4)	110 (22.4)	95 (19.8)	124 (25.9)	25 (18.2)	41 (31.3)	30 (22.9)	5 (10.9)	8 (17.8)	7 (15.6)	183 (20.4)	170 (19.5)	203 (23.3)
mGFR underestimated by 5+ ml/min	106 (47.3)	153 (70.8)	137 (63.4)	91 (18.5)	233 (48.6)	152 (31.7)	9 (6.6)	28 (21.4)	15 (11.5)	8 (17.4)	13 (28.9)	11 (24.4)	214 (23.8)	427 (49.0)	315 (36.2)
mGFR overestimated by 10+ ml/min	45 (20.1)	26 (12.0)	22 (10.2)	243 (49.5)	115 (24.0)	147 (30.7)	88 (64.2)	50 (38.2)	66 (50.4)	29 (63.0)	14 (31.1)	24 (53.3)	405 (45.1)	205 (23.5)	259 (29.7)
eGFR within +/- 10 ml/min of mGFR	93 (41.5)	55 (25.5)	82 (38.0)	193 (39.3)	202 (42.2)	245 (51.1)	44 (32.1)	68 (51.9)	58 (44.3)	14 (30.4)	22 (48.9)	16 (35.6)	344 (38.3)	347 (39.8)	401 (46.0)
mGFR underestimated by 10+ ml/min	86 (38.4)	135 (62.5)	112 (51.9)	55 (11.2)	162 (33.8)	87 (18.2)	5 (3.6)	13 (9.9)	7 (5.3)	3 (6.5)	9 (20.0)	5 (11.1)	149 (16.6)	319 (36.6)	211 (24.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S8 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	166 (74.1)	97 (43.3)	139 (62.1)	207 (42.2)	72 (14.7)	152 (31.0)	25 (18.2)	4 (2.9)	14 (10.2)	5 (10.9)	4 (8.7)	6 (13.0)	403 (44.9)	177 (19.7)	311 (34.6)
eGFR stage 2	55 (24.6)	121 (54.0)	78 (34.8)	258 (52.5)	367 (74.7)	278 (56.6)	77 (56.2)	85 (62.0)	70 (51.1)	15 (32.6)	15 (32.6)	12 (26.1)	405 (45.1)	588 (65.5)	438 (48.8)
eGFR stage 3a	3 (1.3)	5 (2.2)	6 (2.7)	24 (4.9)	48 (9.8)	57 (11.6)	29 (21.2)	37 (27.0)	38 (27.7)	12 (26.1)	11 (23.9)	7 (15.2)	68 (7.6)	101 (11.2)	108 (12.0)
eGFR stage 3b,4,5	0 (0.0)	1 (0.4)	1 (0.4)	2 (0.4)	4 (0.8)	4 (0.8)	6 (4.4)	11 (8.0)	15 (10.9)	14 (30.4)	16 (34.8)	21 (45.7)	22 (2.4)	32 (3.6)	41 (4.6)
eGFR all stages	224 [24.9]	224 [24.9]	224 [24.9]	491 [54.7]	491 [54.7]	491 [54.7]	137 [15.3]	137 [15.3]	137 [15.3]	46 [5.1]	46 [5.1]	46 [5.1]	898 [100]	898 [100]	898 [100]
Correctly classified*	166 (74.1)	97 (43.3)	139 (62.1)	258 (52.5)	367 (74.7)	278 (56.6)	29 (21.2)	37 (27.0)	38 (27.7)	14 (30.4)	16 (34.8)	21 (45.7)	467 (52.0)	517 (57.6)	476 (53.0)
Stage as or more severe than mGFR	224 (100)	224 (100)	224 (100)	284 (57.8)	419 (85.3)	339 (69.0)	35 (25.5)	48 (35.0)	53 (38.7)	14 (30.4)	16 (34.8)	21 (45.7)	557 (62.0)	707 (78.7)	637 (70.9)
Absolute bias (mean: eGFR-mGFR)	1.3	-16	-5.1	13.2	1.6	7.3	19.4	12.1	13.7	22.6	16.0	18.2	11.6	-0.5	5.8
Absolute bias (median: eGFR-mGFR)	1.1	-15	-5.6	11.4	1.2	5.1	16.4	10.8	10.7	16.3	13.6	13.8	10.2	-0.5	3.8
Relative bias (mean: eGFR/mGFR)	1.02	0.85	0.96	1.18	1.03	1.10	1.37	1.23	1.26	1.65	1.47	1.55	1.19	1.04	1.11
Relative bias (median: eGFR/mGFR)	1.01	0.86	0.94	1.15	1.02	1.07	1.30	1.20	1.19	1.42	1.35	1.33	1.14	0.99	1.05
Precision, RMSE	22.3	16.9	24.2	20.1	14.0	21.3	19.2	14.6	20.6	28.2	23.1	30.2	22.0	18.4	23.5
Precision, IQR	-13;15.4	-26;-4.3	-20; 9.3	-0.2;23.5	-7.2;10.4	-5.9;17.5	5.5;31.0	1.4;22.2	-2.2;25.3	3.9;31.9	-1.1;27.4	-1.1;28.8	-1.6;23.1	-11;10.4	-8.5;17.9
P10	0.36	0.35	0.33	0.29	0.41	0.32	0.19	0.25	0.25	0.11	0.13	0.20	0.28	0.36	0.31
P30	0.86	0.86	0.83	0.71	0.89	0.78	0.50	0.60	0.55	0.41	0.41	0.35	0.70	0.81	0.73
mGFR overestimated by 2+ ml/min	109 (48.7)	27 (12.1)	77 (34.4)	350 (71.3)	235 (47.9)	283 (57.6)	112 (81.8)	100 (73.0)	96 (70.1)	35 (76.1)	31 (67.4)	30 (65.2)	606 (67.5)	393 (43.8)	486 (54.1)
eGFR within +/- 2 ml/min of mGFR	19 (8.5)	17 (7.6)	18 (8.0)	37 (7.5)	54 (11.0)	46 (9.4)	14 (10.2)	15 (10.9)	6 (4.4)	4 (8.7)	4 (8.7)	6 (13.0)	74 (8.2)	90 (10.0)	76 (8.5)
mGFR underestimated by 2+ ml/min	96 (42.9)	180 (80.4)	129 (57.6)	104 (21.2)	202 (41.1)	162 (33.0)	11 (8.0)	22 (16.1)	35 (25.5)	7 (15.2)	11 (23.9)	10 (21.7)	218 (24.3)	415 (46.2)	336 (37.4)
mGFR overestimated by 5+ ml/min	100 (44.6)	17 (7.6)	67 (29.9)	319 (65.0)	197 (40.1)	246 (50.1)	104 (75.9)	93 (67.9)	85 (62.0)	32 (69.6)	30 (65.2)	28 (60.9)	555 (61.8)	337 (37.5)	426 (47.4)
eGFR within +/- 5 ml/min of mGFR	42 (18.8)	43 (19.2)	39 (17.4)	94 (19.1)	137 (27.9)	109 (22.2)	26 (19.0)	31 (22.6)	30 (21.9)	9 (19.6)	6 (13.0)	10 (21.7)	171 (19.0)	217 (24.2)	188 (20.9)
mGFR underestimated by 5+ ml/min	82 (36.6)	164 (73.2)	118 (52.7)	78 (15.9)	157 (32.0)	136 (27.7)	7 (5.1)	13 (9.5)	22 (16.1)	5 (10.9)	10 (21.7)	8 (17.4)	172 (19.2)	344 (38.3)	284 (31.6)
mGFR overestimated by 10+ ml/min	77 (34.4)	7 (3.1)	53 (23.7)	258 (52.5)	130 (26.5)	197 (40.1)	91 (66.4)	74 (54.0)	71 (51.8)	28 (60.9)	27 (58.7)	24 (52.2)	454 (50.6)	238 (26.5)	345 (38.4)
eGFR within +/- 10 ml/min of mGFR	80 (35.7)	77 (34.4)	75 (33.5)	190 (38.7)	270 (55.0)	203 (41.3)	42 (30.7)	57 (41.6)	54 (39.4)	15 (32.6)	14 (30.4)	15 (32.6)	327 (36.4)	418 (46.5)	347 (38.6)
mGFR underestimated by 10+ ml/min	67 (29.9)	140 (62.5)	96 (42.9)	43 (8.8)	91 (18.5)	91 (18.5)	4 (2.9)	6 (4.4)	12 (8.8)	3 (6.5)	5 (10.9)	7 (15.2)	117 (13.0)	242 (26.9)	206 (22.9)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S8 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	195 (87.1)	118 (52.7)	177 (79.0)	317 (64.6)	112 (22.8)	283 (57.6)	50 (36.5)	10 (7.3)	43 (31.4)	10 (21.7)	3 (6.5)	8 (17.4)	572 (63.7)	243 (27.1)	511 (56.9)
eGFR stage 2	27 (12.1)	96 (42.9)	43 (19.2)	167 (34.0)	317 (64.6)	195 (39.7)	70 (51.1)	87 (63.5)	77 (56.2)	16 (34.8)	18 (39.1)	21 (45.7)	280 (31.2)	518 (57.7)	336 (37.4)
eGFR stage 3a	2 (0.9)	9 (4.0)	3 (1.3)	5 (1.0)	58 (11.8)	11 (2.2)	13 (9.5)	29 (21.2)	16 (11.7)	9 (19.6)	11 (23.9)	7 (15.2)	29 (3.2)	107 (11.9)	37 (4.1)
eGFR stage 3b,4,5	0 (0.0)	1 (0.4)	1 (0.4)	2 (0.4)	4 (0.8)	2 (0.4)	4 (2.9)	11 (8.0)	1 (0.7)	11 (23.9)	14 (30.4)	10 (21.7)	17 (1.9)	30 (3.3)	14 (1.6)
eGFR all stages	224 [24.9]	224 [24.9]	224 [24.9]	491 [54.7]	491 [54.7]	491 [54.7]	137 [15.3]	137 [15.3]	137 [15.3]	46 [5.1]	46 [5.1]	46 [5.1]	898 [100]	898 [100]	898 [100]
Correctly classified*	195 (87.1)	118 (52.7)	177 (79.0)	167 (34.0)	317 (64.6)	195 (39.7)	13 (9.5)	29 (21.2)	16 (11.7)	11 (23.9)	14 (30.4)	10 (21.7)	386 (43.0)	478 (53.2)	398 (44.3)
Stage as or more severe than mGFR	224 (100)	224 (100)	224 (100)	174 (35.4)	379 (77.2)	208 (42.4)	17 (12.4)	40 (29.2)	17 (12.4)	11 (23.9)	14 (30.4)	10 (21.7)	426 (47.4)	657 (73.2)	459 (51.1)
Absolute bias (mean: eGFR-mGFR)	9.9	-12	7.2	22.6	4.4	21.2	29.3	14.9	29.4	30.1	18.8	30.9	20.8	2.6	19.4
Absolute bias (median: eGFR-mGFR)	11.7	-12	6.9	23.0	2.2	18.8	26.8	12.6	26.1	26.1	16.2	26.9	21.3	1.8	18.0
Relative bias (mean: eGFR/mGFR)	1.11	0.89	1.08	1.31	1.06	1.29	1.55	1.28	1.56	1.85	1.54	1.87	1.32	1.08	1.31
Relative bias (median: eGFR/mGFR)	1.12	0.88	1.06	1.31	1.03	1.25	1.50	1.22	1.48	1.64	1.42	1.72	1.27	1.02	1.24
Precision, RMSE	21.2	21.9	25.6	18.4	19.7	23.6	19.0	17.7	21.3	29.1	23.6	28.4	21.0	22.3	25.2
Precision, IQR	-2.7;25.5	-26; 1.4	-8.8;22.8	9.8;35.2	-7.9;13.6	6.9;31.2	15.2;42.4	3.0;23.3	15.6;40.0	7.6;44.0	0.3;27.6	7.3;40.8	7.6;34.7	-10;14.0	4.3;31.5
P10	0.31	0.31	0.34	0.16	0.34	0.21	0.06	0.21	0.06	0.11	0.07	0.13	0.18	0.30	0.21
P30	0.82	0.78	0.79	0.48	0.82	0.55	0.28	0.58	0.29	0.26	0.33	0.24	0.52	0.75	0.56
mGFR overestimated by 2+ ml/min	148 (66.1)	52 (23.2)	127 (56.7)	428 (87.2)	252 (51.3)	407 (82.9)	129 (94.2)	107 (78.1)	130 (94.9)	37 (80.4)	34 (73.9)	38 (82.6)	742 (82.6)	445 (49.6)	702 (78.2)
eGFR within +/- 2 ml/min of mGFR	18 (8.0)	15 (6.7)	18 (8.0)	20 (4.1)	39 (7.9)	23 (4.7)	3 (2.2)	13 (9.5)	2 (1.5)	3 (6.5)	2 (4.3)	5 (10.9)	44 (4.9)	69 (7.7)	48 (5.3)
mGFR underestimated by 2+ ml/min	58 (25.9)	157 (70.1)	79 (35.3)	43 (8.8)	200 (40.7)	61 (12.4)	5 (3.6)	17 (12.4)	5 (3.6)	6 (13.0)	10 (21.7)	3 (6.5)	112 (12.5)	384 (42.8)	148 (16.5)
mGFR overestimated by 5+ ml/min	137 (61.2)	40 (17.9)	121 (54.0)	408 (83.1)	219 (44.6)	385 (78.4)	125 (91.2)	95 (69.3)	124 (90.5)	35 (76.1)	34 (73.9)	36 (78.3)	705 (78.5)	388 (43.2)	666 (74.2)
eGFR within +/- 5 ml/min of mGFR	39 (17.4)	38 (17.0)	35 (15.6)	50 (10.2)	110 (22.4)	64 (13.0)	7 (5.1)	28 (20.4)	9 (6.6)	8 (17.4)	4 (8.7)	8 (17.4)	104 (11.6)	180 (20.0)	116 (12.9)
mGFR underestimated by 5+ ml/min	48 (21.4)	146 (65.2)	68 (30.4)	33 (6.7)	162 (33.0)	42 (8.6)	5 (3.6)	14 (10.2)	4 (2.9)	3 (6.5)	8 (17.4)	2 (4.3)	89 (9.9)	330 (36.7)	116 (12.9)
mGFR overestimated by 10+ ml/min	118 (52.7)	32 (14.3)	95 (42.4)	367 (74.7)	158 (32.2)	330 (67.2)	115 (83.9)	77 (56.2)	116 (84.7)	34 (73.9)	27 (58.7)	34 (73.9)	634 (70.6)	294 (32.7)	575 (64.0)
eGFR within +/- 10 ml/min of mGFR	67 (29.9)	68 (30.4)	77 (34.4)	106 (21.6)	236 (48.1)	137 (27.9)	20 (14.6)	54 (39.4)	20 (14.6)	11 (23.9)	17 (37.0)	11 (23.9)	204 (22.7)	375 (41.8)	245 (27.3)
mGFR underestimated by 10+ ml/min	39 (17.4)	124 (55.4)	52 (23.2)	18 (3.7)	97 (19.8)	24 (4.9)	2 (1.5)	6 (4.4)	1 (0.7)	1 (2.2)	2 (4.3)	1 (2.2)	60 (6.7)	229 (25.5)	78 (8.7)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

Table S9 (a): South Africa – Performance of eGFR equations compared to iohexol GFR

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[4]P_{10}	[5]P_{30}
Cockroft Gault (adjusted for BSA)	28.3	1.36	31.0	0.16	0.42
MDRD	13.3	1.17	26.3	0.25	0.60
MDRD ethnicity coefficient	32.9	1.42	28.8	0.10	0.34
FAS (creatinine)	21.1	1.27	26.9	0.20	0.51
Lund-Malmö (revised)	8.2	1.10	23.9	0.31	0.68
CKD-EPI (creatinine)	19.4	1.24	24.7	0.23	0.54
CKD-EPI (creatinine) ethnicity coefficient	34.9	1.44	26.0	0.08	0.33
CKD-EPI (creatinine + cystatin C)	11.0	1.14	25.2	0.27	0.64
CKD-EPI (cystatin C)	4.4	1.06	27.9	0.26	0.64

[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3]Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=947 for creatinine-based equations; N=942 for cystatin C-based equations

Table S9 (b): South Africa – GFR stage comparing eGFR equations to iohexol GFR

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	335 (**35**)	715 (**76**)	520 (**55**)	785 (**83**)	671 (**71**)	479 (**51**)	688 (**73**)	816 (**86**)	519 (**55**)	414 (**44**)
G2	395 (**42**)	184 (**19**)	370 (**39**)	143 (**15**)	225 (**24**)	409 (**43**)	217 (**23**)	108 (**11**)	351 (**37**)	356 (**38**)
G3a	132 (**14**)	37 (**4**)	44 (**5**)	14 (**2**)	42 (**4**)	46 (**5**)	31 (**3**)	18 (**2**)	54 (**6**)	132 (**14**)
G3b, G4, G5	85(**9**)	11 (**1**)	13 (**1**)	5 (**1**)	9 (**1**)	13 (**1**)	11 (**1**)	5 (**1**)	18 (**2**)	40 (**4**)
% classified correctly	reference category	402 (**42**)	450 (**48**)	368 (**39**)	437 (**46**)	468 (**49**)	412 (**44**)	363 (**38**)	457 (**49**)	436 **4** **6**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15
[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;
[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;
[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.
N=947 for creatinine-based equations; N=942 for cystatin C-based equations

Table S9 (c): South Africa – overall and by iohexol GFR category: performance of creatinine and cystatin C-based eGFR equations

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	297 (88.7)	214 (64.3)	250 (75.1)	286 (72.4)	148 (37.7)	200 (50.9)	68 (51.5)	40 (30.5)	49 (37.4)	37 (43.5)	12 (14.1)	20 (23.5)	688 (72.7)	414 (43.9)	519 (55.1)
eGFR stage 2	35 (10.4)	96 (28.8)	74 (22.2)	100 (25.3)	177 (45.0)	176 (44.8)	53 (40.2)	47 (35.9)	57 (43.5)	29 (34.1)	36 (42.4)	44 (51.8)	217 (22.9)	356 (37.8)	351 (37.3)
eGFR stage 3a	2 (0.6)	16 (4.8)	8 (2.4)	9 (2.3)	60 (15.3)	16 (4.1)	8 (6.1)	32 (24.4)	20 (15.3)	12 (14.1)	24 (28.2)	10 (11.8)	31 (3.3)	132 (14.0)	54 (5.7)
eGFR stage 3b,4,5	1 (0.3)	7 (2.1)	1 (0.3)	0 (0.0)	8 (2.0)	1 (0.3)	3 (2.3)	12 (9.2)	5 (3.8)	7 (8.2)	13 (15.3)	11 (12.9)	11 (1.2)	40 (4.2)	18 (1.9)
eGFR all stages	335 [35.4]	333 [35.4]	333 [35.4]	395 [41.7]	393 [41.7]	393 [41.7]	132 [13.9]	131 [13.9]	131 [13.9]	85 [9.0]	85 [9.0]	85 [9.0]	947 [100]	942 [100]	942 [100]
Correctly classified*	297 (88.7)	214 (64.3)	250 (75.1)	100 (25.3)	177 (45.0)	176 (44.8)	8 (6.1)	32 (24.4)	20 (15.3)	7 (8.2)	13 (15.3)	11 (12.9)	412 (43.5)	436 (46.3)	457 (48.5)
Stage as or more severe than mGFR	335 (100)	333 (100)	333 (100)	109 (27.6)	245 (62.3)	193 (49.1)	11 (8.3)	44 (33.6)	25 (19.1)	7 (8.2)	13 (15.3)	11 (12.9)	462 (48.8)	635 (67.4)	562 (59.7)
Absolute bias (mean: eGFR-mGFR)	2.1	-11	-4.5	23.0	7.6	14.8	35.5	21.5	27.7	47.0	31.9	38.4	19.5	5.1	11.9
Absolute bias (median: eGFR-mGFR)	6.0	-7.3	-1.9	23.5	6.0	14.1	37.7	17.1	27.0	49.3	31.3	36.9	19.4	4.4	11.0
Relative bias (mean: eGFR/mGFR)	1.04	0.91	0.97	1.32	1.11	1.20	1.68	1.41	1.53	2.53	2.09	2.29	1.38	1.17	1.26
Relative bias (median: eGFR/mGFR)	1.06	0.93	0.98	1.32	1.08	1.19	1.72	1.32	1.49	2.27	1.80	2.00	1.24	1.06	1.14
Precision, RMSE	21.4	25.2	22.2	16.2	22.4	18.0	21.5	24.8	22.3	26.4	26.4	25.5	24.7	27.9	25.2
Precision, IQR	-5.5;16.1	-26; 6.3	-15;10.0	12.3;34.1	-9.4;24.5	1.7;28.1	17.3;52.5	2.1;41.9	9.8;47.7	25.5;61.8	10.2;49.1	21.3;54.0	5.9;34.0	-12;23.2	-3.1;27.0
P10	0.44	0.37	0.42	0.15	0.24	0.24	0.06	0.17	0.15	0.02	0.05	0.05	0.23	0.26	0.27
P30	0.87	0.78	0.87	0.46	0.68	0.66	0.23	0.44	0.34	0.07	0.26	0.11	0.54	0.64	0.64
mGFR overestimated by 2+ ml/min	202 (60.3)	109 (32.7)	136 (40.8)	357 (90.4)	219 (55.7)	292 (74.3)	124 (93.9)	99 (75.6)	111 (84.7)	83 (97.6)	76 (89.4)	79 (92.9)	766 (80.9)	503 (53.4)	618 (65.6)
eGFR within +/- 2 ml/min of mGFR	28 (8.4)	29 (8.7)	34 (10.2)	12 (3.0)	22 (5.6)	26 (6.6)	2 (1.5)	10 (7.6)	5 (3.8)	0 (0.0)	2 (2.4)	3 (3.5)	42 (4.4)	63 (6.7)	68 (7.2)
mGFR underestimated by 2+ ml/min	105 (31.3)	195 (58.6)	163 (48.9)	26 (6.6)	152 (38.7)	75 (19.1)	6 (4.5)	22 (16.8)	15 (11.5)	2 (2.4)	7 (8.2)	3 (3.5)	139 (14.7)	376 (39.9)	256 (27.2)
mGFR overestimated by 5+ ml/min	180 (53.7)	92 (27.6)	115 (34.5)	346 (87.6)	203 (51.7)	276 (70.2)	121 (91.7)	93 (71.0)	105 (80.2)	81 (95.3)	74 (87.1)	78 (91.8)	728 (76.9)	462 (49.0)	574 (60.9)
eGFR within +/- 5 ml/min of mGFR	67 (20.0)	67 (20.1)	76 (22.8)	29 (7.3)	61 (15.5)	62 (15.8)	8 (6.1)	21 (16.0)	19 (14.5)	3 (3.5)	6 (7.1)	5 (5.9)	107 (11.3)	155 (16.5)	162 (17.2)
mGFR underestimated by 5+ ml/min	88 (26.3)	174 (52.3)	142 (42.6)	20 (5.1)	129 (32.8)	55 (14.0)	3 (2.3)	17 (13.0)	7 (5.3)	1 (1.2)	5 (5.9)	2 (2.4)	112 (11.8)	325 (34.5)	206 (21.9)
mGFR overestimated by 10+ ml/min	129 (38.5)	62 (18.6)	85 (25.5)	310 (78.5)	169 (43.0)	235 (59.8)	113 (85.6)	78 (59.5)	96 (73.3)	78 (91.8)	64 (75.3)	74 (87.1)	630 (66.5)	373 (39.6)	490 (52.0)
eGFR within +/- 10 ml/min of mGFR	145 (43.3)	116 (34.8)	144 (43.2)	75 (19.0)	130 (33.1)	126 (32.1)	19 (14.4)	42 (32.1)	32 (24.4)	6 (7.1)	19 (22.4)	9 (10.6)	245 (25.9)	307 (32.6)	311 (33.0)
mGFR underestimated by 10+ ml/min	61 (18.2)	155 (46.5)	104 (31.2)	10 (2.5)	94 (23.9)	32 (8.1)	0 (0.0)	11 (8.4)	3 (2.3)	1 (1.2)	2 (2.4)	2 (2.4)	72 (7.6)	262 (27.8)	141 (15.0)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S9 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	301 (89.9)	238 (71.0)	300 (89.6)	269 (68.1)	173 (43.8)	301 (76.2)	66 (50.0)	49 (37.1)	77 (58.3)	35 (41.2)	19 (22.4)	37 (43.5)	671 (70.9)	479 (50.6)	715 (75.5)
eGFR stage 2	30 (9.0)	90 (26.9)	29 (8.7)	115 (29.1)	209 (52.9)	85 (21.5)	51 (38.6)	67 (50.8)	40 (30.3)	29 (34.1)	43 (50.6)	30 (35.3)	225 (23.8)	409 (43.2)	184 (19.4)
eGFR stage 3a	3 (0.9)	5 (1.5)	5 (1.5)	11 (2.8)	13 (3.3)	9 (2.3)	14 (10.6)	13 (9.8)	11 (8.3)	14 (16.5)	15 (17.6)	12 (14.1)	42 (4.4)	46 (4.9)	37 (3.9)
eGFR stage 3b,4,5	1 (0.3)	2 (0.6)	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.8)	3 (2.3)	4 (3.0)	7 (8.2)	8 (9.4)	6 (7.1)	9 (1.0)	13 (1.4)	11 (1.2)
eGFR all stages	335 [35.4]	335 [35.4]	335 [35.4]	395 [41.7]	395 [41.7]	395 [41.7]	132 [13.9]	132 [13.9]	132 [13.9]	85 [9.0]	85 [9.0]	85 [9.0]	947 [100]	947 [100]	947 [100]
Correctly classified*	301 (89.9)	238 (71.0)	300 (89.6)	115 (29.1)	209 (52.9)	85 (21.5)	14 (10.6)	13 (9.8)	11 (8.3)	7 (8.2)	8 (9.4)	6 (7.1)	437 (46.1)	468 (49.4)	402 (42.4)
Stage as or more severe than mGFR	335 (100)	335 (100)	335 (100)	126 (31.9)	222 (56.2)	94 (23.8)	15 (11.4)	16 (12.1)	15 (11.4)	7 (8.2)	8 (9.4)	6 (7.1)	483 (51.0)	581 (61.4)	450 (47.5)
Absolute bias (mean: eGFR-mGFR)	7.6	-10	15.6	24.3	12.4	33.6	36.8	27.2	43.0	47.0	38.9	52.8	22.2	8.9	30.3
Absolute bias (median: eGFR-mGFR)	9.2	-7.1	17.6	22.3	12.9	30.9	37.5	28.3	46.7	47.1	41.5	50.1	21.1	8.2	28.3
Relative bias (mean: eGFR/mGFR)	1.09	0.92	1.16	1.33	1.17	1.45	1.70	1.52	1.82	2.55	2.28	2.73	1.41	1.23	1.52
Relative bias (median: eGFR/mGFR)	1.09	0.93	1.17	1.31	1.17	1.41	1.69	1.54	1.88	2.29	2.07	2.45	1.27	1.10	1.36
Precision, RMSE	25.8	19.8	29.3	20.3	14.0	26.0	24.8	18.0	30.2	28.6	22.5	33.6	26.9	23.9	31.0
Precision, IQR	-4.1;23.9	-17; 4.0	-0.0;34.7	11.2;36.6	2.9;21.5	16.1;50.1	13.8;55.7	11.2;42.4	17.5;62.9	26.3;63.9	23.4;52.2	31.4;74.3	6.6;37.7	-5.0;22.8	10.4;50.3
P10	0.35	0.47	0.26	0.15	0.32	0.12	0.11	0.10	0.08	0.02	0.04	0.02	0.20	0.31	0.16
P30	0.75	0.89	0.65	0.48	0.76	0.37	0.27	0.32	0.23	0.11	0.12	0.11	0.51	0.68	0.42
mGFR overestimated by 2+ ml/min	220 (65.7)	97 (29.0)	235 (70.1)	345 (87.3)	303 (76.7)	353 (89.4)	121 (91.7)	120 (90.9)	121 (91.7)	83 (97.6)	82 (96.5)	82 (96.5)	769 (81.2)	602 (63.6)	791 (83.5)
eGFR within +/- 2 ml/min of mGFR	20 (6.0)	27 (8.1)	26 (7.8)	16 (4.1)	33 (8.4)	12 (3.0)	7 (5.3)	4 (3.0)	4 (3.0)	1 (1.2)	1 (1.2)	2 (2.4)	44 (4.6)	65 (6.9)	44 (4.6)
mGFR underestimated by 2+ ml/min	95 (28.4)	211 (63.0)	74 (22.1)	34 (8.6)	59 (14.9)	30 (7.6)	4 (3.0)	8 (6.1)	7 (5.3)	1 (1.2)	2 (2.4)	1 (1.2)	134 (14.1)	280 (29.6)	112 (11.8)
mGFR overestimated by 5+ ml/min	198 (59.1)	75 (22.4)	223 (66.6)	330 (83.5)	282 (71.4)	342 (86.6)	118 (89.4)	117 (88.6)	118 (89.4)	82 (96.5)	79 (92.9)	82 (96.5)	728 (76.9)	553 (58.4)	765 (80.8)
eGFR within +/- 5 ml/min of mGFR	59 (17.6)	70 (20.9)	47 (14.0)	39 (9.9)	71 (18.0)	37 (9.4)	11 (8.3)	10 (7.6)	9 (6.8)	2 (2.4)	5 (5.9)	2 (2.4)	111 (11.7)	156 (16.5)	95 (10.0)
mGFR underestimated by 5+ ml/min	78 (23.3)	190 (56.7)	65 (19.4)	26 (6.6)	42 (10.6)	16 (4.1)	3 (2.3)	5 (3.8)	5 (3.8)	1 (1.2)	1 (1.2)	1 (1.2)	108 (11.4)	238 (25.1)	87 (9.2)
mGFR overestimated by 10+ ml/min	163 (48.7)	34 (10.1)	201 (60.0)	303 (76.7)	228 (57.7)	326 (82.5)	106 (80.3)	101 (76.5)	111 (84.1)	76 (89.4)	76 (89.4)	77 (90.6)	648 (68.4)	439 (46.4)	715 (75.5)
eGFR within +/- 10 ml/min of mGFR	111 (33.1)	150 (44.8)	81 (24.2)	76 (19.2)	144 (36.5)	58 (14.7)	26 (19.7)	30 (22.7)	19 (14.4)	8 (9.4)	8 (9.4)	7 (8.2)	221 (23.3)	332 (35.1)	165 (17.4)
mGFR underestimated by 10+ ml/min	61 (18.2)	151 (45.1)	53 (15.8)	16 (4.1)	23 (5.8)	11 (2.8)	0 (0.0)	1 (0.8)	2 (1.5)	1 (1.2)	1 (1.2)	1 (1.2)	78 (8.2)	176 (18.6)	67 (7.1)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S9 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	317 (94.6)	244 (72.8)	310 (92.5)	356 (90.1)	200 (50.6)	343 (86.8)	89 (67.4)	50 (37.9)	82 (62.1)	54 (63.5)	26 (30.6)	50 (58.8)	816 (86.2)	520 (54.9)	785 (82.9)
eGFR stage 2	16 (4.8)	85 (25.4)	23 (6.9)	37 (9.4)	183 (46.3)	50 (12.7)	36 (27.3)	64 (48.5)	44 (33.3)	19 (22.4)	38 (44.7)	26 (30.6)	108 (11.4)	370 (39.1)	143 (15.1)
eGFR stage 3a	1 (0.3)	4 (1.2)	1 (0.3)	2 (0.5)	12 (3.0)	2 (0.5)	6 (4.5)	15 (11.4)	5 (3.8)	9 (10.6)	13 (15.3)	6 (7.1)	18 (1.9)	44 (4.6)	14 (1.5)
eGFR stage 3b,4,5	1 (0.3)	2 (0.6)	1 (0.3)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.8)	3 (2.3)	1 (0.8)	3 (3.5)	8 (9.4)	3 (3.5)	5 (0.5)	13 (1.4)	5 (0.5)
eGFR all stages	335 [35.4]	335 [35.4]	335 [35.4]	395 [41.7]	395 [41.7]	395 [41.7]	132 [13.9]	132 [13.9]	132 [13.9]	85 [9.0]	85 [9.0]	85 [9.0]	947 [100]	947 [100]	947 [100]
Correctly classified*	317 (94.6)	244 (72.8)	310 (92.5)	37 (9.4)	183 (46.3)	50 (12.7)	6 (4.5)	15 (11.4)	5 (3.8)	3 (3.5)	8 (9.4)	3 (3.5)	363 (38.3)	450 (47.5)	368 (38.9)
Stage as or more severe than mGFR	335 (100)	335 (100)	335 (100)	39 (9.9)	195 (49.4)	52 (13.2)	7 (5.3)	18 (13.6)	6 (4.5)	3 (3.5)	8 (9.4)	3 (3.5)	384 (40.5)	556 (58.7)	396 (41.8)
Absolute bias (mean: eGFR-mGFR)	19.4	-2.5	19.6	38.7	16.9	36.4	49.7	30.5	48.2	60.1	41.7	58.1	35.3	14.2	34.1
Absolute bias (median: eGFR-mGFR)	23.7	-0.7	21.0	38.8	15.3	34.5	52.4	29.2	46.4	63.8	42.7	58.3	34.9	13.3	32.9
Relative bias (mean: eGFR/mGFR)	1.20	0.99	1.20	1.52	1.23	1.50	1.94	1.58	1.91	2.93	2.39	2.90	1.59	1.30	1.58
Relative bias (median: eGFR/mGFR)	1.22	0.99	1.20	1.53	1.20	1.46	2.00	1.55	1.87	2.64	2.13	2.58	1.44	1.17	1.42
Precision, RMSE	23.3	25.0	28.7	18.4	18.6	22.1	24.8	21.8	26.4	30.2	25.5	30.3	26.0	26.3	28.8
Precision, IQR	11.1;34.5	-15;14.5	4.6;39.1	26.7;51.5	3.5;28.4	20.9;50.7	28.1;68.8	12.4;47.1	26.2;68.4	36.6;77.3	21.3;55.9	34.7;74.7	20.5;51.3	-1.7;29.9	16.6;51.1
P10	0.13	0.37	0.20	0.05	0.24	0.07	0.04	0.09	0.03	0.01	0.04	0.01	0.08	0.25	0.10
P30	0.64	0.81	0.58	0.21	0.62	0.28	0.09	0.30	0.10	0.04	0.08	0.04	0.33	0.60	0.34
mGFR overestimated by 2+ ml/min	282 (84.2)	152 (45.4)	263 (78.5)	382 (96.7)	308 (78.0)	377 (95.4)	129 (97.7)	121 (91.7)	129 (97.7)	84 (98.8)	82 (96.5)	84 (98.8)	877 (92.6)	663 (70.0)	853 (90.1)
eGFR within +/- 2 ml/min of mGFR	4 (1.2)	20 (6.0)	10 (3.0)	6 (1.5)	27 (6.8)	7 (1.8)	3 (2.3)	5 (3.8)	2 (1.5)	0 (0.0)	1 (1.2)	0 (0.0)	13 (1.4)	53 (5.6)	19 (2.0)
mGFR underestimated by 2+ ml/min	49 (14.6)	163 (48.7)	62 (18.5)	7 (1.8)	60 (15.2)	11 (2.8)	0 (0.0)	6 (4.5)	1 (0.8)	1 (1.2)	2 (2.4)	1 (1.2)	57 (6.0)	231 (24.4)	75 (7.9)
mGFR overestimated by 5+ ml/min	276 (82.4)	136 (40.6)	248 (74.0)	377 (95.4)	288 (72.9)	373 (94.4)	127 (96.2)	118 (89.4)	128 (97.0)	83 (97.6)	81 (95.3)	83 (97.6)	863 (91.1)	623 (65.8)	832 (87.9)
eGFR within +/- 5 ml/min of mGFR	15 (4.5)	54 (16.1)	34 (10.1)	12 (3.0)	66 (16.7)	16 (4.1)	5 (3.8)	11 (8.3)	4 (3.0)	1 (1.2)	3 (3.5)	1 (1.2)	33 (3.5)	134 (14.1)	55 (5.8)
mGFR underestimated by 5+ ml/min	44 (13.1)	145 (43.3)	53 (15.8)	6 (1.5)	41 (10.4)	6 (1.5)	0 (0.0)	3 (2.3)	0 (0.0)	1 (1.2)	1 (1.2)	1 (1.2)	51 (5.4)	190 (20.1)	60 (6.3)
mGFR overestimated by 10+ ml/min	257 (76.7)	108 (32.2)	230 (68.7)	366 (92.7)	250 (63.3)	359 (90.9)	124 (93.9)	105 (79.5)	126 (95.5)	82 (96.5)	77 (90.6)	82 (96.5)	829 (87.5)	540 (57.0)	797 (84.2)
eGFR within +/- 10 ml/min of mGFR	40 (11.9)	119 (35.5)	63 (18.8)	25 (6.3)	125 (31.6)	32 (8.1)	8 (6.1)	26 (19.7)	6 (4.5)	2 (2.4)	7 (8.2)	3 (3.5)	75 (7.9)	277 (29.3)	104 (11.0)
mGFR underestimated by 10+ ml/min	38 (11.3)	108 (32.2)	42 (12.5)	4 (1.0)	20 (5.1)	4 (1.0)	0 (0.0)	1 (0.8)	0 (0.0)	1 (1.2)	1 (1.2)	0 (0.0)	43 (4.5)	130 (13.7)	46 (4.9)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

Table S10 (a): Uganda – Performance of eGFR equations compared to iohexol GFR

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[4]P_{10}	[5]P_{30}
Cockroft Gault (adjusted for BSA)	0.0	1.00	32.2	0.24	0.65
MDRD	3.0	1.03	31.3	0.25	0.65
MDRD ethnicity coefficient	22.4	1.25	34.4	0.20	0.50
FAS (creatinine)	-0.4	0.99	29.7	0.26	0.68
Lund-Malmö (revised)	-3.0	0.97	27.5	0.27	0.72
CKD-EPI (creatinine)	5.9	1.07	28.0	0.27	0.68
CKD-EPI (creatinine) ethnicity coefficient	21.6	1.24	29.2	0.21	0.52
CKD-EPI (creatinine + cystatin C)	-0.5	0.99	25.8	0.32	0.74
CKD-EPI (cystatin C)	-6.7	0.92	26.9	0.29	0.72

[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3]Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=733 for creatinine-based equations; N=620 for cystatin C-based equations

Table S10 (b): Uganda – GFR stage comparing eGFR equations to iohexol GFR

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	350 (**48**)	360 (**49**)	377 (**51**)	575 (**78**)	331 (**45**)	306 (**42**)	478 (**65**)	605 (**83**)	288 (**47**)	216 (**35**)
G2	282 (**39**)	306 (**42**)	312 (**43**)	137 (**19**)	341 (**47**)	385 (**53**)	223 (**30**)	105 (**14**)	287 (**46**)	312 (**50**)
G3a	71 (**10**)	45 (**6**)	30 (**4**)	9 (**1**)	39 (**5**)	27 (**4**)	18 (**3**)	10 (**1**)	32 (**5**)	71 (**12**)
G3b, G4, G5	30 (**4**)	22 (**3**)	14 (**2**)	12 (**2**)	22 (**3**)	15 (**2**)	14 (**2**)	13 (**2**)	13 (**2**)	21 (**3**)
% classified correctly	reference category	386 (**53**)	390 (**53**)	384 (**52**)	393 (**54**)	406 (**55**)	398 (**54**)	377 (**51**)	349 (**56**)	326 (**53**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15
[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;
[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;
[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.
N=733 for creatinine-based equations; N=620 for cystatin C-based equations

Table S10 (c): Uganda – overall and by iohexol GFR category: performance of creatinine and cystatin C-based eGFR equations

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	279 (79.7)	149 (51.6)	189 (65.4)	158 (56.0)	54 (21.7)	83 (33.3)	30 (42.3)	7 (12.1)	11 (19.0)	11 (36.7)	6 (25.0)	5 (20.8)	478 (65.2)	216 (34.8)	288 (46.5)
eGFR stage 2	67 (19.1)	121 (41.9)	93 (32.2)	110 (39.0)	152 (61.0)	147 (59.0)	32 (45.1)	27 (46.6)	33 (56.9)	14 (46.7)	12 (50.0)	14 (58.3)	223 (30.4)	312 (50.3)	287 (46.3)
eGFR stage 3a	1 (0.3)	14 (4.8)	6 (2.1)	11 (3.9)	36 (14.5)	16 (6.4)	5 (7.0)	20 (34.5)	9 (15.5)	1 (3.3)	1 (4.2)	1 (4.2)	18 (2.5)	71 (11.5)	32 (5.2)
eGFR stage 3b,4,5	3 (0.9)	5 (1.7)	1 (0.3)	3 (1.1)	7 (2.8)	3 (1.2)	4 (5.6)	4 (6.9)	5 (8.6)	4 (13.3)	5 (20.8)	4 (16.7)	14 (1.9)	21 (3.4)	13 (2.1)
eGFR all stages	350 [47.7]	289 [46.6]	289 [46.6]	282 [38.5]	249 [40.2]	249 [40.2]	71 [9.7]	58 [9.4]	58 [9.4]	30 [4.1]	24 [3.9]	24 [3.9]	733 [100]	620 [100]	620 [100]
Correctly classified*	279 (79.7)	149 (51.6)	189 (65.4)	110 (39.0)	152 (61.0)	147 (59.0)	5 (7.0)	20 (34.5)	9 (15.5)	4 (13.3)	5 (20.8)	4 (16.7)	398 (54.3)	326 (52.6)	349 (56.3)
Stage as or more severe than mGFR	350 (100)	289 (100)	289 (100)	124 (44.0)	195 (78.3)	166 (66.7)	9 (12.7)	24 (41.4)	14 (24.1)	4 (13.3)	5 (20.8)	4 (16.7)	487 (66.4)	513 (82.7)	473 (76.3)
Absolute bias (mean: eGFR-mGFR)	-12	-23	-17	14.6	0.8	7.1	30.0	13.1	19.2	45.6	34.0	39.7	4.6	-8.0	-1.6
Absolute bias (median: eGFR-mGFR)	-9.5	-21	-14	15.6	1.9	7.4	31.0	10.4	18.9	49.4	33.1	40.7	5.9	-6.7	-0.5
Relative bias (mean: eGFR/mGFR)	0.91	0.81	0.87	1.20	1.02	1.10	1.57	1.25	1.36	2.76	2.60	2.75	1.16	1.00	1.08
Relative bias (median: eGFR/mGFR)	0.91	0.81	0.87	1.21	1.03	1.09	1.61	1.21	1.36	2.35	1.90	2.12	1.07	0.92	0.99
Precision, RMSE	24.5	23.7	22.6	17.8	18.9	17.0	21.7	18.9	19.3	26.6	30.9	29.7	28.0	26.9	25.8
Precision, IQR	-25; 4.0	-38;-7.1	-29;-1.0	4.5;26.3	-12;14.6	-4.6;18.6	20.9;43.5	0.4;21.4	6.7;30.5	34.0;57.7	9.7;51.0	23.1;54.4	-12;22.8	-24; 8.9	-17;14.3
P10	0.39	0.28	0.37	0.20	0.31	0.33	0.07	0.33	0.16	0.10	0.13	0.08	0.27	0.29	0.32
P30	0.85	0.73	0.83	0.65	0.78	0.77	0.15	0.62	0.43	0.10	0.17	0.17	0.68	0.72	0.74
mGFR overestimated by 2+ ml/min	101 (28.9)	35 (12.1)	55 (19.0)	222 (78.7)	123 (49.4)	152 (61.0)	65 (91.5)	40 (69.0)	47 (81.0)	26 (86.7)	19 (79.2)	20 (83.3)	414 (56.5)	217 (35.0)	274 (44.2)
eGFR within +/- 2 ml/min of mGFR	33 (9.4)	9 (3.1)	22 (7.6)	11 (3.9)	21 (8.4)	24 (9.6)	0 (0.0)	9 (15.5)	5 (8.6)	3 (10.0)	2 (8.3)	1 (4.2)	47 (6.4)	41 (6.6)	52 (8.4)
mGFR underestimated by 2+ ml/min	216 (61.7)	245 (84.8)	212 (73.4)	49 (17.4)	105 (42.2)	73 (29.3)	6 (8.5)	9 (15.5)	6 (10.3)	1 (3.3)	3 (12.5)	3 (12.5)	272 (37.1)	362 (58.4)	294 (47.4)
mGFR overestimated by 5+ ml/min	79 (22.6)	26 (9.0)	44 (15.2)	208 (73.8)	104 (41.8)	136 (54.6)	62 (87.3)	36 (62.1)	45 (77.6)	26 (86.7)	18 (75.0)	20 (83.3)	375 (51.2)	184 (29.7)	245 (39.5)
eGFR within +/- 5 ml/min of mGFR	69 (19.7)	35 (12.1)	51 (17.6)	34 (12.1)	51 (20.5)	52 (20.9)	5 (7.0)	16 (27.6)	9 (15.5)	3 (10.0)	3 (12.5)	2 (8.3)	111 (15.1)	105 (16.9)	114 (18.4)
mGFR underestimated by 5+ ml/min	202 (57.7)	228 (78.9)	194 (67.1)	40 (14.2)	94 (37.8)	61 (24.5)	4 (5.6)	6 (10.3)	4 (6.9)	1 (3.3)	3 (12.5)	2 (8.3)	247 (33.7)	331 (53.4)	261 (42.1)
mGFR overestimated by 10+ ml/min	54 (15.4)	16 (5.5)	26 (9.0)	183 (64.9)	81 (32.5)	106 (42.6)	61 (85.9)	29 (50.0)	41 (70.7)	26 (86.7)	18 (75.0)	20 (83.3)	324 (44.2)	144 (23.2)	193 (31.1)
eGFR within +/- 10 ml/min of mGFR	128 (36.6)	77 (26.6)	101 (34.9)	72 (25.5)	103 (41.4)	106 (42.6)	6 (8.5)	25 (43.1)	14 (24.1)	3 (10.0)	4 (16.7)	3 (12.5)	209 (28.5)	209 (33.7)	224 (36.1)
mGFR underestimated by 10+ ml/min	168 (48.0)	196 (67.8)	162 (56.1)	27 (9.6)	65 (26.1)	37 (14.9)	4 (5.6)	4 (6.9)	3 (5.2)	1 (3.3)	2 (8.3)	1 (4.2)	200 (27.3)	267 (43.1)	203 (32.7)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S10 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	219 (62.6)	208 (59.4)	228 (65.1)	91 (32.3)	80 (28.4)	109 (38.7)	15 (21.1)	13 (18.3)	15 (21.1)	6 (20.0)	5 (16.7)	8 (26.7)	331 (45.2)	306 (41.7)	360 (49.1)
eGFR stage 2	120 (34.3)	135 (38.6)	110 (31.4)	162 (57.4)	185 (65.6)	143 (50.7)	41 (57.7)	45 (63.4)	38 (53.5)	18 (60.0)	20 (66.7)	15 (50.0)	341 (46.5)	385 (52.5)	306 (41.7)
eGFR stage 3a	8 (2.3)	4 (1.1)	9 (2.6)	23 (8.2)	13 (4.6)	24 (8.5)	7 (9.9)	9 (12.7)	10 (14.1)	1 (3.3)	1 (3.3)	2 (6.7)	39 (5.3)	27 (3.7)	45 (6.1)
eGFR stage 3b,4,5	3 (0.9)	3 (0.9)	3 (0.9)	6 (2.1)	4 (1.4)	6 (2.1)	8 (11.3)	4 (5.6)	8 (11.3)	5 (16.7)	4 (13.3)	5 (16.7)	22 (3.0)	15 (2.0)	22 (3.0)
eGFR all stages	350 [47.7]	350 [47.7]	350 [47.7]	282 [38.5]	282 [38.5]	282 [38.5]	71 [9.7]	71 [9.7]	71 [9.7]	30 [4.1]	30 [4.1]	30 [4.1]	733 [100]	733 [100]	733 [100]
Correctly classified*	219 (62.6)	208 (59.4)	228 (65.1)	162 (57.4)	185 (65.6)	143 (50.7)	7 (9.9)	9 (12.7)	10 (14.1)	5 (16.7)	4 (13.3)	5 (16.7)	393 (53.6)	406 (55.4)	386 (52.7)
Stage as or more severe than mGFR	350 (100)	350 (100)	350 (100)	191 (67.7)	202 (71.6)	173 (61.3)	15 (21.1)	13 (18.3)	18 (25.4)	5 (16.7)	4 (13.3)	5 (16.7)	561 (76.5)	569 (77.6)	546 (74.5)
Absolute bias (mean: eGFR-mGFR)	-17	-22	-13	8.1	6.4	11.0	20.6	21.2	21.4	38.9	38.3	40.3	-1.2	-4.7	1.8
Absolute bias (median: eGFR-mGFR)	-16	-20	-13	6.5	6.6	8.5	20.4	22.5	24.0	39.1	40.4	43.5	-0.4	-3.0	0.0
Relative bias (mean: eGFR/mGFR)	0.87	0.82	0.90	1.11	1.09	1.15	1.39	1.40	1.41	2.51	2.50	2.52	1.08	1.05	1.11
Relative bias (median: eGFR/mGFR)	0.85	0.82	0.88	1.09	1.09	1.11	1.39	1.43	1.46	2.01	2.07	2.26	0.99	0.97	1.00
Precision, RMSE	27.9	23.3	31.1	21.2	16.3	25.4	22.0	18.8	26.0	27.0	24.0	28.1	29.7	27.5	32.2
Precision, IQR	-33; 0.1	-35;-6.7	-31; 4.8	-5.8;20.5	-3.0;16.4	-6.3;25.4	6.6;31.6	11.6;32.5	2.9;36.1	23.8;50.8	27.6;51.3	18.9;61.7	-19;18.0	-21;12.3	-19;21.1
P10	0.26	0.27	0.25	0.31	0.34	0.27	0.10	0.08	0.18	0.10	0.10	0.07	0.26	0.27	0.24
P30	0.76	0.77	0.74	0.72	0.82	0.66	0.37	0.31	0.38	0.10	0.10	0.13	0.68	0.72	0.65
mGFR overestimated by 2+ ml/min	82 (23.4)	44 (12.6)	93 (26.6)	173 (61.3)	179 (63.5)	174 (61.7)	60 (84.5)	62 (87.3)	57 (80.3)	26 (86.7)	26 (86.7)	26 (86.7)	341 (46.5)	311 (42.4)	350 (47.7)
eGFR within +/- 2 ml/min of mGFR	17 (4.9)	11 (3.1)	14 (4.0)	20 (7.1)	27 (9.6)	24 (8.5)	1 (1.4)	2 (2.8)	2 (2.8)	3 (10.0)	2 (6.7)	2 (6.7)	41 (5.6)	42 (5.7)	42 (5.7)
mGFR underestimated by 2+ ml/min	251 (71.7)	295 (84.3)	243 (69.4)	89 (31.6)	76 (27.0)	84 (29.8)	10 (14.1)	7 (9.9)	12 (16.9)	1 (3.3)	2 (6.7)	2 (6.7)	351 (47.9)	380 (51.8)	341 (46.5)
mGFR overestimated by 5+ ml/min	71 (20.3)	28 (8.0)	86 (24.6)	154 (54.6)	162 (57.4)	158 (56.0)	56 (78.9)	60 (84.5)	49 (69.0)	26 (86.7)	26 (86.7)	26 (86.7)	307 (41.9)	276 (37.7)	319 (43.5)
eGFR within +/- 5 ml/min of mGFR	41 (11.7)	49 (14.0)	42 (12.0)	56 (19.9)	58 (20.6)	53 (18.8)	7 (9.9)	6 (8.5)	12 (16.9)	3 (10.0)	3 (10.0)	2 (6.7)	107 (14.6)	116 (15.8)	109 (14.9)
mGFR underestimated by 5+ ml/min	238 (68.0)	273 (78.0)	222 (63.4)	72 (25.5)	62 (22.0)	71 (25.2)	8 (11.3)	5 (7.0)	10 (14.1)	1 (3.3)	1 (3.3)	2 (6.7)	319 (43.5)	341 (46.5)	305 (41.6)
mGFR overestimated by 10+ ml/min	54 (15.4)	15 (4.3)	76 (21.7)	112 (39.7)	115 (40.8)	134 (47.5)	48 (67.6)	56 (78.9)	47 (66.2)	26 (86.7)	26 (86.7)	25 (83.3)	240 (32.7)	212 (28.9)	282 (38.5)
eGFR within +/- 10 ml/min of mGFR	90 (25.7)	94 (26.9)	81 (23.1)	121 (42.9)	122 (43.3)	98 (34.8)	19 (26.8)	11 (15.5)	18 (25.4)	3 (10.0)	3 (10.0)	4 (13.3)	233 (31.8)	230 (31.4)	201 (27.4)
mGFR underestimated by 10+ ml/min	206 (58.9)	241 (68.9)	193 (55.1)	49 (17.4)	45 (16.0)	50 (17.7)	4 (5.6)	4 (5.6)	6 (8.5)	1 (3.3)	1 (3.3)	1 (3.3)	260 (35.5)	291 (39.7)	250 (34.1)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S10 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	317 (90.6)	233 (66.6)	307 (87.7)	223 (79.1)	115 (40.8)	205 (72.7)	47 (66.2)	22 (31.0)	45 (63.4)	18 (60.0)	7 (23.3)	18 (60.0)	605 (82.5)	377 (51.4)	575 (78.4)
eGFR stage 2	30 (8.6)	108 (30.9)	40 (11.4)	52 (18.4)	147 (52.1)	71 (25.2)	16 (22.5)	39 (54.9)	19 (26.8)	7 (23.3)	18 (60.0)	7 (23.3)	105 (14.3)	312 (42.6)	137 (18.7)
eGFR stage 3a	1 (0.3)	6 (1.7)	1 (0.3)	4 (1.4)	17 (6.0)	3 (1.1)	4 (5.6)	6 (8.5)	3 (4.2)	1 (3.3)	1 (3.3)	2 (6.7)	10 (1.4)	30 (4.1)	9 (1.2)
eGFR stage 3b,4,5	2 (0.6)	3 (0.9)	2 (0.6)	3 (1.1)	3 (1.1)	3 (1.1)	4 (5.6)	4 (5.6)	4 (5.6)	4 (13.3)	4 (13.3)	3 (10.0)	13 (1.8)	14 (1.9)	12 (1.6)
eGFR all stages	350 [47.7]	350 [47.7]	350 [47.7]	282 [38.5]	282 [38.5]	282 [38.5]	71 [9.7]	71 [9.7]	71 [9.7]	30 [4.1]	30 [4.1]	30 [4.1]	733 [100]	733 [100]	733 [100]
Correctly classified*	317 (90.6)	233 (66.6)	307 (87.7)	52 (18.4)	147 (52.1)	71 (25.2)	4 (5.6)	6 (8.5)	3 (4.2)	4 (13.3)	4 (13.3)	3 (10.0)	377 (51.4)	390 (53.2)	384 (52.4)
Stage as or more severe than mGFR	350 (100)	350 (100)	350 (100)	59 (20.9)	167 (59.2)	77 (27.3)	8 (11.3)	10 (14.1)	7 (9.9)	4 (13.3)	4 (13.3)	3 (10.0)	421 (57.4)	531 (72.4)	437 (59.6)
Absolute bias (mean: eGFR-mGFR)	4.2	-13	8.7	29.1	12.2	30.9	43.2	27.0	44.1	58.3	44.7	61.4	19.8	2.9	22.9
Absolute bias (median: eGFR-mGFR)	6.6	-12	6.9	29.5	9.5	27.1	43.8	27.8	45.2	63.4	42.9	60.5	21.6	3.0	22.4
Relative bias (mean: eGFR/mGFR)	1.06	0.90	1.10	1.39	1.17	1.41	1.81	1.51	1.83	3.19	2.75	3.33	1.35	1.14	1.38
Relative bias (median: eGFR/mGFR)	1.06	0.88	1.07	1.40	1.13	1.37	1.87	1.50	1.82	2.73	2.18	2.65	1.24	1.03	1.25
Precision, RMSE	26.4	29.4	33.7	20.4	22.9	27.5	25.1	21.8	26.5	30.6	31.9	38.4	29.2	31.3	34.4
Precision, IQR	-9.6;21.0	-30; 4.7	-12;27.5	17.5;42.8	-1.7;24.8	14.1;46.3	33.6;58.5	14.4;42.0	28.8;62.9	46.3;73.1	27.9;50.0	43.0;68.6	1.9;38.2	-16;22.9	1.4;44.2
P10	0.35	0.28	0.30	0.11	0.27	0.14	0.01	0.07	0.01	0.00	0.07	0.00	0.21	0.25	0.20
P30	0.81	0.75	0.71	0.32	0.69	0.39	0.08	0.23	0.08	0.10	0.10	0.07	0.52	0.65	0.50
mGFR overestimated by 2+ ml/min	199 (56.9)	94 (26.9)	202 (57.7)	254 (90.1)	191 (67.7)	246 (87.2)	67 (94.4)	64 (90.1)	67 (94.4)	29 (96.7)	27 (90.0)	29 (96.7)	549 (74.9)	376 (51.3)	544 (74.2)
eGFR within +/- 2 ml/min of mGFR	27 (7.7)	25 (7.1)	17 (4.9)	8 (2.8)	22 (7.8)	9 (3.2)	0 (0.0)	1 (1.4)	0 (0.0)	0 (0.0)	2 (6.7)	0 (0.0)	35 (4.8)	50 (6.8)	26 (3.5)
mGFR underestimated by 2+ ml/min	124 (35.4)	231 (66.0)	131 (37.4)	20 (7.1)	69 (24.5)	27 (9.6)	4 (5.6)	6 (8.5)	4 (5.6)	1 (3.3)	1 (3.3)	1 (3.3)	149 (20.3)	307 (41.9)	163 (22.2)
mGFR overestimated by 5+ ml/min	185 (52.9)	87 (24.9)	185 (52.9)	249 (88.3)	170 (60.3)	236 (83.7)	66 (93.0)	62 (87.3)	66 (93.0)	28 (93.3)	26 (86.7)	29 (96.7)	528 (72.0)	345 (47.1)	516 (70.4)
eGFR within +/- 5 ml/min of mGFR	61 (17.4)	47 (13.4)	47 (13.4)	18 (6.4)	53 (18.8)	25 (8.9)	1 (1.4)	4 (5.6)	1 (1.4)	1 (3.3)	3 (10.0)	0 (0.0)	81 (11.1)	107 (14.6)	73 (10.0)
mGFR underestimated by 5+ ml/min	104 (29.7)	216 (61.7)	118 (33.7)	15 (5.3)	59 (20.9)	21 (7.4)	4 (5.6)	5 (7.0)	4 (5.6)	1 (3.3)	1 (3.3)	1 (3.3)	124 (16.9)	281 (38.3)	144 (19.6)
mGFR overestimated by 10+ ml/min	153 (43.7)	70 (20.0)	166 (47.4)	236 (83.7)	137 (48.6)	226 (80.1)	64 (90.1)	59 (83.1)	64 (90.1)	26 (86.7)	26 (86.7)	26 (86.7)	479 (65.3)	292 (39.8)	482 (65.8)
eGFR within +/- 10 ml/min of mGFR	110 (31.4)	89 (25.4)	93 (26.6)	36 (12.8)	102 (36.2)	47 (16.7)	4 (5.6)	8 (11.3)	4 (5.6)	3 (10.0)	3 (10.0)	3 (10.0)	153 (20.9)	202 (27.6)	147 (20.1)
mGFR underestimated by 10+ ml/min	87 (24.9)	191 (54.6)	91 (26.0)	10 (3.5)	43 (15.2)	9 (3.2)	3 (4.2)	4 (5.6)	3 (4.2)	1 (3.3)	1 (3.3)	1 (3.3)	101 (13.8)	239 (32.6)	104 (14.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

Figure S13 (a): Iohexol GFR overall and by country, restricted for r>0.985

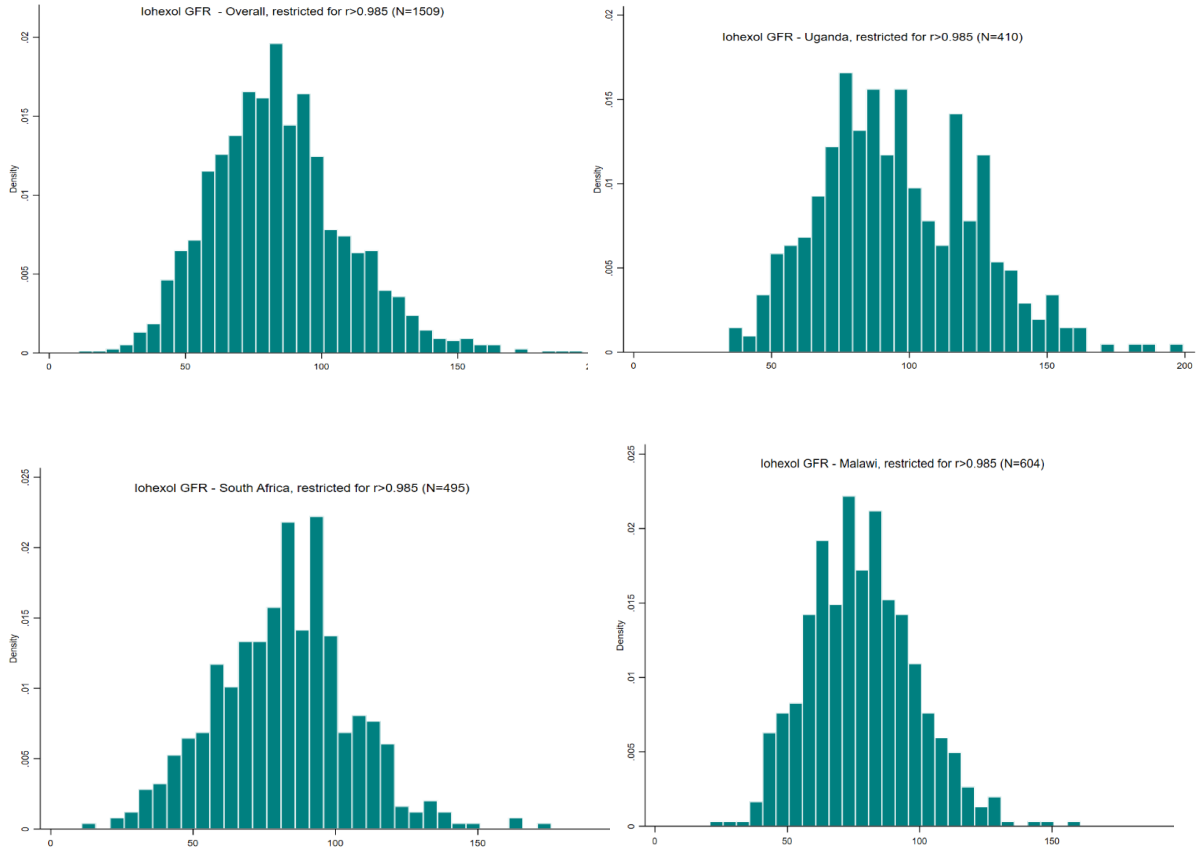
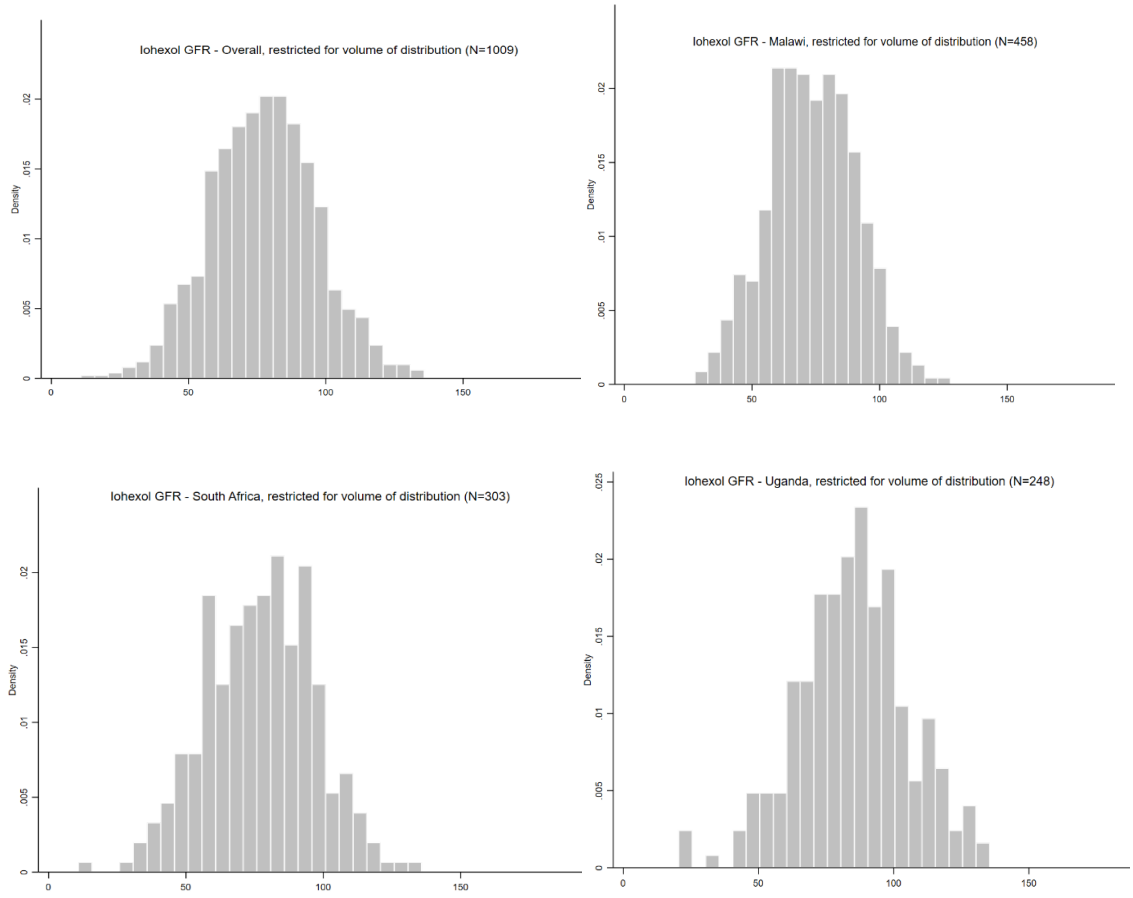
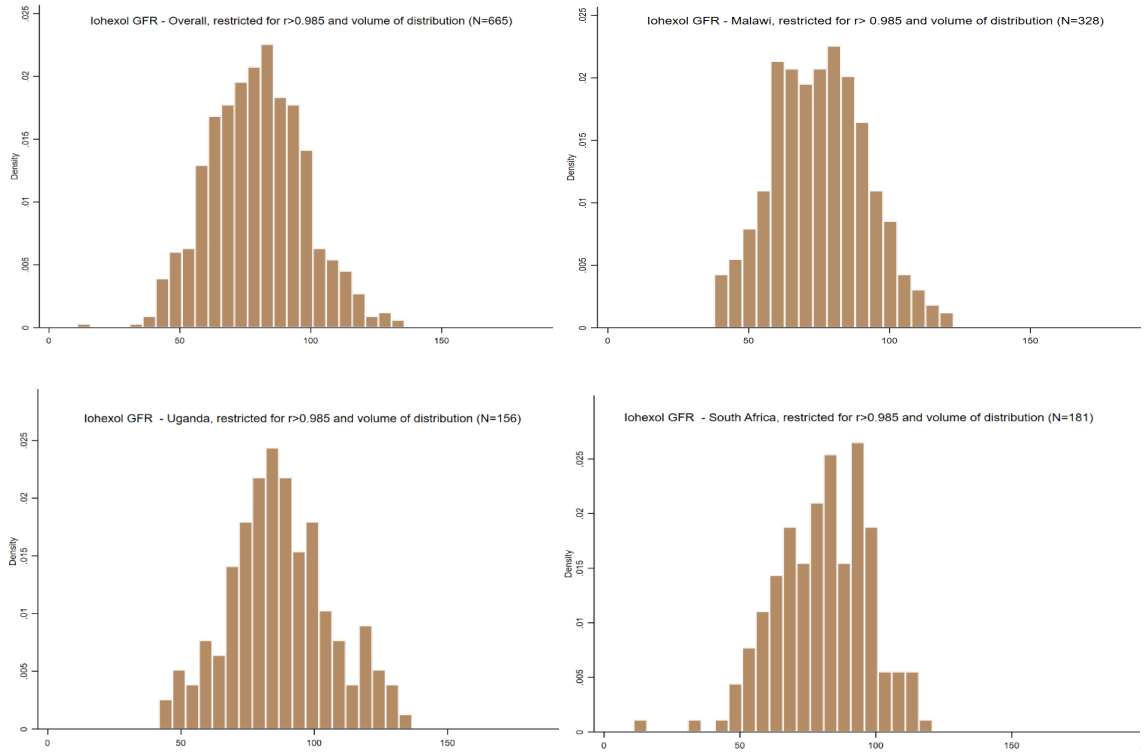


Figure S13 (b): Iohexol GFR overall and by country, restricted for VD*



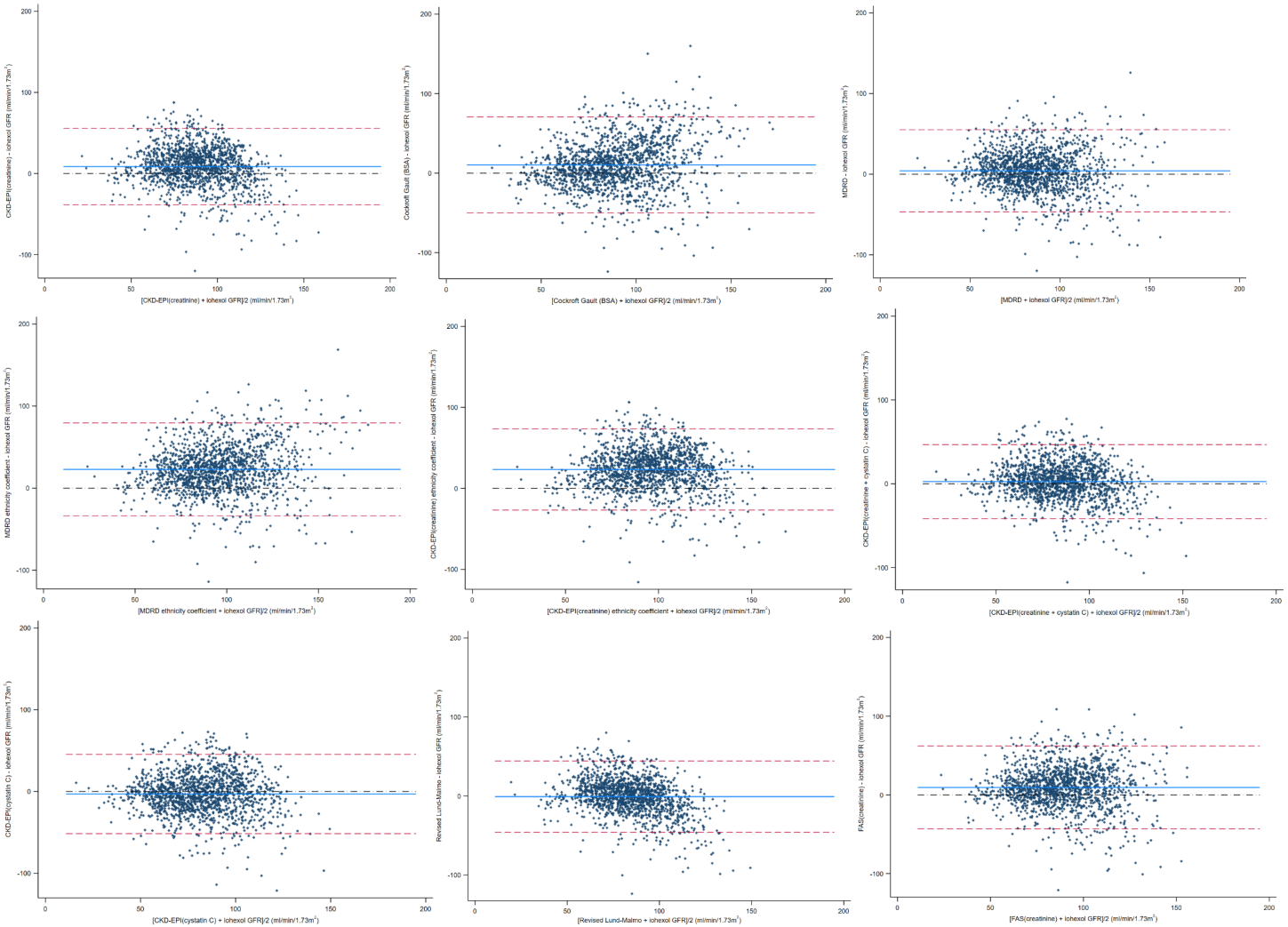
*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Figure S13 (c): Iohexol GFR overall and by country, restricted for r>0.985 + VD*



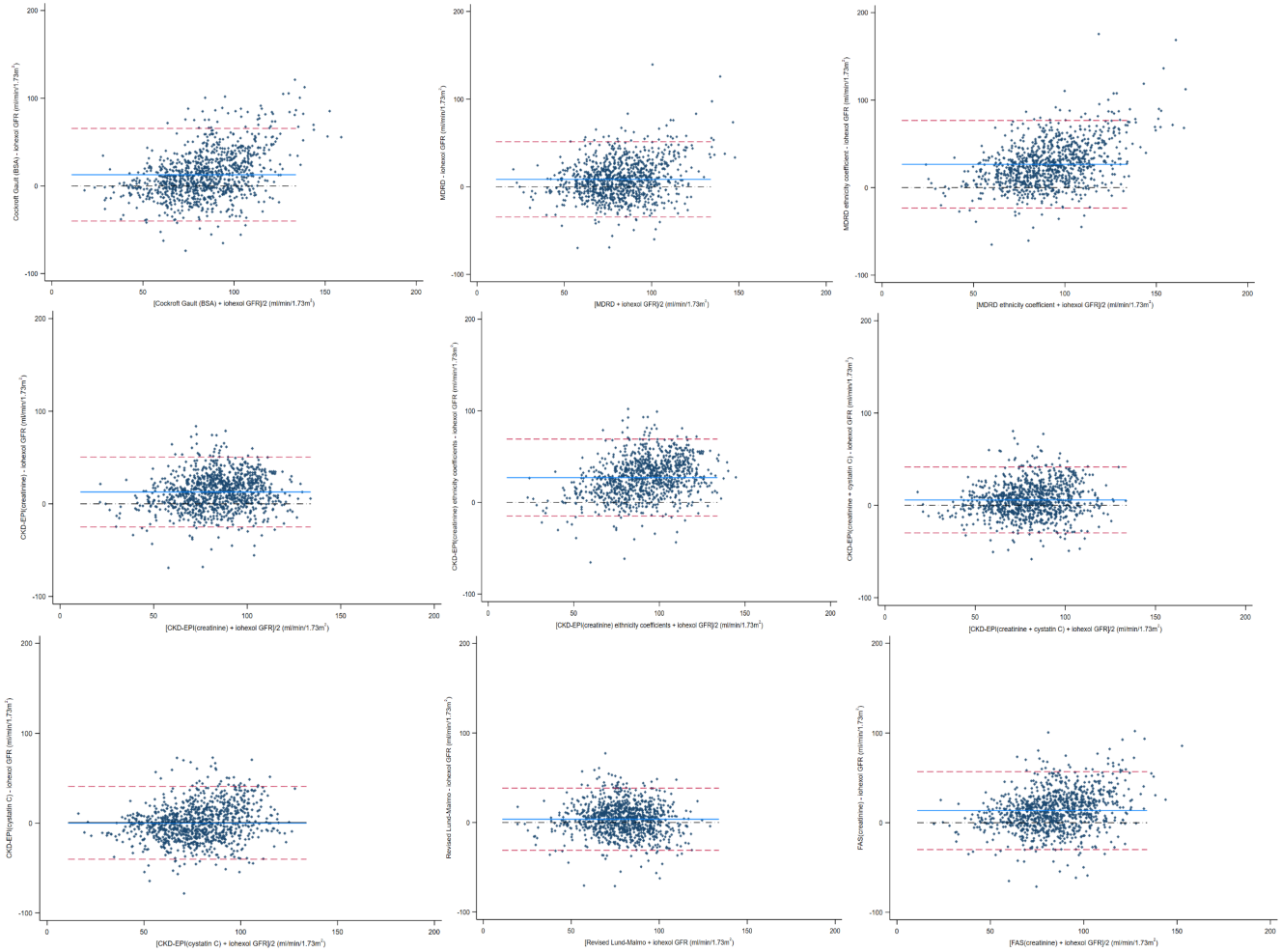
*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Figure S14 (a): Agreement: eGFR equations and iohexol GFR, restricted for r>0.985



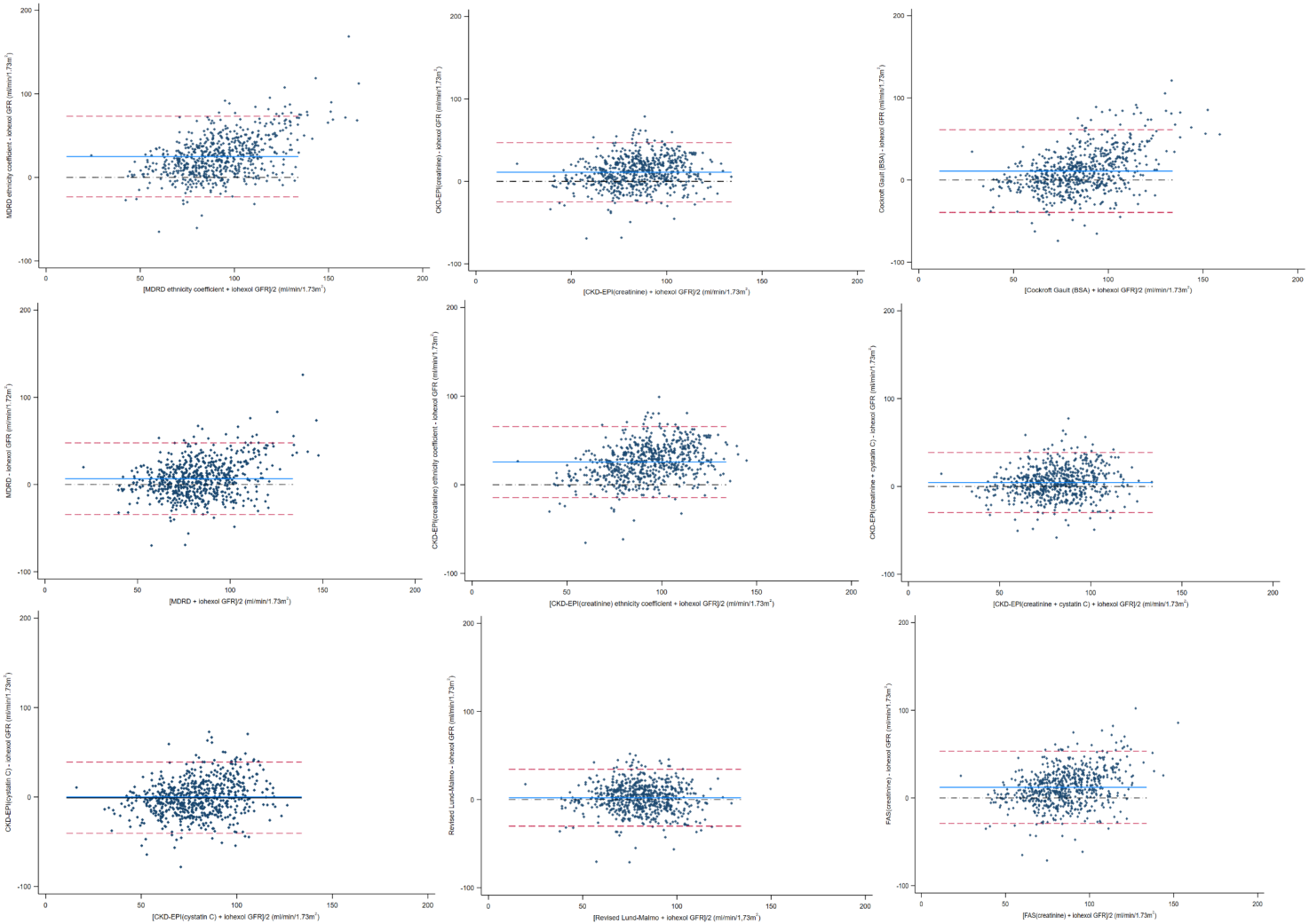
Red dashed lines indicate limits of agreement (+-2 SD); solid blue line indicates bias (ml/min/1.72m^2); normal ranges for sex-specific volume of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Figure S14 (b): Agreement: eGFR equations and iohexol GFR restricted for VD*



Red dashed lines indicate limits of agreement (+-2 SD); solid blue line indicates bias (ml/min/1.72m²); *VD: normal ranges for sex-specific volume of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Figure 5.14 (c): Agreement: eGFR equations and iohexol GFR restricted for r>0.985 + VD*



Red dashed lines indicate limits of agreement (+-2 SD); solid blue line indicates bias (ml/min/1.72m^2); *VD: normal ranges for sex-specific volume of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Table S11 (a): GFR stage: comparing eGFR equations to iohexol GFR restricted for r>0.985

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	566 (**38**)	783 (**52**)	660 (**44**)	1082 (**72**)	814 (**54**)	553 (**37**)	868 (**58**)	1157 (**77**)	618 (**43**)	481 (**34**)
G2	702 (**47**)	571 (**38**)	711 (**47**)	378 (**25**)	587 (**39**)	831 (**55**)	549 (**36**)	302 (**20**)	680 (**48**)	661 (**46**)
G3a	181 (**12**)	116 (**8**)	110 (**7**)	32 (**2**)	80 (**5**)	93 (**6**)	64 (**4**)	31 (**2**)	96 (**7**)	221 (**15**)
G3b, G4, G5	60 (**4**)	39 (**3**)	28 (**2**)	17(**1**)	28 (**2**)	32 (**2**)	28 (**2**)	19 (**1**)	37 (**3**)	68 (**5**)
% classified correctly	reference category	784 (**52**)	821 (**54**)	726 (**48**)	823 (**55**)	865 (**57**)	795 (**53**)	709 (**47**)	809 (**57**)	733 (**51**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15
[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;
[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;
[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.
N=1509 for creatinine-based equations; N=1431 for cystatin C-based equations

Table S11 (b): GFR stage: comparing eGFR equations to iohexol GFR for VD*

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	259 (**26**)	473 (**47**)	395 (**39**)	681 (**68**)	493 (**49**)	324 (**32**)	541 (**54**)	732 (**73**)	372 (**38**)	283 (**29**)
G2	566 (**56**)	407 (**40**)	497 (**49**)	284 (**28**)	422 (**42**)	574 (**57**)	380 (**38**)	228 (**23**)	475 (**49**)	470 (**49**)
G3a	135 (**13**)	85 (**8**)	86 (**9**)	29 (**3**)	65 (**6**)	80 (**8**)	59 (**6**)	31 (**3**)	86 (**9**)	157 (**16**)
G3b, G4, G5	49 (**5**)	44 (**4**)	31 (**3**)	15(**2**)	29 (**3**)	31 (**3**)	29 (**3**)	18 (**2**)	36 (**4**)	59 (**6**)
% classified correctly	reference category	504 (**50**)	548 (**54**)	442 (**44**)	524 (**52**)	591 (**59**)	510 (**51**)	415 (**41**)	551 (**57**)	499 (**52**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15

[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;

[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;

[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation

N=1009 for creatinine-based equations; N=969 for cystatin C-based equations.

*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Table S11 (c): GFR stage: comparing eGFR equations to iohexol GFR restricted for r>0.985 + VD*

GFR stage[1]	mGFR[2]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	185 (**28**)	309 (**47**)	260 (**39**)	449 (**68**)	331 (**50**)	212 (**32**)	353 (**53**)	487 (**73**)	247 (**39**)	190 (**30**)
G2	386 (**58**)	276 (**42**)	331 (**50**)	189 (**28**)	280 (**42**)	388 (**58**)	261 (**39**)	149 (**22**)	321 (**50**)	322 (**50**)
G3a	79 (**12**)	57 (**9**)	57 (**9**)	21 (**3**)	38 (**6**)	47 (**7**)	35 (**5**)	21 (**3**)	51 (**8**)	97 (**15**)
G3b, G4, G5	15 (**2**)	23 (**4**)	17 (**3**)	6(**1**)	16 (**2**)	18 (**3**)	16 (**2**)	8 (**1**)	22 (**3**)	32 (**5**)
% classified correctly	reference category	355 (**53**)	387 (**58**)	313 (**47**)	378 (**57**)	423 (**64**)	362 (**54**)	293 (**44**)	379 (**59**)	344 (**54**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15

[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;

[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;

[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.

N=665 for creatinine-based equations; N=641 for cystatin C-based equations.

*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Table S12 (a): eGFR equations compared to iohexol GFR restricted for r>0.985

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[5]P_{10}	[5]P_{30}
Cockroft Gault (adjusted for BSA)	8.9	1.11	30.8	0.25	0.64
MDRD	4.1	1.05	26.0	0.28	0.70
MDRD ethnicity coefficient	22.4	1.27	28.8	0.20	0.50
FAS (creatinine)	9.3	1.12	26.9	0.27	0.66
Lund-Malmö (revised)	0.5	1.01	23.0	0.34	0.78
CKD-EPI (creatinine)	9.2	1.11	24.1	0.30	0.69
CKD-EPI (creatinine) ethnicity coefficient	24.5	1.28	25.5	0.17	0.49
CKD-EPI (creatinine + cystatin C)	2.7	1.03	22.5	0.34	0.76
CKD-EPI (cystatin C)	-3.0	0.96	24.8	0.30	0.73

[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3]Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=1509 creatinine; N=1431 cystatin C

Table S12 (b): eGFR equations compared to iohexol GFR for VD*

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[4]P_{10}	[5]P_{30}
Cockroft Gault (adjusted for BSA)	10.0	1.13	26.9	0.27	0.66
MDRD	7.2	1.10	21.9	0.28	0.72
MDRD ethnicity coefficient	25.1	1.33	25.5	0.18	0.46
FAS (creatinine)	12.2	1.17	22.2	0.27	0.66
Lund-Malmö (revised)	4.1	1.05	17.7	0.37	0.82
CKD-EPI (creatinine)	12.8	1.17	19.3	0.29	0.67
CKD-EPI (creatinine) ethnicity coefficient	27.0	1.35	21.4	0.14	0.42
CKD-EPI (creatinine + cystatin C)	4.9	1.06	18.2	0.36	0.79
CKD-EPI (cystatin C)	-0.8	0.99	20.7	0.30	0.76

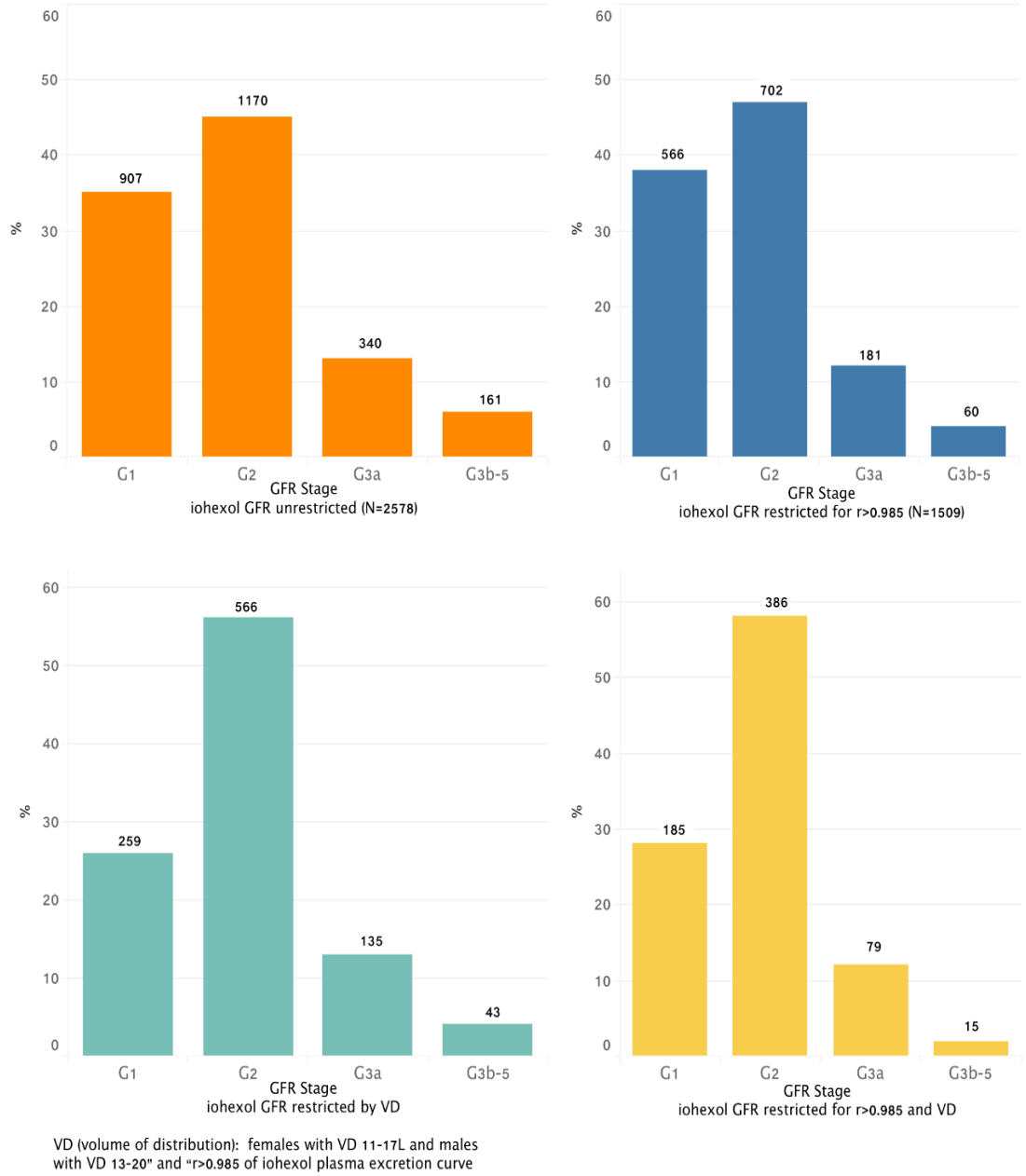
[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3]Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=1009 for creatinine; N=969 for cystatin C
*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Table S12 (c): eGFR equations compared to iohexol GFR for r>0.985 + VD*

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[4]P_{10}	[5]P_{30}
Cockroft Gault (adjusted for BSA)	8.4	1.11	25.7	0.28	0.70
MDRD	5.7	1.07	20.9	0.30	0.75
MDRD ethnicity coefficient	23.0	1.30	24.7	0.20	0.49
FAS (creatinine)	10.7	1.14	21.0	0.29	0.69
Lund-Malmö (revised)	2.5	1.03	16.5	0.40	0.87
CKD-EPI (creatinine)	11.4	1.15	18.3	0.31	0.71
CKD-EPI (creatinine) ethnicity coefficient	25.5	1.33	20.4	0.15	0.46
CKD-EPI (creatinine + cystatin C)	4.3	1.05	17.4	0.37	0.81
CKD-EPI (cystatin C)	-1.4	0.98	20.4	0.32	0.78

[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3]Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=665 creatinine; N=641 cystatin C
*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males.[4]

Figure S15: GFR stage for iohexol GFR, restricted for r<0.985, VD*, r<0.985 + VD*



*VD: Normal ranges for sex-specific volumes of distribution were derived from the British Nuclear Medicine Society Guidelines defined as 11-17 litres for females; 13-30 litres for males[4]

Table S13 (a): Overall and by iohexol GFR category: performance of creatinine and cystatin C-based eGFR equations restricted for r>0.985

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	452 (79.9)	284 (54.6)	355 (68.3)	352 (50.1)	169 (24.9)	223 (32.9)	47 (26.0)	25 (14.3)	32 (18.3)	17 (28.3)	3 (5.2)	8 (13.8)	868 (57.5)	481 (33.6)	618 (43.2)
eGFR stage 2	105 (18.6)	201 (38.7)	152 (29.2)	315 (44.9)	369 (54.4)	406 (59.9)	102 (56.4)	68 (38.9)	92 (52.6)	27 (45.0)	23 (39.7)	30 (51.7)	549 (36.4)	661 (46.2)	680 (47.5)
eGFR stage 3a	5 (0.9)	24 (4.6)	11 (2.1)	31 (4.4)	117 (17.3)	43 (6.3)	20 (11.0)	64 (36.6)	35 (20.0)	8 (13.3)	16 (27.6)	7 (12.1)	64 (4.2)	221 (15.4)	96 (6.7)
eGFR stage 3b,4,5	4 (0.7)	11 (2.1)	2 (0.4)	4 (0.6)	23 (3.4)	6 (0.9)	12 (6.6)	18 (10.3)	16 (9.1)	8 (13.3)	16 (27.6)	13 (22.4)	28 (1.9)	68 (4.8)	37 (2.6)
eGFR all stages	566 [37.5]	520 [36.3]	520 [36.3]	702 [46.5]	678 [47.4]	678 [47.4]	181 [12.0]	175 [12.2]	175 [12.2]	60 [4.0]	58 [4.1]	58 [4.1]	1509 [100]	1431 [100]	1431 [100]
Correctly classified*	452 (79.9)	284 (54.6)	355 (68.3)	315 (44.9)	369 (54.4)	406 (59.9)	20 (11.0)	64 (36.6)	35 (20.0)	8 (13.3)	16 (27.6)	13 (22.4)	795 (52.7)	733 (51.2)	809 (56.5)
Stage as or more severe than mGFR	566 (100)	520 (100)	520 (100)	350 (49.9)	509 (75.1)	455 (67.1)	32 (17.7)	82 (46.9)	51 (29.1)	8 (13.3)	16 (27.6)	13 (22.4)	956 (63.4)	1127 (78.8)	1039 (72.6)
Absolute bias (mean: eGFR-mGFR)	-5.7	-17	-11	13.6	1.3	6.7	23.9	12.7	17.4	37.4	20.3	27.5	8.5	-3.1	2.5
Absolute bias (median: eGFR-mGFR)	-2.3	-14	-8.0	13.9	-0.7	5.8	23.1	7.6	12.7	36.7	16.4	24.4	9.2	-3.0	2.7
Relative bias (mean: eGFR/mGFR)	0.97	0.86	0.92	1.19	1.02	1.09	1.45	1.24	1.33	2.07	1.61	1.80	1.17	1.01	1.09
Relative bias (median: eGFR/mGFR)	0.98	0.86	0.92	1.18	0.99	1.07	1.43	1.14	1.23	1.96	1.40	1.61	1.11	0.96	1.03
Precision, RMSE	23.8	24.1	21.9	17.4	20.5	17.3	21.2	21.2	19.7	24.8	22.6	22.9	24.1	24.8	22.5
Precision, IQR	-17; 9.7	-31;-0.3	-22; 4.5	1.4;25.0	-13;15.0	-5.5;17.7	9.1;36.7	-1.2;23.6	3.9;30.5	19.6;56.2	4.7;36.1	10.8;47.7	-4.7;23.2	-18;11.3	-9.8;15.5
P10	0.42	0.33	0.40	0.26	0.29	0.35	0.14	0.30	0.21	0.05	0.09	0.09	0.30	0.30	0.34
P30	0.87	0.78	0.87	0.67	0.75	0.79	0.39	0.63	0.53	0.15	0.38	0.24	0.69	0.73	0.76
mGFR overestimated by 2+ ml/min	241 (42.6)	115 (22.1)	147 (28.3)	519 (73.9)	308 (45.4)	403 (59.4)	156 (86.2)	111 (63.4)	137 (78.3)	55 (91.7)	46 (79.3)	49 (84.5)	971 (64.3)	580 (40.5)	736 (51.4)
eGFR within +/- 2 ml/min of mGFR	40 (7.1)	34 (6.5)	36 (6.9)	49 (7.0)	50 (7.4)	57 (8.4)	8 (4.4)	23 (13.1)	14 (8.0)	1 (1.7)	3 (5.2)	4 (6.9)	98 (6.5)	110 (7.7)	111 (7.8)
mGFR underestimated by 2+ ml/min	285 (50.4)	371 (71.3)	337 (64.8)	134 (19.1)	320 (47.2)	218 (32.2)	17 (9.4)	41 (23.4)	24 (13.7)	4 (6.7)	9 (15.5)	5 (8.6)	440 (29.2)	741 (51.8)	584 (40.8)
mGFR overestimated by 5+ ml/min	204 (36.0)	97 (18.7)	122 (23.5)	481 (68.5)	262 (38.6)	354 (52.2)	145 (80.1)	99 (56.6)	125 (71.4)	54 (90.0)	42 (72.4)	47 (81.0)	884 (58.6)	500 (34.9)	648 (45.3)
eGFR within +/- 5 ml/min of mGFR	107 (18.9)	80 (15.4)	99 (19.0)	122 (17.4)	131 (19.3)	149 (22.0)	25 (13.8)	47 (26.9)	35 (20.0)	3 (5.0)	9 (15.5)	7 (12.1)	257 (17.0)	267 (18.7)	290 (20.3)
mGFR underestimated by 5+ ml/min	255 (45.1)	343 (66.0)	299 (57.5)	99 (14.1)	285 (42.0)	175 (25.8)	11 (6.1)	29 (16.6)	15 (8.6)	3 (5.0)	7 (12.1)	4 (6.9)	368 (24.4)	664 (46.4)	493 (34.5)
mGFR overestimated by 10+ ml/min	139 (24.6)	60 (11.5)	84 (16.2)	414 (59.0)	206 (30.4)	267 (39.4)	131 (72.4)	77 (44.0)	103 (58.9)	50 (83.3)	35 (60.3)	45 (77.6)	734 (48.6)	378 (26.4)	499 (34.9)
eGFR within +/- 10 ml/min of mGFR	230 (40.6)	164 (31.5)	202 (38.8)	230 (32.8)	270 (39.8)	303 (44.7)	44 (24.3)	86 (49.1)	63 (36.0)	10 (16.7)	19 (32.8)	11 (19.0)	514 (34.1)	539 (37.7)	579 (40.5)
mGFR underestimated by 10+ ml/min	197 (34.8)	296 (56.9)	234 (45.0)	58 (8.3)	202 (29.8)	108 (15.9)	6 (3.3)	12 (6.9)	9 (5.1)	0 (0.0)	4 (6.9)	2 (3.4)	261 (17.3)	514 (35.9)	353 (24.7)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
$ sensitivity analysis with any records excluded where r<0.985 in the iohexol-based slope-intercept mGFR derivation.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S13 (a) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	435 (76.9)	344 (60.8)	412 (72.8)	315 (44.9)	171 (24.4)	314 (44.7)	48 (26.5)	30 (16.6)	41 (22.7)	16 (26.7)	8 (13.3)	16 (26.7)	814 (53.9)	553 (36.6)	783 (51.9)
eGFR stage 2	119 (21.0)	207 (36.6)	137 (24.2)	350 (49.9)	483 (68.8)	325 (46.3)	93 (51.4)	108 (59.7)	86 (47.5)	25 (41.7)	33 (55.0)	23 (38.3)	587 (38.9)	831 (55.1)	571 (37.8)
eGFR stage 3a	8 (1.4)	11 (1.9)	13 (2.3)	31 (4.4)	43 (6.1)	57 (8.1)	30 (16.6)	29 (16.0)	36 (19.9)	11 (18.3)	10 (16.7)	10 (16.7)	80 (5.3)	93 (6.2)	116 (7.7)
eGFR stage 3b,4,5	4 (0.7)	4 (0.7)	4 (0.7)	6 (0.9)	5 (0.7)	6 (0.9)	10 (5.5)	14 (7.7)	18 (9.9)	8 (13.3)	9 (15.0)	11 (18.3)	28 (1.9)	32 (2.1)	39 (2.6)
eGFR all stages	566 [37.5]	566 [37.5]	566 [37.5]	702 [46.5]	702 [46.5]	702 [46.5]	181 [12.0]	181 [12.0]	181 [12.0]	60 [4.0]	60 [4.0]	60 [4.0]	1509 [100]	1509 [100]	1509 [100]
Correctly classified*	435 (76.9)	344 (60.8)	412 (72.8)	350 (49.9)	483 (68.8)	325 (46.3)	30 (16.6)	29 (16.0)	36 (19.9)	8 (13.3)	9 (15.0)	11 (18.3)	823 (54.5)	865 (57.3)	784 (52.0)
Stage as or more severe than mGFR	566 (100)	566 (100)	566 (100)	387 (55.1)	531 (75.6)	388 (55.3)	40 (22.1)	43 (23.8)	54 (29.8)	8 (13.3)	9 (15.0)	11 (18.3)	1001 (66.3)	1149 (76.1)	1019 (67.5)
Absolute bias (mean: eGFR-mGFR)	-3.6	-17	-1.3	14.1	4.9	14.4	22.8	16.7	21.7	36.7	30.5	37.6	9.4	-0.8	10.3
Absolute bias (median: eGFR-mGFR)	-1.7	-13	-2.1	12.4	5.3	11.9	18.4	15.4	15.3	32.0	29.3	33.3	9.3	0.5	8.9
Relative bias (mean: eGFR/mGFR)	0.99	0.86	1.01	1.19	1.07	1.20	1.43	1.32	1.41	2.06	1.88	2.10	1.18	1.05	1.19
Relative bias (median: eGFR/mGFR)	0.98	0.88	0.98	1.17	1.07	1.16	1.36	1.28	1.29	1.88	1.76	1.87	1.12	1.01	1.11
Precision, RMSE	28.8	22.0	32.6	20.6	14.8	25.5	23.2	18.1	29.2	26.1	21.8	33.1	26.9	23.0	30.8
Precision, IQR	-20;15.2	-26;-2.6	-21;21.1	0.8;26.2	-4.9;14.5	-2.1;28.1	5.7;38.3	5.0;27.5	2.1;34.4	16.3;55.1	16.5;48.0	13.2;59.5	-4.9;25.1	-12;12.5	-7.8;27.2
P10	0.32	0.36	0.28	0.27	0.39	0.26	0.18	0.18	0.22	0.07	0.08	0.08	0.27	0.34	0.25
P30	0.78	0.83	0.71	0.67	0.87	0.66	0.45	0.49	0.49	0.17	0.18	0.20	0.66	0.78	0.64
mGFR overestimated by 2+ ml/min	252 (44.5)	103 (18.2)	249 (44.0)	508 (72.4)	408 (58.1)	474 (67.5)	151 (83.4)	142 (78.5)	136 (75.1)	55 (91.7)	52 (86.7)	51 (85.0)	966 (64.0)	705 (46.7)	910 (60.3)
eGFR within +/- 2 ml/min of mGFR	33 (5.8)	32 (5.7)	34 (6.0)	48 (6.8)	68 (9.7)	51 (7.3)	14 (7.7)	13 (7.2)	8 (4.4)	3 (5.0)	4 (6.7)	4 (6.7)	98 (6.5)	117 (7.8)	97 (6.4)
mGFR underestimated by 2+ ml/min	281 (49.6)	431 (76.1)	283 (50.0)	146 (20.8)	226 (32.2)	177 (25.2)	16 (8.8)	26 (14.4)	37 (20.4)	2 (3.3)	4 (6.7)	5 (8.3)	445 (29.5)	687 (45.5)	502 (33.3)
mGFR overestimated by 5+ ml/min	226 (39.9)	78 (13.8)	232 (41.0)	463 (66.0)	358 (51.0)	427 (60.8)	139 (76.8)	136 (75.1)	121 (66.9)	53 (88.3)	50 (83.3)	51 (85.0)	881 (58.4)	622 (41.2)	831 (55.1)
eGFR within +/- 5 ml/min of mGFR	88 (15.5)	94 (16.6)	75 (13.3)	127 (18.1)	170 (24.2)	135 (19.2)	32 (17.7)	29 (16.0)	34 (18.8)	5 (8.3)	7 (11.7)	5 (8.3)	252 (16.7)	300 (19.9)	249 (16.5)
mGFR underestimated by 5+ ml/min	252 (44.5)	394 (69.6)	259 (45.8)	112 (16.0)	174 (24.8)	140 (19.9)	10 (5.5)	16 (8.8)	26 (14.4)	2 (3.3)	3 (5.0)	4 (6.7)	376 (24.9)	587 (38.9)	429 (28.4)
mGFR overestimated by 10+ ml/min	183 (32.3)	37 (6.5)	200 (35.3)	384 (54.7)	254 (36.2)	375 (53.4)	119 (65.7)	113 (62.4)	109 (60.2)	49 (81.7)	49 (81.7)	48 (80.0)	735 (48.7)	453 (30.0)	732 (48.5)
eGFR within +/- 10 ml/min of mGFR	177 (31.3)	196 (34.6)	146 (25.8)	251 (35.8)	344 (49.0)	229 (32.6)	56 (30.9)	60 (33.1)	58 (32.0)	11 (18.3)	9 (15.0)	9 (15.0)	495 (32.8)	609 (40.4)	442 (29.3)
mGFR underestimated by 10+ ml/min	206 (36.4)	333 (58.8)	220 (38.9)	67 (9.5)	104 (14.8)	98 (14.0)	6 (3.3)	8 (4.4)	14 (7.7)	0 (0.0)	2 (3.3)	3 (5.0)	279 (18.5)	447 (29.6)	335 (22.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
$ sensitivity analysis with any records excluded where r<0.985 in the iohexol-based slope-intercept mGFR derivation.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S13 (a) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	520 (91.9)	378 (66.8)	497 (87.8)	522 (74.4)	234 (33.3)	478 (68.1)	86 (47.5)	37 (20.4)	79 (43.6)	29 (48.3)	11 (18.3)	28 (46.7)	1157 (76.7)	660 (43.7)	1082 (71.7)
eGFR stage 2	41 (7.2)	171 (30.2)	62 (11.0)	169 (24.1)	407 (58.0)	211 (30.1)	74 (40.9)	102 (56.4)	83 (45.9)	18 (30.0)	31 (51.7)	22 (36.7)	302 (20.0)	711 (47.1)	378 (25.0)
eGFR stage 3a	2 (0.4)	13 (2.3)	4 (0.7)	8 (1.1)	57 (8.1)	10 (1.4)	14 (7.7)	29 (16.0)	13 (7.2)	7 (11.7)	11 (18.3)	5 (8.3)	31 (2.1)	110 (7.3)	32 (2.1)
eGFR stage 3b,4,5	3 (0.5)	4 (0.7)	3 (0.5)	3 (0.4)	4 (0.6)	3 (0.4)	7 (3.9)	13 (7.2)	6 (3.3)	6 (10.0)	7 (11.7)	5 (8.3)	19 (1.3)	28 (1.9)	17 (1.1)
eGFR all stages	566 [37.5]	566 [37.5]	566 [37.5]	702 [46.5]	702 [46.5]	702 [46.5]	181 [12.0]	181 [12.0]	181 [12.0]	60 [4.0]	60 [4.0]	60 [4.0]	1509 [100]	1509 [100]	1509 [100]
Correctly classified*	520 (91.9)	378 (66.8)	497 (87.8)	169 (24.1)	407 (58.0)	211 (30.1)	14 (7.7)	29 (16.0)	13 (7.2)	6 (10.0)	7 (11.7)	5 (8.3)	709 (47.0)	821 (54.4)	726 (48.1)
Stage as or more severe than mGFR	566 (100)	566 (100)	566 (100)	180 (25.6)	468 (66.7)	224 (31.9)	21 (11.6)	42 (23.2)	19 (10.5)	6 (10.0)	7 (11.7)	5 (8.3)	773 (51.2)	1083 (71.8)	814 (53.9)
Absolute bias (mean: eGFR-mGFR)	10.8	-9.3	12.0	27.8	8.4	26.3	36.2	20.2	35.8	49.3	32.9	47.9	23.3	4.2	22.9
Absolute bias (median: eGFR-mGFR)	14.7	-8.4	12.0	28.9	7.4	24.8	34.9	18.0	33.0	48.8	29.7	45.3	24.5	4.1	22.4
Relative bias (mean: eGFR/mGFR)	1.12	0.93	1.13	1.37	1.12	1.35	1.68	1.38	1.67	2.40	1.94	2.36	1.36	1.11	1.35
Relative bias (median: eGFR/mGFR)	1.14	0.92	1.11	1.37	1.10	1.33	1.65	1.33	1.61	2.27	1.79	2.17	1.28	1.05	1.27
Precision, RMSE	25.9	27.4	31.4	19.9	19.9	23.8	24.5	20.8	25.2	28.5	22.4	26.8	25.5	26.0	28.8
Precision, IQR	-2.7;27.8	-23; 7.3	-5.5;30.1	13.4;40.7	-5.0;19.7	10.4;40.0	19.7;51.2	5.6;31.3	18.3;49.0	27.6;70.7	17.6;51.2	26.9;69.3	9.1;38.8	-10;19.5	5.6;39.5
P10	0.28	0.32	0.29	0.13	0.30	0.17	0.05	0.17	0.04	0.07	0.03	0.07	0.17	0.28	0.20
P30	0.75	0.78	0.68	0.39	0.75	0.46	0.20	0.44	0.21	0.10	0.17	0.08	0.49	0.70	0.50
mGFR overestimated by 2+ ml/min	392 (69.3)	182 (32.2)	364 (64.3)	638 (90.9)	424 (60.4)	612 (87.2)	171 (94.5)	151 (83.4)	172 (95.0)	57 (95.0)	55 (91.7)	58 (96.7)	1258 (83.4)	812 (53.8)	1206 (79.9)
eGFR within +/- 2 ml/min of mGFR	31 (5.5)	35 (6.2)	32 (5.7)	19 (2.7)	54 (7.7)	23 (3.3)	4 (2.2)	11 (6.1)	2 (1.1)	2 (3.3)	1 (1.7)	2 (3.3)	56 (3.7)	101 (6.7)	59 (3.9)
mGFR underestimated by 2+ ml/min	143 (25.3)	349 (61.7)	170 (30.0)	45 (6.4)	224 (31.9)	67 (9.5)	6 (3.3)	19 (10.5)	7 (3.9)	1 (1.7)	4 (6.7)	0 (0.0)	195 (12.9)	596 (39.5)	244 (16.2)
mGFR overestimated by 5+ ml/min	371 (65.5)	161 (28.4)	340 (60.1)	615 (87.6)	382 (54.4)	582 (82.9)	166 (91.7)	139 (76.8)	167 (92.3)	56 (93.3)	55 (91.7)	56 (93.3)	1208 (80.1)	737 (48.8)	1145 (75.9)
eGFR within +/- 5 ml/min of mGFR	74 (13.1)	86 (15.2)	80 (14.1)	53 (7.5)	143 (20.4)	76 (10.8)	9 (5.0)	27 (14.9)	9 (5.0)	4 (6.7)	2 (3.3)	4 (6.7)	140 (9.3)	258 (17.1)	169 (11.2)
mGFR underestimated by 5+ ml/min	121 (21.4)	319 (56.4)	146 (25.8)	34 (4.8)	177 (25.2)	44 (6.3)	6 (3.3)	15 (8.3)	5 (2.8)	0 (0.0)	3 (5.0)	0 (0.0)	161 (10.7)	514 (34.1)	195 (12.9)
mGFR overestimated by 10+ ml/min	322 (56.9)	124 (21.9)	301 (53.2)	567 (80.8)	299 (42.6)	531 (75.6)	155 (85.6)	119 (65.7)	159 (87.8)	55 (91.7)	49 (81.7)	55 (91.7)	1099 (72.8)	591 (39.2)	1046 (69.3)
eGFR within +/- 10 ml/min of mGFR	144 (25.4)	172 (30.4)	152 (26.9)	115 (16.4)	295 (42.0)	148 (21.1)	23 (12.7)	56 (30.9)	19 (10.5)	5 (8.3)	11 (18.3)	5 (8.3)	287 (19.0)	534 (35.4)	324 (21.5)
mGFR underestimated by 10+ ml/min	100 (17.7)	270 (47.7)	113 (20.0)	20 (2.8)	108 (15.4)	23 (3.3)	3 (1.7)	6 (3.3)	3 (1.7)	0 (0.0)	0 (0.0)	0 (0.0)	123 (8.2)	384 (25.4)	139 (9.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
$ sensitivity analysis with any records excluded where r<0.985 in the iohexol-based slope-intercept mGFR derivation.
All results based on GFR values measured in ml/min/1.73m^2

Table S13 (b): Overall and by iohexol GFR category: performance of creatinine and cystatin C-based eGFR equations restricted for VD

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	216 (83.4)	124 (51.5)	163 (67.6)	283 (50.0)	140 (25.4)	186 (33.8)	34 (25.2)	15 (11.6)	17 (13.2)	8 (16.3)	4 (8.3)	6 (12.5)	541 (53.6)	283 (29.2)	372 (38.4)
eGFR stage 2	40 (15.4)	107 (44.4)	75 (31.1)	256 (45.2)	308 (55.9)	329 (59.7)	69 (51.1)	49 (38.0)	59 (45.7)	15 (30.6)	6 (12.5)	12 (25.0)	380 (37.7)	470 (48.5)	475 (49.0)
eGFR stage 3a	1 (0.4)	6 (2.5)	2 (0.8)	24 (4.2)	92 (16.7)	33 (6.0)	23 (17.0)	44 (34.1)	40 (31.0)	11 (22.4)	15 (31.3)	11 (22.9)	59 (5.8)	157 (16.2)	86 (8.9)
eGFR stage 3b,4,5	2 (0.8)	4 (1.7)	1 (0.4)	3 (0.5)	11 (2.0)	3 (0.5)	9 (6.7)	21 (16.3)	13 (10.1)	15 (30.6)	23 (47.9)	19 (39.6)	29 (2.9)	59 (6.1)	36 (3.7)
eGFR all stages	259 [25.7]	241 [24.9]	241 [24.9]	566 [56.1]	551 [56.9]	551 [56.9]	135 [13.4]	129 [13.3]	129 [13.3]	49 [4.9]	48 [5.0]	48 [5.0]	1009 [100]	969 [100]	969 [100]
Correctly classified*	216 (83.4)	124 (51.5)	163 (67.6)	256 (45.2)	308 (55.9)	329 (59.7)	23 (17.0)	44 (34.1)	40 (31.0)	15 (30.6)	23 (47.9)	19 (39.6)	510 (50.5)	499 (51.5)	551 (56.9)
Stage as or more severe than mGFR	259 (100)	241 (100)	241 (100)	283 (50.0)	411 (74.6)	365 (66.2)	32 (23.7)	65 (50.4)	53 (41.1)	15 (30.6)	23 (47.9)	19 (39.6)	589 (58.4)	740 (76.4)	678 (70.0)
Absolute bias (mean: eGFR-mGFR)	3.6	-10	-3.6	14.0	2.2	7.4	21.2	8.5	13.4	23.0	10.2	15.4	12.7	0.4	5.9
Absolute bias (median: eGFR-mGFR)	4.6	-11	-4.1	14.9	0.9	6.4	16.8	4.6	9.7	18.8	5.6	11.2	12.8	-0.8	4.9
Relative bias (mean: eGFR/mGFR)	1.04	0.90	0.97	1.19	1.03	1.10	1.40	1.16	1.25	1.65	1.30	1.44	1.20	1.03	1.11
Relative bias (median: eGFR/mGFR)	1.04	0.89	0.96	1.19	1.01	1.09	1.32	1.08	1.18	1.44	1.14	1.27	1.17	0.99	1.06
Precision, RMSE	18.1	19.2	17.2	16.8	19.3	16.0	21.2	20.5	19.2	26.4	23.1	23.9	19.3	20.7	18.2
Precision, IQR	-7.4;16.7	-22; 2.4	-14; 7.0	1.7;25.5	-11;15.1	-3.9;18.5	4.9;37.0	-3.5;16.2	-0.4;23.1	4.4;37.9	-5.7;11.0	-1.4;23.8	0.3;24.9	-13;12.8	-6.1;17.0
P10	0.45	0.34	0.46	0.25	0.29	0.35	0.19	0.32	0.26	0.12	0.19	0.15	0.29	0.30	0.36
P30	0.90	0.83	0.92	0.65	0.77	0.79	0.46	0.67	0.63	0.33	0.58	0.46	0.67	0.76	0.79
mGFR overestimated by 2+ ml/min	150 (57.9)	62 (25.7)	87 (36.1)	421 (74.4)	264 (47.9)	334 (60.6)	113 (83.7)	72 (55.8)	90 (69.8)	37 (75.5)	29 (60.4)	31 (64.6)	721 (71.5)	427 (44.1)	542 (55.9)
eGFR within +/- 2 ml/min of mGFR	22 (8.5)	15 (6.2)	23 (9.5)	41 (7.2)	42 (7.6)	63 (11.4)	7 (5.2)	20 (15.5)	14 (10.9)	4 (8.2)	5 (10.4)	6 (12.5)	74 (7.3)	82 (8.5)	106 (10.9)
mGFR underestimated by 2+ ml/min	87 (33.6)	164 (68.0)	131 (54.4)	104 (18.4)	245 (44.5)	154 (27.9)	15 (11.1)	37 (28.7)	25 (19.4)	8 (16.3)	14 (29.2)	11 (22.9)	214 (21.2)	460 (47.5)	321 (33.1)
mGFR overestimated by 5+ ml/min	127 (49.0)	55 (22.8)	72 (29.9)	397 (70.1)	225 (40.8)	296 (53.7)	100 (74.1)	62 (48.1)	82 (63.6)	36 (73.5)	25 (52.1)	30 (62.5)	660 (65.4)	367 (37.9)	480 (49.5)
eGFR within +/- 5 ml/min of mGFR	55 (21.2)	36 (14.9)	53 (22.0)	91 (16.1)	105 (19.1)	128 (23.2)	25 (18.5)	39 (30.2)	32 (24.8)	6 (12.2)	11 (22.9)	8 (16.7)	177 (17.5)	191 (19.7)	221 (22.8)
mGFR underestimated by 5+ ml/min	77 (29.7)	150 (62.2)	116 (48.1)	78 (13.8)	221 (40.1)	127 (23.0)	10 (7.4)	28 (21.7)	15 (11.6)	7 (14.3)	12 (25.0)	10 (20.8)	172 (17.0)	411 (42.4)	268 (27.7)
mGFR overestimated by 10+ ml/min	94 (36.3)	39 (16.2)	52 (21.6)	345 (61.0)	179 (32.5)	235 (42.6)	90 (66.7)	48 (37.2)	62 (48.1)	31 (63.3)	14 (29.2)	26 (54.2)	560 (55.5)	280 (28.9)	375 (38.7)
eGFR within +/- 10 ml/min of mGFR	114 (44.0)	76 (31.5)	113 (46.9)	174 (30.7)	224 (40.7)	249 (45.2)	40 (29.6)	64 (49.6)	59 (45.7)	14 (28.6)	24 (50.0)	16 (33.3)	342 (33.9)	388 (40.0)	437 (45.1)
mGFR underestimated by 10+ ml/min	51 (19.7)	126 (52.3)	76 (31.5)	47 (8.3)	148 (26.9)	67 (12.2)	5 (3.7)	17 (13.2)	8 (6.2)	4 (8.2)	10 (20.8)	6 (12.5)	107 (10.6)	301 (31.1)	157 (16.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
$ sensitivity analysis with participants excluded if the calculated volume of distribution (based on the S-I method) is outside the interval [11-17L] for women and [13-20L] for men.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S13 (b) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	198 (76.4)	157 (60.6)	189 (73.0)	256 (45.2)	142 (25.1)	244 (43.1)	32 (23.7)	20 (14.8)	32 (23.7)	7 (14.3)	5 (10.2)	8 (16.3)	493 (48.9)	324 (32.1)	473 (46.9)
eGFR stage 2	57 (22.0)	99 (38.2)	65 (25.1)	284 (50.2)	386 (68.2)	269 (47.5)	68 (50.4)	73 (54.1)	60 (44.4)	13 (26.5)	16 (32.7)	13 (26.5)	422 (41.8)	574 (56.9)	407 (40.3)
eGFR stage 3a	2 (0.8)	1 (0.4)	3 (1.2)	22 (3.9)	35 (6.2)	48 (8.5)	27 (20.0)	32 (23.7)	26 (19.3)	14 (28.6)	12 (24.5)	8 (16.3)	65 (6.4)	80 (7.9)	85 (8.4)
eGFR stage 3b,4,5	2 (0.8)	2 (0.8)	2 (0.8)	4 (0.7)	3 (0.5)	5 (0.9)	8 (5.9)	10 (7.4)	17 (12.6)	15 (30.6)	16 (32.7)	20 (40.8)	29 (2.9)	31 (3.1)	44 (4.4)
eGFR all stages	259 [25.7]	259 [25.7]	259 [25.7]	566 [56.1]	566 [56.1]	566 [56.1]	135 [13.4]	135 [13.4]	135 [13.4]	49 [4.9]	49 [4.9]	49 [4.9]	1009 [100]	1009 [100]	1009 [100]
Correctly classified*	198 (76.4)	157 (60.6)	189 (73.0)	284 (50.2)	386 (68.2)	269 (47.5)	27 (20.0)	32 (23.7)	26 (19.3)	15 (30.6)	16 (32.7)	20 (40.8)	524 (51.9)	591 (58.6)	504 (50.0)
Stage as or more severe than mGFR	259 (100)	259 (100)	259 (100)	310 (54.8)	424 (74.9)	322 (56.9)	35 (25.9)	42 (31.1)	43 (31.9)	15 (30.6)	16 (32.7)	20 (40.8)	619 (61.3)	741 (73.4)	644 (63.8)
Absolute bias (mean: eGFR-mGFR)	5.6	-8.1	4.6	14.7	5.4	14.0	20.1	14.3	20.1	22.2	17.4	21.2	13.4	3.7	12.7
Absolute bias (median: eGFR-mGFR)	5.7	-7.4	1.4	14.1	5.9	11.3	16.4	11.9	15.7	15.0	16.3	14.4	12.2	4.1	10.0
Relative bias (mean: eGFR/mGFR)	1.06	0.93	1.05	1.20	1.08	1.19	1.37	1.27	1.37	1.64	1.50	1.63	1.21	1.09	1.20
Relative bias (median: eGFR/mGFR)	1.06	0.93	1.01	1.18	1.08	1.15	1.30	1.22	1.29	1.36	1.40	1.38	1.17	1.05	1.13
Precision, RMSE	24.3	16.3	27.0	19.9	14.5	25.4	21.2	17.4	27.2	26.8	23.1	31.9	22.2	17.7	26.9
Precision, IQR	-9.7;22.0	-18; 3.1	-13;21.1	1.5;26.2	-4.9;15.0	-2.3;28.2	4.7;33.7	1.2;28.0	-1.1;36.5	3.9;31.9	-0.6;28.3	-1.0;32.9	-0.5;25.8	-7.2;14.1	-4.3;26.9
P10	0.34	0.45	0.32	0.26	0.38	0.27	0.19	0.23	0.20	0.14	0.16	0.22	0.27	0.37	0.27
P30	0.81	0.95	0.79	0.66	0.86	0.66	0.47	0.57	0.48	0.41	0.41	0.39	0.66	0.82	0.66
mGFR overestimated by 2+ ml/min	146 (56.4)	74 (28.6)	126 (48.6)	419 (74.0)	342 (60.4)	370 (65.4)	108 (80.0)	99 (73.3)	99 (73.3)	38 (77.6)	35 (71.4)	32 (65.3)	711 (70.5)	550 (54.5)	627 (62.1)
eGFR within +/- 2 ml/min of mGFR	16 (6.2)	13 (5.0)	17 (6.6)	39 (6.9)	56 (9.9)	50 (8.8)	13 (9.6)	14 (10.4)	5 (3.7)	5 (10.2)	5 (10.2)	8 (16.3)	73 (7.2)	88 (8.7)	80 (7.9)
mGFR underestimated by 2+ ml/min	97 (37.5)	172 (66.4)	116 (44.8)	108 (19.1)	168 (29.7)	146 (25.8)	14 (10.4)	22 (16.3)	31 (23.0)	6 (12.2)	9 (18.4)	9 (18.4)	225 (22.3)	371 (36.8)	302 (29.9)
mGFR overestimated by 5+ ml/min	135 (52.1)	58 (22.4)	117 (45.2)	385 (68.0)	302 (53.4)	339 (59.9)	101 (74.8)	95 (70.4)	91 (67.4)	34 (69.4)	33 (67.3)	30 (61.2)	655 (64.9)	488 (48.4)	577 (57.2)
eGFR within +/- 5 ml/min of mGFR	37 (14.3)	53 (20.5)	43 (16.6)	102 (18.0)	126 (22.3)	111 (19.6)	24 (17.8)	27 (20.0)	23 (17.0)	10 (20.4)	8 (16.3)	11 (22.4)	173 (17.1)	214 (21.2)	188 (18.6)
mGFR underestimated by 5+ ml/min	87 (33.6)	148 (57.1)	99 (38.2)	79 (14.0)	138 (24.4)	116 (20.5)	10 (7.4)	13 (9.6)	21 (15.6)	5 (10.2)	8 (16.3)	8 (16.3)	181 (17.9)	307 (30.4)	244 (24.2)
mGFR overestimated by 10+ ml/min	106 (40.9)	33 (12.7)	103 (39.8)	324 (57.2)	216 (38.2)	296 (52.3)	88 (65.2)	76 (56.3)	81 (60.0)	30 (61.2)	29 (59.2)	26 (53.1)	548 (54.3)	354 (35.1)	506 (50.1)
eGFR within +/- 10 ml/min of mGFR	90 (34.7)	115 (44.4)	81 (31.3)	193 (34.1)	267 (47.2)	188 (33.2)	42 (31.1)	53 (39.3)	42 (31.1)	15 (30.6)	15 (30.6)	17 (34.7)	340 (33.7)	450 (44.6)	328 (32.5)
mGFR underestimated by 10+ ml/min	63 (24.3)	111 (42.9)	75 (29.0)	49 (8.7)	83 (14.7)	82 (14.5)	5 (3.7)	6 (4.4)	12 (8.9)	4 (8.2)	5 (10.2)	6 (12.2)	121 (12.0)	205 (20.3)	175 (17.3)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
$ sensitivity analysis with participants excluded if the calculated volume of distribution (based on the S-I method) is outside the interval [11-17L] for women and [13-20L] for men.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S13 (b) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	245 (94.6)	176 (68.0)	237 (91.5)	417 (73.7)	189 (33.4)	380 (67.1)	56 (41.5)	24 (17.8)	52 (38.5)	14 (28.6)	6 (12.2)	12 (24.5)	732 (72.5)	395 (39.1)	681 (67.5)
eGFR stage 2	12 (4.6)	79 (30.5)	19 (7.3)	143 (25.3)	331 (58.5)	179 (31.6)	59 (43.7)	72 (53.3)	65 (48.1)	14 (28.6)	15 (30.6)	21 (42.9)	228 (22.6)	497 (49.3)	284 (28.1)
eGFR stage 3a	1 (0.4)	2 (0.8)	2 (0.8)	4 (0.7)	43 (7.6)	5 (0.9)	16 (11.9)	27 (20.0)	16 (11.9)	10 (20.4)	14 (28.6)	6 (12.2)	31 (3.1)	86 (8.5)	29 (2.9)
eGFR stage 3b,4,5	1 (0.4)	2 (0.8)	1 (0.4)	2 (0.4)	3 (0.5)	2 (0.4)	4 (3.0)	12 (8.9)	2 (1.5)	11 (22.4)	14 (28.6)	10 (20.4)	18 (1.8)	31 (3.1)	15 (1.5)
eGFR all stages	259 [25.7]	259 [25.7]	259 [25.7]	566 [56.1]	566 [56.1]	566 [56.1]	135 [13.4]	135 [13.4]	135 [13.4]	49 [4.9]	49 [4.9]	49 [4.9]	1009 [100]	1009 [100]	1009 [100]
Correctly classified*	245 (94.6)	176 (68.0)	237 (91.5)	143 (25.3)	331 (58.5)	179 (31.6)	16 (11.9)	27 (20.0)	16 (11.9)	11 (22.4)	14 (28.6)	10 (20.4)	415 (41.1)	548 (54.3)	442 (43.8)
Stage as or more severe than mGFR	259 (100)	259 (100)	259 (100)	149 (26.3)	377 (66.6)	186 (32.9)	20 (14.8)	39 (28.9)	18 (13.3)	11 (22.4)	14 (28.6)	10 (20.4)	439 (43.5)	689 (68.3)	473 (46.9)
Absolute bias (mean: eGFR-mGFR)	20.4	0.1	21.7	28.2	9.0	26.9	33.2	16.8	31.9	32.7	22.3	35.0	27.1	8.4	26.6
Absolute bias (median: eGFR-mGFR)	21.3	-1.6	19.8	29.3	7.6	25.2	28.1	13.8	28.0	26.5	18.2	28.4	27.0	7.2	25.1
Relative bias (mean: eGFR/mGFR)	1.21	1.01	1.22	1.38	1.13	1.36	1.62	1.32	1.59	1.91	1.64	1.98	1.39	1.15	1.39
Relative bias (median: eGFR/mGFR)	1.21	0.98	1.19	1.38	1.10	1.34	1.53	1.26	1.53	1.67	1.43	1.73	1.35	1.10	1.33
Precision, RMSE	20.4	23.1	27.5	19.3	19.8	23.8	24.6	19.2	23.2	30.5	28.6	34.5	21.4	21.9	25.5
Precision, IQR	8.2;34.4	-17;12.8	2.2;36.4	13.7;41.7	-4.5;20.0	11.3;40.3	14.7;51.3	3.0;29.0	15.3;47.1	12.2;48.8	7.4;34.0	18.4;47.7	12.3;40.9	-6.1;20.7	9.0;40.8
P10	0.23	0.34	0.29	0.12	0.30	0.16	0.07	0.20	0.07	0.06	0.06	0.10	0.14	0.28	0.18
P30	0.66	0.83	0.63	0.37	0.75	0.45	0.25	0.54	0.27	0.24	0.35	0.20	0.42	0.72	0.46
mGFR overestimated by 2+ ml/min	216 (83.4)	115 (44.4)	198 (76.4)	514 (90.8)	353 (62.4)	498 (88.0)	126 (93.3)	106 (78.5)	127 (94.1)	41 (83.7)	38 (77.6)	42 (85.7)	897 (88.9)	612 (60.7)	865 (85.7)
eGFR within +/- 2 ml/min of mGFR	11 (4.2)	16 (6.2)	12 (4.6)	17 (3.0)	46 (8.1)	20 (3.5)	4 (3.0)	11 (8.1)	3 (2.2)	3 (6.1)	2 (4.1)	4 (8.2)	35 (3.5)	75 (7.4)	39 (3.9)
mGFR underestimated by 2+ ml/min	32 (12.4)	128 (49.4)	49 (18.9)	35 (6.2)	167 (29.5)	48 (8.5)	5 (3.7)	18 (13.3)	5 (3.7)	5 (10.2)	9 (18.4)	3 (6.1)	77 (7.6)	322 (31.9)	105 (10.4)
mGFR overestimated by 5+ ml/min	206 (79.5)	102 (39.4)	185 (71.4)	497 (87.8)	313 (55.3)	472 (83.4)	122 (90.4)	95 (70.4)	121 (89.6)	39 (79.6)	37 (75.5)	40 (81.6)	864 (85.6)	547 (54.2)	818 (81.1)
eGFR within +/- 5 ml/min of mGFR	33 (12.7)	42 (16.2)	38 (14.7)	44 (7.8)	116 (20.5)	62 (11.0)	8 (5.9)	25 (18.5)	9 (6.7)	6 (12.2)	5 (10.2)	6 (12.2)	91 (9.0)	188 (18.6)	115 (11.4)
mGFR underestimated by 5+ ml/min	20 (7.7)	115 (44.4)	36 (13.9)	25 (4.4)	137 (24.2)	32 (5.7)	5 (3.7)	15 (11.1)	5 (3.7)	4 (8.2)	7 (14.3)	3 (6.1)	54 (5.4)	274 (27.2)	76 (7.5)
mGFR overestimated by 10+ ml/min	188 (72.6)	81 (31.3)	164 (63.3)	458 (80.9)	248 (43.8)	432 (76.3)	112 (83.0)	81 (60.0)	113 (83.7)	37 (75.5)	29 (59.2)	37 (75.5)	795 (78.8)	439 (43.5)	746 (73.9)
eGFR within +/- 10 ml/min of mGFR	56 (21.6)	85 (32.8)	75 (29.0)	95 (16.8)	233 (41.2)	117 (20.7)	20 (14.8)	48 (35.6)	19 (14.1)	10 (20.4)	17 (34.7)	11 (22.4)	181 (17.9)	383 (38.0)	222 (22.0)
mGFR underestimated by 10+ ml/min	15 (5.8)	93 (35.9)	20 (7.7)	13 (2.3)	85 (15.0)	17 (3.0)	3 (2.2)	6 (4.4)	3 (2.2)	2 (4.1)	3 (6.1)	1 (2.0)	33 (3.3)	187 (18.5)	41 (4.1)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
$ sensitivity analysis with participants excluded if the calculated volume of distribution (based on the S-I method) is outside the interval [11-17L] for women and [13-20L] for men.
All results based on GFR values measured in ml/min/1.73m^2

Table S13 (c): Overall and by iohexol GFR category: Performance of creatinine and cystatin C-based eGFR equations restricted for r>0.985 and VD

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC	CKD-EPI creat only	CKD-EPI cystatin	CKD-EPI Creat+CysC
eGFR stage 1	154 (83.2)	90 (52.3)	118 (68.6)	182 (47.2)	91 (24.1)	120 (31.8)	17 (21.5)	8 (10.4)	8 (10.4)	0 (0.0)	1 (6.7)	1 (6.7)	353 (53.1)	190 (29.6)	247 (38.5)
eGFR stage 2	28 (15.1)	76 (44.2)	52 (30.2)	187 (48.4)	216 (57.3)	231 (61.3)	39 (49.4)	28 (36.4)	35 (45.5)	7 (46.7)	2 (13.3)	3 (20.0)	261 (39.2)	322 (50.2)	321 (50.1)
eGFR stage 3a	1 (0.5)	3 (1.7)	1 (0.6)	15 (3.9)	60 (15.9)	23 (6.1)	16 (20.3)	30 (39.0)	23 (29.9)	3 (20.0)	4 (26.7)	4 (26.7)	35 (5.3)	97 (15.1)	51 (8.0)
eGFR stage 3b,4,5	2 (1.1)	3 (1.7)	1 (0.6)	2 (0.5)	10 (2.7)	3 (0.8)	7 (8.9)	11 (14.3)	11 (14.3)	5 (33.3)	8 (53.3)	7 (46.7)	16 (2.4)	32 (5.0)	22 (3.4)
eGFR all stages	185 [27.8]	172 [26.8]	172 [26.8]	386 [58.0]	377 [58.8]	377 [58.8]	79 [11.9]	77 [12.0]	77 [12.0]	15 [2.3]	15 [2.3]	15 [2.3]	665 [100]	641 [100]	641 [100]
Correctly classified*	154 (83.2)	90 (52.3)	118 (68.6)	187 (48.4)	216 (57.3)	231 (61.3)	16 (20.3)	30 (39.0)	23 (29.9)	5 (33.3)	8 (53.3)	7 (46.7)	362 (54.4)	344 (53.7)	379 (59.1)
Stage as or more severe than mGFR	185 (100)	172 (100)	172 (100)	204 (52.8)	286 (75.9)	257 (68.2)	23 (29.1)	41 (53.2)	34 (44.2)	5 (33.3)	8 (53.3)	7 (46.7)	417 (62.7)	507 (79.1)	470 (73.3)
Absolute bias (mean: eGFR-mGFR)	4.3	-10	-3.5	12.7	1.4	6.3	18.6	8.0	12.1	16.9	6.3	10.6	11.1	-0.8	4.5
Absolute bias (median: eGFR-mGFR)	4.9	-10	-4.8	13.2	0.2	5.8	15.0	3.7	9.4	18.8	4.7	10.6	11.4	-1.4	4.3
Relative bias (mean: eGFR/mGFR)	1.05	0.90	0.97	1.17	1.02	1.09	1.35	1.15	1.23	1.52	1.22	1.34	1.17	1.01	1.08
Relative bias (median: eGFR/mGFR)	1.04	0.90	0.95	1.17	1.00	1.08	1.26	1.07	1.17	1.44	1.11	1.25	1.15	0.98	1.05
Precision, RMSE	18.1	19.5	17.1	16.6	19.2	15.8	21.9	20.7	19.7	16.5	19.0	17.3	18.3	20.4	17.4
Precision, IQR	-5.9;17.0	-22; 4.1	-15; 8.6	0.5;24.1	-12;13.5	-4.0;17.1	2.3;30.5	-3.5;16.2	-0.4;22.3	4.4;27.1	-7.5;11.4	-1.9;19.7	-0.5;23.2	-14;11.3	-6.9;14.7
P10	0.46	0.35	0.44	0.26	0.30	0.37	0.25	0.35	0.25	0.13	0.20	0.27	0.31	0.32	0.37
P30	0.90	0.83	0.92	0.68	0.78	0.81	0.52	0.69	0.65	0.40	0.53	0.60	0.71	0.78	0.81
mGFR overestimated by 2+ ml/min	109 (58.9)	48 (27.9)	66 (38.4)	279 (72.3)	175 (46.4)	224 (59.4)	62 (78.5)	41 (53.2)	51 (66.2)	12 (80.0)	8 (53.3)	9 (60.0)	462 (69.5)	272 (42.4)	350 (54.6)
eGFR within +/- 2 ml/min of mGFR	15 (8.1)	10 (5.8)	13 (7.6)	30 (7.8)	31 (8.2)	44 (11.7)	5 (6.3)	13 (16.9)	9 (11.7)	1 (6.7)	2 (13.3)	3 (20.0)	51 (7.7)	56 (8.7)	69 (10.8)
mGFR underestimated by 2+ ml/min	61 (33.0)	114 (66.3)	93 (54.1)	77 (19.9)	171 (45.4)	109 (28.9)	12 (15.2)	23 (29.9)	17 (22.1)	2 (13.3)	5 (33.3)	3 (20.0)	152 (22.9)	313 (48.8)	222 (34.6)
mGFR overestimated by 5+ ml/min	92 (49.7)	42 (24.4)	57 (33.1)	260 (67.4)	148 (39.3)	195 (51.7)	53 (67.1)	35 (45.5)	47 (61.0)	11 (73.3)	6 (40.0)	9 (60.0)	416 (62.6)	231 (36.0)	308 (48.0)
eGFR within +/- 5 ml/min of mGFR	41 (22.2)	27 (15.7)	29 (16.9)	71 (18.4)	74 (19.6)	91 (24.1)	19 (24.1)	25 (32.5)	19 (24.7)	2 (13.3)	5 (33.3)	4 (26.7)	133 (20.0)	131 (20.4)	143 (22.3)
mGFR underestimated by 5+ ml/min	52 (28.1)	103 (59.9)	86 (50.0)	55 (14.2)	155 (41.1)	91 (24.1)	7 (8.9)	17 (22.1)	11 (14.3)	2 (13.3)	4 (26.7)	2 (13.3)	116 (17.4)	279 (43.5)	190 (29.6)
mGFR overestimated by 10+ ml/min	66 (35.7)	27 (15.7)	40 (23.3)	224 (58.0)	113 (30.0)	149 (39.5)	47 (59.5)	28 (36.4)	34 (44.2)	9 (60.0)	5 (33.3)	8 (53.3)	346 (52.0)	173 (27.0)	231 (36.0)
eGFR within +/- 10 ml/min of mGFR	84 (45.4)	57 (33.1)	76 (44.2)	126 (32.6)	160 (42.4)	177 (46.9)	29 (36.7)	40 (51.9)	36 (46.8)	6 (40.0)	7 (46.7)	6 (40.0)	245 (36.8)	264 (41.2)	295 (46.0)
mGFR underestimated by 10+ ml/min	35 (18.9)	88 (51.2)	56 (32.6)	36 (9.3)	104 (27.6)	51 (13.5)	3 (3.8)	9 (11.7)	7 (9.1)	0 (0.0)	3 (20.0)	1 (6.7)	74 (11.1)	204 (31.8)	115 (17.9)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

(table continued on next page)

Table S13 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault	FAS creat only	Lund-Malmo	Cockgroft-Gault
eGFR stage 1	148 (80.0)	118 (63.8)	137 (74.1)	165 (42.7)	84 (21.8)	159 (41.2)	18 (22.8)	10 (12.7)	13 (16.5)	0 (0.0)	0 (0.0)	0 (0.0)	331 (49.8)	212 (31.9)	309 (46.5)
eGFR stage 2	35 (18.9)	64 (34.6)	44 (23.8)	205 (53.1)	278 (72.0)	192 (49.7)	35 (44.3)	40 (50.6)	35 (44.3)	5 (33.3)	6 (40.0)	5 (33.3)	280 (42.1)	388 (58.3)	276 (41.5)
eGFR stage 3a	0 (0.0)	1 (0.5)	2 (1.1)	13 (3.4)	22 (5.7)	33 (8.5)	20 (25.3)	21 (26.6)	19 (24.1)	5 (33.3)	3 (20.0)	3 (20.0)	38 (5.7)	47 (7.1)	57 (8.6)
eGFR stage 3b,4,5	2 (1.1)	2 (1.1)	2 (1.1)	3 (0.8)	2 (0.5)	2 (0.5)	6 (7.6)	8 (10.1)	12 (15.2)	5 (33.3)	6 (40.0)	7 (46.7)	16 (2.4)	18 (2.7)	23 (3.5)
eGFR all stages	185 [27.8]	185 [27.8]	185 [27.8]	386 [58.0]	386 [58.0]	386 [58.0]	79 [11.9]	79 [11.9]	79 [11.9]	15 [2.3]	15 [2.3]	15 [2.3]	665 [100]	665 [100]	665 [100]
Correctly classified*	148 (80.0)	118 (63.8)	137 (74.1)	205 (53.1)	278 (72.0)	192 (49.7)	20 (25.3)	21 (26.6)	19 (24.1)	5 (33.3)	6 (40.0)	7 (46.7)	378 (56.8)	423 (63.6)	355 (53.4)
Stage as or more severe than mGFR	185 (100)	185 (100)	185 (100)	221 (57.3)	302 (78.2)	227 (58.8)	26 (32.9)	29 (36.7)	31 (39.2)	5 (33.3)	6 (40.0)	7 (46.7)	437 (65.7)	522 (78.5)	450 (67.7)
Absolute bias (mean: eGFR-mGFR)	6.8	-7.6	5.5	13.5	4.3	12.5	17.4	11.9	16.2	15.6	12.1	9.6	12.2	2.1	10.9
Absolute bias (median: eGFR-mGFR)	5.8	-6.8	2.1	12.4	5.2	10.5	14.5	10.0	11.2	14.8	16.3	6.8	10.7	2.5	8.4
Relative bias (mean: eGFR/mGFR)	1.07	0.93	1.06	1.19	1.06	1.17	1.32	1.22	1.30	1.50	1.39	1.40	1.18	1.05	1.16
Relative bias (median: eGFR/mGFR)	1.06	0.93	1.02	1.17	1.07	1.14	1.27	1.18	1.20	1.36	1.40	1.16	1.14	1.03	1.11
Precision, RMSE	23.6	16.1	27.6	19.3	14.1	24.3	21.5	17.8	27.1	15.0	14.6	16.2	21.0	16.5	25.7
Precision, IQR	-8.8;23.3	-16; 4.1	-13;23.9	1.0;25.4	-5.4;13.8	-2.9;26.2	2.1;29.5	-0.7;22.8	-2.6;28.0	3.9;29.9	-0.6;23.7	-2.5;20.7	-0.9;24.7	-7.7;12.2	-5.2;25.7
P10	0.35	0.46	0.30	0.27	0.41	0.27	0.24	0.27	0.25	0.20	0.27	0.27	0.29	0.40	0.28
P30	0.81	0.97	0.78	0.68	0.89	0.69	0.53	0.62	0.56	0.47	0.47	0.53	0.69	0.87	0.70
mGFR overestimated by 2+ ml/min	108 (58.4)	55 (29.7)	94 (50.8)	284 (73.6)	225 (58.3)	247 (64.0)	60 (75.9)	52 (65.8)	52 (65.8)	12 (80.0)	10 (66.7)	8 (53.3)	464 (69.8)	342 (51.4)	401 (60.3)
eGFR within +/- 2 ml/min of mGFR	10 (5.4)	9 (4.9)	8 (4.3)	26 (6.7)	40 (10.4)	37 (9.6)	8 (10.1)	10 (12.7)	5 (6.3)	2 (13.3)	3 (20.0)	3 (20.0)	46 (6.9)	62 (9.3)	53 (8.0)
mGFR underestimated by 2+ ml/min	67 (36.2)	121 (65.4)	83 (44.9)	76 (19.7)	121 (31.3)	102 (26.4)	11 (13.9)	17 (21.5)	22 (27.8)	1 (6.7)	2 (13.3)	4 (26.7)	155 (23.3)	261 (39.2)	211 (31.7)
mGFR overestimated by 5+ ml/min	99 (53.5)	44 (23.8)	87 (47.0)	259 (67.1)	195 (50.5)	225 (58.3)	54 (68.4)	50 (63.3)	47 (59.5)	10 (66.7)	9 (60.0)	8 (53.3)	422 (63.5)	298 (44.8)	367 (55.2)
eGFR within +/- 5 ml/min of mGFR	26 (14.1)	39 (21.1)	27 (14.6)	71 (18.4)	92 (23.8)	81 (21.0)	18 (22.8)	19 (24.1)	17 (21.5)	4 (26.7)	4 (26.7)	4 (26.7)	119 (17.9)	154 (23.2)	129 (19.4)
mGFR underestimated by 5+ ml/min	60 (32.4)	102 (55.1)	71 (38.4)	56 (14.5)	99 (25.6)	80 (20.7)	7 (8.9)	10 (12.7)	15 (19.0)	1 (6.7)	2 (13.3)	3 (20.0)	124 (18.6)	213 (32.0)	169 (25.4)
mGFR overestimated by 10+ ml/min	78 (42.2)	25 (13.5)	76 (41.1)	210 (54.4)	135 (35.0)	196 (50.8)	45 (57.0)	39 (49.4)	42 (53.2)	8 (53.3)	8 (53.3)	7 (46.7)	341 (51.3)	207 (31.1)	321 (48.3)
eGFR within +/- 10 ml/min of mGFR	65 (35.1)	84 (45.4)	55 (29.7)	141 (36.5)	192 (49.7)	133 (34.5)	30 (38.0)	36 (45.6)	29 (36.7)	7 (46.7)	6 (40.0)	6 (40.0)	243 (36.5)	318 (47.8)	223 (33.5)
mGFR underestimated by 10+ ml/min	42 (22.7)	76 (41.1)	54 (29.2)	35 (9.1)	59 (15.3)	57 (14.8)	4 (5.1)	4 (5.1)	8 (10.1)	0 (0.0)	1 (6.7)	2 (13.3)	81 (12.2)	140 (21.1)	121 (18.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

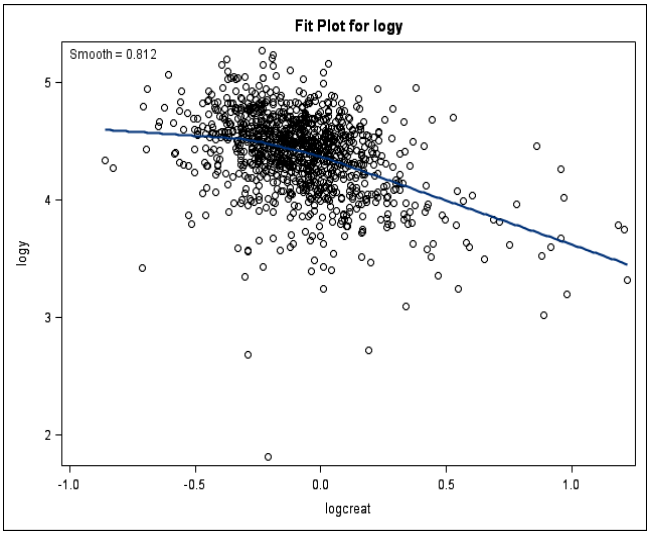
(table continued on next page)

Table S13 (c) (cont.)

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity	CKD-EPI w/ ethnicity	MDRD	MDRD w/ ethnicity
eGFR stage 1	176 (95.1)	129 (69.7)	170 (91.9)	279 (72.3)	118 (30.6)	250 (64.8)	30 (38.0)	13 (16.5)	27 (34.2)	2 (13.3)	0 (0.0)	2 (13.3)	487 (73.2)	260 (39.1)	449 (67.5)
eGFR stage 2	7 (3.8)	53 (28.6)	12 (6.5)	102 (26.4)	235 (60.9)	130 (33.7)	34 (43.0)	38 (48.1)	39 (49.4)	6 (40.0)	5 (33.3)	8 (53.3)	149 (22.4)	331 (49.8)	189 (28.4)
eGFR stage 3a	1 (0.5)	1 (0.5)	2 (1.1)	4 (1.0)	31 (8.0)	5 (1.3)	12 (15.2)	19 (24.1)	11 (13.9)	4 (26.7)	6 (40.0)	3 (20.0)	21 (3.2)	57 (8.6)	21 (3.2)
eGFR stage 3b,4,5	1 (0.5)	2 (1.1)	1 (0.5)	1 (0.3)	2 (0.5)	1 (0.3)	3 (3.8)	9 (11.4)	2 (2.5)	3 (20.0)	4 (26.7)	2 (13.3)	8 (1.2)	17 (2.6)	6 (0.9)
eGFR all stages	185 [27.8]	185 [27.8]	185 [27.8]	386 [58.0]	386 [58.0]	386 [58.0]	79 [11.9]	79 [11.9]	79 [11.9]	15 [2.3]	15 [2.3]	15 [2.3]	665 [100]	665 [100]	665 [100]
Correctly classified*	176 (95.1)	129 (69.7)	170 (91.9)	102 (26.4)	235 (60.9)	130 (33.7)	12 (15.2)	19 (24.1)	11 (13.9)	3 (20.0)	4 (26.7)	2 (13.3)	293 (44.1)	387 (58.2)	313 (47.1)
Stage as or more severe than mGFR	185 (100)	185 (100)	185 (100)	107 (27.7)	268 (69.4)	136 (35.2)	15 (19.0)	28 (35.4)	13 (16.5)	3 (20.0)	4 (26.7)	2 (13.3)	310 (46.6)	485 (72.9)	336 (50.5)
Absolute bias (mean: eGFR-mGFR)	21.0	1.0	22.6	26.8	7.5	25.1	30.2	14.5	29.0	25.9	15.7	27.4	25.6	6.7	24.9
Absolute bias (median: eGFR-mGFR)	21.4	-0.7	20.9	28.0	6.1	23.5	26.4	11.8	26.1	26.5	15.7	26.4	25.5	5.7	23.0
Relative bias (mean: eGFR/mGFR)	1.21	1.01	1.23	1.36	1.11	1.34	1.56	1.27	1.54	1.76	1.49	1.80	1.35	1.11	1.34
Relative bias (median: eGFR/mGFR)	1.21	0.99	1.20	1.36	1.08	1.31	1.47	1.21	1.47	1.67	1.40	1.70	1.33	1.07	1.30
Precision, RMSE	20.4	23.2	27.8	19.0	19.5	23.4	25.3	19.4	23.4	18.9	15.5	18.5	20.4	20.9	24.7
Precision, IQR	9.6;35.6	-16;15.3	3.6;38.4	12.8;40.2	-5.2;19.4	8.7;38.6	11.0;43.8	-0.1;25.5	11.9;41.5	12.2;38.0	7.4;21.3	18.4;34.5	11.8;39.0	-7.6;19.4	7.9;38.8
P10	0.22	0.32	0.31	0.13	0.31	0.18	0.09	0.23	0.09	0.13	0.07	0.13	0.15	0.30	0.20
P30	0.66	0.83	0.62	0.39	0.76	0.48	0.32	0.57	0.32	0.27	0.40	0.20	0.46	0.75	0.49
mGFR overestimated by 2+ ml/min	155 (83.8)	86 (46.5)	142 (76.8)	348 (90.2)	226 (58.5)	336 (87.0)	73 (92.4)	58 (73.4)	74 (93.7)	13 (86.7)	12 (80.0)	14 (93.3)	589 (88.6)	382 (57.4)	566 (85.1)
eGFR within +/- 2 ml/min of mGFR	8 (4.3)	8 (4.3)	11 (5.9)	12 (3.1)	33 (8.5)	12 (3.1)	3 (3.8)	7 (8.9)	2 (2.5)	2 (13.3)	1 (6.7)	1 (6.7)	25 (3.8)	49 (7.4)	26 (3.9)
mGFR underestimated by 2+ ml/min	22 (11.9)	91 (49.2)	32 (17.3)	26 (6.7)	127 (32.9)	38 (9.8)	3 (3.8)	14 (17.7)	3 (3.8)	0 (0.0)	2 (13.3)	0 (0.0)	51 (7.7)	234 (35.2)	73 (11.0)
mGFR overestimated by 5+ ml/min	149 (80.5)	77 (41.6)	132 (71.4)	337 (87.3)	199 (51.6)	315 (81.6)	69 (87.3)	52 (65.8)	69 (87.3)	13 (86.7)	12 (80.0)	13 (86.7)	568 (85.4)	340 (51.1)	529 (79.5)
eGFR within +/- 5 ml/min of mGFR	24 (13.0)	26 (14.1)	31 (16.8)	29 (7.5)	86 (22.3)	46 (11.9)	7 (8.9)	16 (20.3)	7 (8.9)	2 (13.3)	1 (6.7)	2 (13.3)	62 (9.3)	129 (19.4)	86 (12.9)
mGFR underestimated by 5+ ml/min	12 (6.5)	82 (44.3)	22 (11.9)	20 (5.2)	101 (26.2)	25 (6.5)	3 (3.8)	11 (13.9)	3 (3.8)	0 (0.0)	2 (13.3)	0 (0.0)	35 (5.3)	196 (29.5)	50 (7.5)
mGFR overestimated by 10+ ml/min	137 (74.1)	60 (32.4)	116 (62.7)	308 (79.8)	158 (40.9)	287 (74.4)	61 (77.2)	42 (53.2)	63 (79.7)	12 (80.0)	9 (60.0)	12 (80.0)	518 (77.9)	269 (40.5)	478 (71.9)
eGFR within +/- 10 ml/min of mGFR	38 (20.5)	58 (31.4)	57 (30.8)	67 (17.4)	165 (42.7)	85 (22.0)	16 (20.3)	34 (43.0)	14 (17.7)	3 (20.0)	6 (40.0)	3 (20.0)	124 (18.6)	263 (39.5)	159 (23.9)
mGFR underestimated by 10+ ml/min	10 (5.4)	67 (36.2)	12 (6.5)	11 (2.8)	63 (16.3)	14 (3.6)	2 (2.5)	3 (3.8)	2 (2.5)	0 (0.0)	0 (0.0)	0 (0.0)	23 (3.5)	133 (20.0)	28 (4.2)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

Figure S16: Lowess plot for log iohexol GFR (y axis) against log creatinine for men



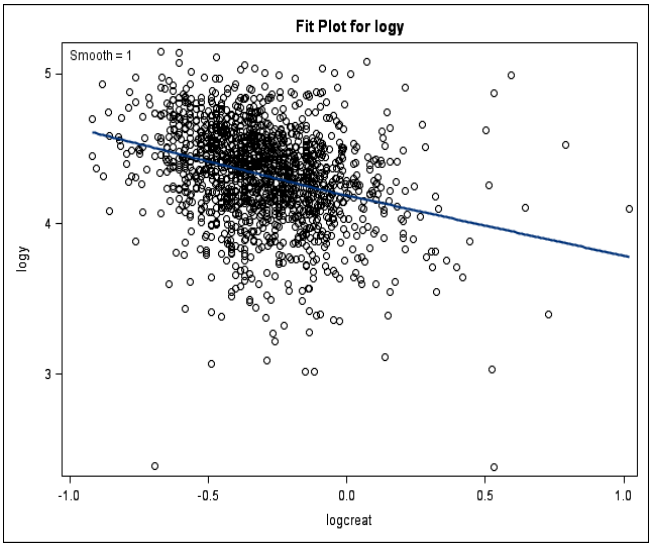
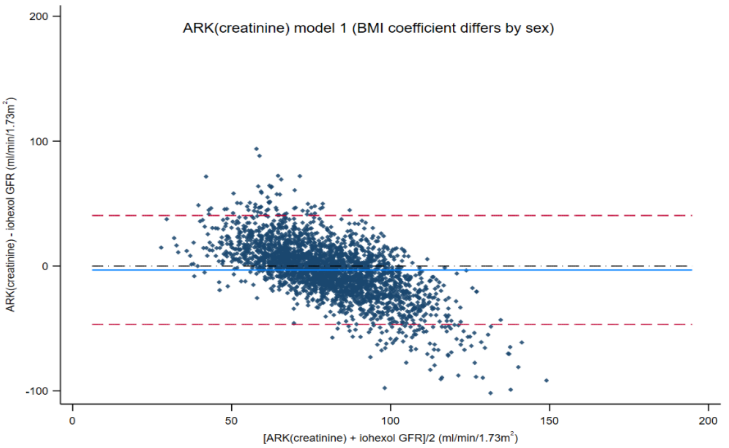
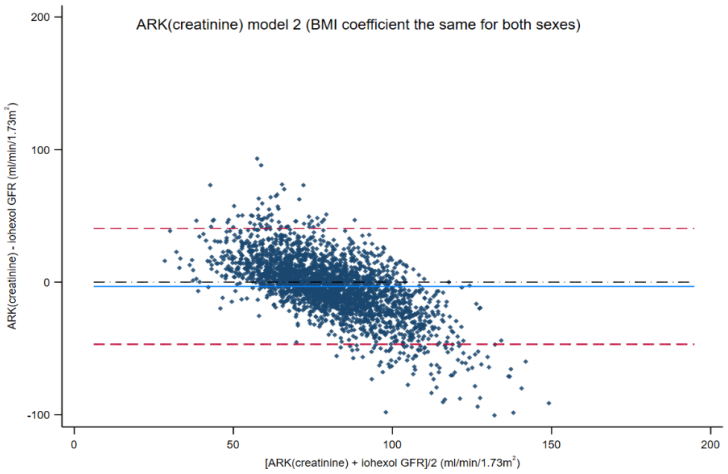
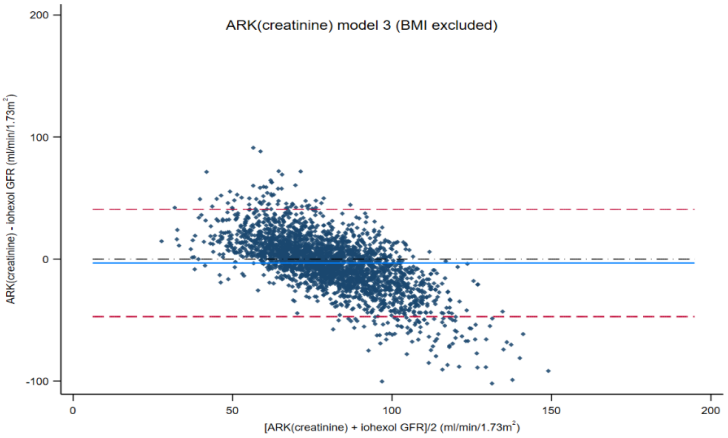


Figure S17: Agreement: models 1-3 for ARK (creatinine) eGFR equations and Iohexol GFR







Red dashed lines indicate limits of agreement (+-2 SD); solid blue line indicates bias (ml/min/1.72m^2)

Table S14: Models for the ARK (creatinine) eGFR equations

ARK eGFR Model #1 – with bmi (allowed to differ by sex):

n	-2log likelihood'	df	R^2	adj_R^2	Note
2578	1074.3	5	0.2380	0.2365	highest adj-r2

If male: eGFR = 126 x $\min(1, SCr/0.82)^{-0.344}$ x $\max(1,SCr/0.82)^{-0.571}$ x 0.993^{age} x 0.999^{bmi}
If female: eGFR = 126 x $(SCr/0.82)^{-0.344}$ x 0.993^{age} x 0.992^{bmi}

ARK eGFR Model #1prime if forcing BMI coefficient among males to be 1 then:

n	-2log likelihood'	df	R^2	adj_R^2	Note
2578	1074.8	4	0.2378	0.236615	highest adj-r2

If male: eGFR = 124 x min(1, SCr/0.82)-0.345 x max(1,SCr/0.82)-0.578 x 0.993age
If female: eGFR = 124 x (SCr/0.82)-0.345 x 0.993age x 0.993bmi

ARK eGFR Model #1prime if forcing BMI coefficient among males to be 1 then:

n	-2log likelihood'	df	R^2	adj_R^2	Note
2578	1074.8	4	0.2378	0.236615	highest adj-r2

If male: eGFR = 124 x min(1, SCr/0.82)-0.345 x max(1,SCr/0.82)-0.578 x 0.993age
If female: eGFR = 124 x (SCr/0.82)-0.345 x 0.993age x 0.993bmi

ARK eGFR Model #2 – with BMI (common coefficient for women and men):

n	-2log likelihood	df	R^2	adj_R^2	Note
2578	1081.5	5	0.2359	0.234415	increased face validity

If male: eGFR = 142 x min(1, SCr/0.82)-0.340 x max(1,SCr/0.82)-0.559 x 0.993age x 0.994bmi
If female: eGFR = 121 x (SCr/0.82)-0.340 x 0.993age x 0.994bmi

ARK eGFR Model #3 – without bmi:

n	-2log likelihood'	df	R^2	adj_R^2	Note
2578	1112.7	4	0.2266	0.225398	simpler model without bmi

If male: eGFR = 124 x min(1, SCr/0.82)-0.339 x max(1,SCr/0.82)-0.574 x 0.993age
If female: eGFR = 103 x (SCr/0.82)-0.339 x 0.993age

For all models, we checked for evidence of interactions between sex*creatinine.

Table S15: Overall: comparing ARK (creatinine) models to eGFR equations with iohexol GFR as the reference

GFR estimating equation	[1]Absolute bias	[2]Relative bias	[3]Precision (RMSE)	[4]P_{10}	[5]P_{30}
ARK (creatinine) model 1[a]	-1.7	0.98	22.3	0.32	0.77
ARK (creatinine) model 2[b]	-1.7	0.98	22.3	0.32	0.77
ARK (creatinine) model 3[c]	-1.9	0.98	22.4	0.33	0.77
Cockroft Gault (adjusted for BSA)	11.3	1.15	31.7	0.23	0.60
MDRD	6.6	1.08	27.1	0.27	0.66
MDRD ethnicity coefficient	24.7	1.31	30.1	0.17	0.46
FAS (creatinine)	11.2	1.14	27.8	0.25	0.62
Lund-Malmö (revised)	2.3	1.03	24.0	0.32	0.74
CKD-EPI (creatinine)	11.5	1.15	24.9	0.27	0.65
CKD-EPI (creatinine) ethnicity coefficient	26.7	1.33	26.4	0.15	0.45
CKD-EPI (creatinine + cystatin C)	4.1	1.05	23.8	0.32	0.73
CKD-EPI (cystatin C)	-1.7	0.98	25.9	0.28	0.70

[1]Absolute bias: median of the difference between (estimated GFR - iohexol GFR)
[2]Relative bias: median of the difference between ([estimated GFR - iohexol GFR]/iohexol GFR)
[3]Precision RMSE (Root Mean Square Error): standard deviation of (estimated GFR - iohexol GFR)
[4]Precision IQR (Interquartile Range): IQR for (estimated GFR - iohexol GFR)
[5]Accuracy: proportion of eGFR results within 10% (P_{10}) and 30% (P_{30}) of iohexol GFR
N=2578 for creatinine-based equations; N=2433 for cystatin C-based equations
[a]ARK (creatinine) model 1: model based on age, sex, and sex specific coefficients for BMI
[b]ARK (creatinine) model 2: model based on age, sex, and the same coefficient for BMI for both sexes
[c]ARK (creatinine) model 3: model based on age and sex only, no BMI

Table S16: Overall: comparing GFR stage by ARK (creatinine) models to eGFR equations with iohexol GFR

GFR stage[1]	mGFR[2]	ARK model 1[a]	ARK model 2[b]	ARK model 3[c]	CG(BSA)[3]	MDRD[4]	MDRD (ec)[5]	FAS (Cr)[6]	Lund-Malmö[7]	CKD-EPI (Cr)[8]	CKD-EPI (Cr, ec)[9]	CKD-EPI (Cr+cysC)[10]	CKD-EPI (cysC)[11]
G1	909 (**35**)	504 (**20**)	502 (**20**)	486 (**19**)	1386(**54**)	1140 (**44**)	1871 (**73**)	1405 (**55**)	962 (**37**)	1534 (**60**)	1993 (**77**)	1052 (**43**)	826 (**34**)
G2	1168 (**45**)	1920 (**75**)	1925 (**75**)	1949 (**76**)	928 (**36**)	1200 (**47**)	616 (**24**)	971 (**38**)	1382 (**54**)	874 (**34**)	493 (**19**)	1131(**47**)	1098 (**45**)
G3a	340 (**13**)	140 (**5**)	137 (**5**)	130 (**5**)	190 (**7**)	181 (**7**)	60 (**2**)	149 (**6**)	174 (**7**)	116 (**5**)	57 (**2**)	182 (**8**)	381(**16**)
G3b, G4, G5	161 (**6**)	14 (**1**)	14 (**1**)	13 (**1**)	74 (**3**)	57 (**2**)	31(**1**)	53 (**2**)	60 (**2**)	54 (**2**)	35 (**1**)	68 (**3**)	128 (**5**)
% classified correctly	reference category	1404 (**55**)	1401 (**54**)	1389 (**54**)	1264 (**49**)	1318 (**51**)	1150 (**45**)	1297 (**50**)	1391 (**54**)	1265 (**49**)	1126 (**44**)	1304 (**54**)	1196 (**49**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15
[2]mGFR: iohexol GFR; [3]CG(BSA): Cockroft Gault equation adjusted for body surface area; [4]MDRD: MDRD equation no ethnicity coefficient;
[5]MDRD (ec): MDRD equation with ethnicity coefficient; [6]FAS (cr): Full Age Spectrum equation for creatinine; [7]Lund-Malmö: Revised Lund-Malmö Study equation; [8]CKD-EPI(Cr): CKD-EPI (creatinine) equation no ethnicity coefficient; [9]CKD-EPI (Cr, ec): CKD-EPI (creatinine) equation with ethnicity coefficient; [10]CKD-EPI (Cr+cysC): CKD-EPI (creatinine + cystatin C) equation no ethnicity coefficient;
[11]CKD-EPI (cysC): CKD-EPI (cystatin C) equation.
N=2578 for creatinine-based equations; N=2433 for cystatin C-based equations
[a]ARK (creatinine) model 1: model based on age, sex, and sex specific coefficients for BMI
[b]ARK (creatinine) model 2: model based on age, sex, and the same coefficient for BMI for both sexes
[c]ARK (creatinine) model 3: model based on age and sex only, no BMI

Table S17: Overall and by iohexol GFR category: Comparing ARK (creatinine) models to eGFR equations

	mGFR>=90: Stage 1			60<=mGFR<90: Stage 2			45<=mGFR<60: Stage 3a			mGFR<45: Stage 3b,4,5			All stages		
	ARK Model 1 BMI x SEX	ARK Model 2 BMI pooled	ARK Model 3 no BMI	ARK Model 1 BMI x SEX	ARK Model 2 BMI pooled	ARK Model 3 no BMI	ARK Model 1 BMI x SEX	ARK Model 2 BMI pooled	ARK Model 3 no BMI	ARK Model 1 BMI x SEX	ARK Model 2 BMI pooled	ARK Model 3 no BMI	ARK Model 1 BMI x SEX	ARK Model 2 BMI pooled	ARK Model 3 no BMI
eGFR stage 1	345 (38.0)	344 (37.8)	328 (36.1)	129 (11.0)	129 (11.0)	128 (11.0)	18 (5.3)	18 (5.3)	19 (5.6)	12 (7.5)	11 (6.8)	11 (6.8)	504 (19.6)	502 (19.5)	486 (18.9)
eGFR stage 2	557 (61.3)	558 (61.4)	574 (63.1)	996 (85.3)	997 (85.4)	1003 (85.9)	267 (78.5)	270 (79.4)	270 (79.4)	100 (62.1)	100 (62.1)	102 (63.4)	1920 (74.5)	1925 (74.7)	1949 (75.6)
eGFR stage 3a	7 (0.8)	7 (0.8)	7 (0.8)	43 (3.7)	42 (3.6)	37 (3.2)	52 (15.3)	49 (14.4)	48 (14.1)	38 (23.6)	39 (24.2)	38 (23.6)	140 (5.4)	137 (5.3)	130 (5.0)
eGFR stage 3b,4,5	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	3 (0.9)	3 (0.9)	3 (0.9)	11 (6.8)	11 (6.8)	10 (6.2)	14 (0.5)	14 (0.5)	13 (0.5)
eGFR all stages	909 [35.3]	909 [35.3]	909 [35.3]	1168 [45.3]	1168 [45.3]	1168 [45.3]	340 [13.2]	340 [13.2]	340 [13.2]	161 [6.2]	161 [6.2]	161 [6.2]	2578 [100]	2578 [100]	2578 [100]
Correctly classified*	345 (38.0)	344 (37.8)	328 (36.1)	996 (85.3)	997 (85.4)	1003 (85.9)	52 (15.3)	49 (14.4)	48 (14.1)	11 (6.8)	11 (6.8)	10 (6.2)	1404 (54.5)	1401 (54.3)	1389 (53.9)
Stage as or more severe than mGFR	909 (100)	909 (100)	909 (100)	1039 (89.0)	1039 (89.0)	1040 (89.0)	55 (16.2)	52 (15.3)	51 (15.0)	11 (6.8)	11 (6.8)	10 (6.2)	2014 (78.1)	2011 (78.0)	2010 (78.0)
Absolute bias (mean: eGFR-mGFR)	-23	-23	-23	1.3	1.2	1.3	17.4	17.5	17.6	32.0	32.1	32.6	-3.2	-3.2	-3.2
Absolute bias (median: eGFR-mGFR)	-20	-20	-20	1.0	1.0	1.2	16.2	16.4	16.4	29.9	30.0	31.3	-1.7	-1.7	-1.9
Relative bias (mean: eGFR/mGFR)	0.80	0.80	0.80	1.03	1.02	1.03	1.33	1.33	1.34	2.09	2.09	2.10	1.05	1.05	1.05
Relative bias (median: eGFR/mGFR)	0.80	0.80	0.80	1.01	1.01	1.02	1.31	1.31	1.31	1.76	1.75	1.84	0.98	0.98	0.98
Precision, RMSE	18.6	18.6	18.6	11.3	11.3	11.1	11.9	11.8	11.7	17.6	17.5	17.5	22.3	22.3	22.4
Precision, IQR	-32; -11	-32; -10	-32; -11	-6.7; 8.2	-6.7; 8.1	-6.4; 8.3	9.0;24.5	9.0;24.6	9.1;24.6	19.1;43.4	19.4;41.8	19.6;44.7	-15; 9.7	-15; 9.8	-16; 9.7
P10	0.23	0.23	0.22	0.50	0.50	0.51	0.11	0.12	0.13	0.02	0.02	0.02	0.32	0.32	0.33
P30	0.78	0.78	0.78	0.94	0.94	0.95	0.48	0.48	0.48	0.09	0.09	0.08	0.77	0.77	0.77
mGFR overestimated by 2+ ml/min	54 (5.9)	51 (5.6)	44 (4.8)	541 (46.3)	535 (45.8)	553 (47.3)	318 (93.5)	322 (94.7)	324 (95.3)	156 (96.9)	157 (97.5)	156 (96.9)	1069 (41.5)	1065 (41.3)	1077 (41.8)
eGFR within +/- 2 ml/min of mGFR	39 (4.3)	41 (4.5)	40 (4.4)	180 (15.4)	184 (15.8)	166 (14.2)	14 (4.1)	11 (3.2)	9 (2.6)	3 (1.9)	2 (1.2)	3 (1.9)	236 (9.2)	238 (9.2)	218 (8.5)
mGFR underestimated by 2+ ml/min	816 (89.8)	817 (89.9)	825 (90.8)	447 (38.3)	449 (38.4)	449 (38.4)	8 (2.4)	7 (2.1)	7 (2.1)	2 (1.2)	2 (1.2)	2 (1.2)	1273 (49.4)	1275 (49.5)	1283 (49.8)
mGFR overestimated by 5+ ml/min	32 (3.5)	33 (3.6)	27 (3.0)	419 (35.9)	419 (35.9)	420 (36.0)	297 (87.4)	297 (87.4)	298 (87.6)	155 (96.3)	155 (96.3)	155 (96.3)	903 (35.0)	904 (35.1)	900 (34.9)
eGFR within +/- 5 ml/min of mGFR	94 (10.3)	91 (10.0)	92 (10.1)	398 (34.1)	396 (33.9)	404 (34.6)	38 (11.2)	38 (11.2)	37 (10.9)	4 (2.5)	5 (3.1)	4 (2.5)	534 (20.7)	530 (20.6)	537 (20.8)
mGFR underestimated by 5+ ml/min	783 (86.1)	785 (86.4)	790 (86.9)	351 (30.1)	353 (30.2)	344 (29.5)	5 (1.5)	5 (1.5)	5 (1.5)	2 (1.2)	1 (0.6)	2 (1.2)	1141 (44.3)	1144 (44.4)	1141 (44.3)
mGFR overestimated by 10+ ml/min	13 (1.4)	14 (1.5)	12 (1.3)	231 (19.8)	230 (19.7)	230 (19.7)	243 (71.5)	242 (71.2)	245 (72.1)	149 (92.5)	149 (92.5)	152 (94.4)	636 (24.7)	635 (24.6)	639 (24.8)
eGFR within +/- 10 ml/min of mGFR	206 (22.7)	202 (22.2)	197 (21.7)	748 (64.0)	753 (64.5)	766 (65.6)	93 (27.4)	94 (27.6)	91 (26.8)	12 (7.5)	12 (7.5)	9 (5.6)	1059 (41.1)	1061 (41.2)	1063 (41.2)
mGFR underestimated by 10+ ml/min	690 (75.9)	693 (76.2)	700 (77.0)	189 (16.2)	185 (15.8)	172 (14.7)	4 (1.2)	4 (1.2)	4 (1.2)	0 (0.0)	0 (0.0)	0 (0.0)	883 (34.3)	882 (34.2)	876 (34.0)

Unless otherwise indicated, numbers are N (%). Numbers in curved parentheses () are column percentages; square brackets [] indicate row percentages.
P10 and P30 indicate the proportion of eGFR results that fall within 10% and 30% respectively of the measured GFR.
Precision, RMSE (root mean square error) indicates the standard deviation of the errors, i.e. of the differences between eGFR and mGFR.
Precision, IQR indicates inter-quartile range of the errors, i.e. of the differences between eGFR and mGFR.
*Indicates number (%) where eGFR and mGFR-based CKD staging agree (disagreements between eGF and mGFR staging within stages 3b,4 and 5 not counted).
#Indicates number (%) where eGFR staging is as *or more* severe than mGFR staging.
All results based on GFR values measured in ml/min/1.73m^2

Table S18: Participant characteristics for external validation dataset (N=651)

Sample description	All participants Overall	Cohort Study HIV Positive Adults[18]	Clinical Sample Evaluating potential living kidney donors	Clinical Sample Evaluating kidney function for suspected CKD	Randomised Controlled Trial[19]	Cohort Study Adults with CKD[20]
Participant number N (%)	651 (100)	97 (14.9)	309 (47.5)	50 (7.7)	96 (14.8)	99 (15.2)
Radionuclide filtration marker		51Cr-EDTA[1]	51Cr-EDTA 99Tc-DTPA[2]	51Cr-EDTA 99Tc-DTPA	51Cr-EDTA	51Cr-EDTA
Dates of testing		2010-2011[3]	2007-2020[4]	2013-2020[5]	1994-1997[6]	2006[7]
Age - yr	43.3 (13.3)	37.0 (9.6)	40.3 (11.2)	49.3 (16.8)	52.5 (9.8)	46.5 (16.5)
Age category N (%)						
<40 yr	274 (42.1)	63 (65.0)	146 (47.3)	13 (26.0)	11 (11.5)	41 (41.4)
40-60 yr	310 (47.6)	33 (34.0)	149 (48.2)	26 (52.0)	64 (66.7)	38 (38.4)
60 yr	67 (10.3)	1 (1.0)	14 (4.5)	11 (22.0)	21 (21.9)	20 (20.2)
Female sex N (%)	365 (56.1)	40 (41.2)	185 (59.9)	16 (32.0)	76 (79.2)	48 (48.9)
Population Group						
Black	412 (63.3)	97 (100)	114 (36.9)	6 (12.0)	96 (100)	99 (100)
White	188 (28.9)	0 (0)	157 (50.8)	31 (62.0)	0 (0)	0 (0)
Indian/Asian	28 (4.3)	0 (0)	20 (6.5)	8 (16.0)	0 (0)	0 (0)
Mixed	13 (2.0)	0 (0)	12 (3.9)	1 (2.0)	0 (0)	0 (0)
Height (cm)	165.9 (9.6)	165.5 (8.5)	167.6 (9.2)	170.5 (10.5)	158.4 (8.0)	165.9 (9.0)
Weight (kg)	71.8 (14.9)	60.3 (12.4)	73.8 (14.3)	76.7 (15.6)	76.6 (13.6)	69.6 (13.8)
[9]Body mass index	26.2 (5.37)	22.1 (4.8)	26.2 (4.4)	26.3 (4.9)	30.6 (5.6)	25.5 (5.5)
BMI category N (%)						
<18.5 (underweight)	42 (6.5)	21 (21.7)	10 (3.2)	3 (6.0)	2 (2.1)	6 (6.1)
18.5-24.9 (normal)	229 (35.2)	53 (54.6)	99 (32.0)	13 (26.0)	13 (13.5)	51 (51.5)
25.0-29.9 (overweight)	210 (32.3)	11 (11.3)	124 (40.1)	25 (50)	28 (29.2)	22 (22.2)
>=30.0 (obese)	170 (26.1)	12 (12.4)	76 (24.6)	9 (18.0)	53 (55.2)	20 (20.2)
[10]Radionuclide GFR	82.0 (26.3)	90.2 (28.1)	90.3 (18.0)	56.2 (24.1)	82.2 (20.9)	60.7 (31.9)
[11]Serum creatinine (mg/dL)	1.08 (0.95)	1.00 (1.04)	0.87 (0.17)	1.60 (1.23)	0.85 (0.16)	1.80 (1.77)

All variables reported as mean (standard deviation); categories reported as frequency (percent); percentages might sum to +-100 due to rounding; for population group N=641; [1]51Cr-EDTA: chromium-51 labelled ethylene diamine tetra-acetic acid; [2]99Tc-DTPA: technetium-99 labelled diethylene triamine penta-acetic acid; [1,2]Before April 2019, the radionuclide testing comprised plasma excretion of 51Cr-EDTA, which was switched to plasma excretion of 99Tc-DTPA due to supply chain difficulties; [3]Multisample GFR testing at two and four hours; [4]Potential living kidney donors were evaluated using single sample GFR testing at three hours on the premise that potential donors are healthy with normal kidney function; [5]Those with suspected CKD were evaluated using multi-sample GFR testing at two and four hours; [6]Used single sample GFR testing at three hours; [7]Used multi-sample GFR testing at two and four hours if estimated GFR greater than 30ml/min/1.73m^2; and at two and five hours if estimated GFR less than or equal to 30ml/min/1.73m^2; [6]Body surface area calculated according to the Du Bois method[21]; Body mass index (BMI) calculated by dividing weight (kilograms) by height squared (metres); [8]Radionuclide GFR corrected for BSA (ml/min/1.73m^2); [9] To convert serum creatinine measurements from mg/dL to µmol/L, multiply by 88.4.All testing was performed in the Division of Nuclear Medicine, Department of Radiation Sciences, Charlotte Maxeke Johannesburg Academic Hospital, South Africa

Table S19: Performance of eGFR equations compared to radionuclide GFR (external validation)

GFR estimating equation	Absolute bias[1]	Relative Bias[2]	Precision[3]	Accuracy P_{10}[4]	Accuracy P_{30}[4]
CKD-EPI(creatinine)	5.5	1.10	18.6	0.34	0.77
Lund-Malmö (revised)	-3.3	1.00	17.4	0.37	0.85
ARK (creatinine)	-4.3	1.06	18.7	0.34	0.81

[1] Absolute bias: median of the difference (estimated GFR - radionuclide GFR)
[2] Relative bias: median of the difference (estimated GFR - radionuclide GFR/radionuclide GFR)
[3] Precision: RMSE (Root Mean Square Error): standard deviation of (estimated GFR - radionuclide GFR)
[4]Accuracy: the proportion of eGFR results that fall within 10% (P_{10}) and 30% (P_{30}) respectively, of radionuclide GFR

Table S20: GFR stage for eGFR equations compared to radionuclide GFR (external validation)

GFR stage[1]	Radionuclide mGFR	CKD-EPI(creatinine)	Revised Lund-Malmö	ARK(creatinine)
G1	256 (**39**)	324 (**50**)	206 (**32**)	115 (**18**)
G2	283 (**43**)	233 (**36**)	346 (**53**)	477 (**73**)
G3a	48 (**7**)	45 (**7**)	43 (**7**)	34 (**5**)
G3b	31 (**5**)	20 (**3**)	22 (**3**)	21 (**3**)
G4	27 (**4**)	14 (**2**)	23 (**4**)	4 (**1**)
G5	6 (**1**)	15 (**2**)	11 (**2**)	0 (**0**)
% participants classified correctly for all stages	Reference category	390 (**60**)	384 (**59**)	355 (**55**)

Data shown are number (%); [1]GFR (ml/min/1.73m^2)G1 >=90; G2 60-89; G3a 45-59; G3b 30-44; G4 15-29; G5 <15

Table S21: Predicted population prevalence of impaired kidney function in ARK countries comparing imputation and GFR estimates

GFR staging	ARK Uganda (N=5715)			ARK Malawi (N=4719)			ARK South Africa (N=2020)		
GFR (ml/min/1.73m^2)	**Imputation[1]**	**CKD-EPI[2]**	**Lund-Malmö[3]**	**Imputation**	**CKD-EPI[2]**	**Lund-Malmö[3]**	**Imputation**	**CKD-EPI[2]**	**Lund-Malmö[3]**
Stage G1 (>=90)	3687 (65)	4510 (79)	3546 (62)	2328 (49)	3852 (82)	2999 (64)	903 (45)	1760 (87)	1432 (71)
Stage G2 (60-89)	1667 (29)	1101 (19)	2008 (35)	1829 (39)	808 (17)	1616 (34)	806 (40)	221 (11)	534 (26)
Stage G3a (45-59)	256 (4)	83 (1)	132 (2)	379 (8)	45 (1)	85 (2)	207 (10)	30 (1)	42 (2)
Stage G3b (30-45)	80 (1)	14 (0)	22 (0)	138 (3)	10 (0)	14 (0)	83 (4)	6 (0)	8 (0)
Stage G4 (15-29)	19 (0)	4 (0)	4 (0)	36 (1)	4 (0)	5 (0)	18 (1)	3 (0)	4 (0)
Stage G5 (<15)	6 (0)	3 (0)	3 (0)	9 (0)	0 (0)	0 (0)	4 (0)	0 (0)	0 (0)
Proportion with impaired kidney function (G3-G5) % (95% confidence interval)	6.32% (5.07-7.56)	1.77% (1.49- 2.20)	2.82% (2.40-3.28)	11.91% (10.13-13.69)	1.25% (0.95-1.61)	2.21% (1.80-2.66)	15.41% (13.30-17.53)	1.93% (1.38-2.63)	2.67% (2.01-3.47)

Data shown are number (**%**); Due to rounding percentages may sum to + or -100; Due to rounding the sum of people across the GFR categories may differ by +1 or -1 from the total N indicated at the top;
[1]Predictions based on 100 imputed datasets; [2]CKD-EPI (creatinine) equation without adjustment for ethnicity; [3]Revised Lund-Malmö Study Equation

Table S22: Predicted population prevalence of impaired kidney function in AWI-Gen countries comparing imputation and GFR estimates

GFR staging	AWI-Gen Ghana (N=2011)			AWI-Gen Burkino Faso (N=2072)			AWI-Gen Kenya (N=2000)			AWI-Gen South Africa (N=5618)		
GFR (ml/min/1.73m²)	**Imputation[1]**	**CKD-EPI[2]**	**Lund-Malmö[3]**	**Imputation**	**CKD-EPI[2]**	**Lund-Malmö[3]**	**Imputation**	**CKD-EPI[2]**	**Lund-Malmö[3]**	**Imputation**	**CKD-EPI[2]**	**Lund-Malmö[3]**
Stage G1 (>=90)	795 (40)	1510 (75)	993 (49)	881 (43)	1729 (83)	1315 (63)	831 (42)	1519 (76)	1087 (54)	2022 (36)	3401 (61)	2155 (38)
Stage G2 (60-89)	889 (44)	451 (22)	961 (48)	899 (43)	298 (14)	706 (34)	867 (43)	424 (21)	846 (42)	2514 (45)	1986 (35)	3177 (57)
Stage G3a (45-59)	219 (11)	33 (2)	38 (2)	199 (10)	31 (1)	37 (2)	201 (10)	39 (2)	47 (2)	688 (12)	181 (3)	225 (4)
Stage G3b (30-45)	82 (4)	14 (1)	13 (1)	70 (3)	10 (0)	9 (0)	76 (4)	13 (1)	13 (1)	280 (5)	34 (1)	42 (1)
Stage G4 (15-29)	21 (1)	2 (0)	5 (0)	17 (1)	2 (0)	4 (0)	20 (1)	4 (0)	7 (0)	83 (1)	6 (0)	9 (0)
Stage G5 (<15)	6 (0)	1 (0)	1 (0)	6 (0)	2 (0)	1 (0)	5 (0)	1 (0)	0 (0)	30 (1)	10 (0)	10 (0)
Proportion with impaired kidney function (G3-G5) % (95% confidence interval)	16.30% (13.96-18.63)	2.49% (1.85-3.27)	2.83% (2.15-3.66)	14.12% (11.83-16.42)	2.17% (1.59-2.90)	2.46% (1.84-3.22)	15.08% (12.70-17.46)	2.85% (2.17-3.68)	3.35% (2.61-4.24)	19.25% (17.48-21.03)	4.12% (2.61-4.66)	5.10% (4.53-5.70)

Data shown are number (**%**); Due to rounding percentages may sum to + or -100; Due to rounding the sum of people across the GFR categories may differ by +1 or -1 from the total N indicated at the top;
[1]Predictions based on 100 imputed datasets; [2]CKD-EPI (creatinine) equation without adjustment for ethnicity; [3]Revised Lund-Malmö Study Equation

REFERENCES

1. Kalyesubula R, Fabian J, Nakanga W, et al. How to estimate glomerular filtration rate in sub-Saharan Africa: design and methods of the African Research into Kidney Diseases (ARK) study. BMC Nephrology 2020;21:1-12.

2. Nakanga WP, Prynn JE, Banda L, et al. Prevalence of impaired renal function among rural and urban populations: findings of a cross-sectional study in Malawi. Wellcome Open Res 2019;4:92.

3. Kalyesubula R, Hau JP, Asiki G, et al. Impaired renal function in a rural Ugandan population cohort. Wellcome Open Res 2018;3:149.

4. Fleming JS, Zivanovic MA, Blake GM, Burniston M, Cosgriff PS. Guidelines for the measurement of glomerular filtration rate using plasma sampling. Nuclear medicine communications 2004;25:759-69.

5. Haycock GB, Schwartz GJ, Wisotsky DH. Geometric method for measuring body surface area: a height-weight formula validated in infants, children, and adults. The Journal of pediatrics 1978;93:62-6.

6. Bröchner-Mortensen J. A simple method for the determination of glomerular filtration rate. Scandinavian journal of clinical and laboratory investigation 1972;30:271-4.

7. Nordin G, Ekvall S, Kristoffersson C, et al. Accuracy of determination of the glomerular filtration marker iohexol by European laboratories as monitored by external quality assessment. Clinical Chemistry and Laboratory Medicine (CCLM) 2019;57:1006-11.

8. Annesley TM, Clayton LT. Ultraperformance liquid chromatography–tandem mass spectrometry assay for iohexol in human serum. Clinical chemistry 2009;55:1196-202.

9. Parrinello CM, Grams ME, Couper D, et al. Recalibration of blood analytes over 25 years in the atherosclerosis risk in communities study: impact of recalibration on chronic kidney disease prevalence and incidence. Clinical chemistry 2015;61:938-47.

10. Kidney Disease Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 Clinical Practice Guideline for the Evaluation and Management of Chronic Kidney Disease. Kidney Int Suppl 2013;3:1-150.

11. Cockcroft DW, Gault MH. Prediction of creatinine clearance from serum creatinine. Nephron 1976;16:31-41.

12. Levey AS, Coresh J, Greene T, et al. Using standardized serum creatinine values in the modification of diet in renal disease study equation for estimating glomerular filtration rate. Annals of internal medicine 2006;145:247-54.

13. Levey AS, Stevens LA, Schmid CH, et al. A new equation to estimate glomerular filtration rate. Annals of internal medicine 2009;150:604-12.

14. Inker LA, Schmid CH, Tighiouart H, et al. Estimating glomerular filtration rate from serum creatinine and cystatin C. New England Journal of Medicine 2012;367:20-9.

15. Nyman U, Grubb A, Larsson A, et al. The revised Lund-Malmö GFR estimating equation outperforms MDRD and CKD-EPI across GFR, age and BMI intervals in a large Swedish population. Clinical Chemistry and Laboratory Medicine (CCLM) 2014;52:815-24.

16. Pottel H, Hoste L, Dubourg L, et al. An estimated glomerular filtration rate equation for the full age spectrum. Nephrology Dialysis Transplantation 2016;31:798-806.

17. Pottel H, Vrydags N, Mahieu B, Vandewynckele E, Croes K, Martens F. Establishing age/sex related serum creatinine reference intervals from hospital laboratory data based on different statistical methods. Clinica chimica acta 2008;396:49-55.

18. Seape T, Gounden V, van Deventer HE, Candy GP, George JA. Cystatin C-and creatinine-based equations in the assessment of renal function in HIV-positive patients prior to commencing highly active antiretroviral therapy. Annals of Clinical Biochemistry: An international journal of biochemistry and laboratory medicine 2016;53:58-66.

19. Sareli P, Radevski IV, Valtchanova ZP, et al. Efficacy of different drug classes used to initiate antihypertensive treatment in black subjects: results of a randomized trial in Johannesburg, South Africa. Archives of Internal Medicine 2001;161:965-71.

20. van Deventer HE, Paiker JE, Katz IJ, George JA. A comparison of cystatin C-and creatinine-based prediction equations for the estimation of glomerular filtration rate in black South Africans. Nephrology Dialysis Transplantation 2010:gfq621.

21. Du Bois D, Du Bois E. A formula to estimate the approximate surface area if height and weight are known. Archives of Internal Medicine 1917;17:863-71.

22. World Health Organisation. Managing the Global Epidemic of Obesity: Report of the World Health Organization Consultation of Obesity. Geneva, June 1997.

R14/49 «Tit init name»

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M170583

NAME: (Principal Investigator)	Dr June Fabian et al
DEPARTMENT:	Internal Medicine Wits Donald Gordon Medical Centre
PROJECT TITLE:	Determining the Prevalence of Chronic Kidney Disease (CKD) in a Rural Population in South Africa
DATE CONSIDERED:	Adhoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-Study under Primary Studies M1403107, M160939, M160937 and M160939 Principal Investigator Prof Saraladevi Naicker et al
SUPERVISOR:	
APPROVED BY:	Professor P. Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	05/06/2017

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand. I/We fully understand the conditions under which I am/we are authorised to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit to the Committee. **I agree to submit a yearly progress report**. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. in this case, the study was initially review in May and will therefore be due in the month of May each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

Date 24/06/2017

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG | WITS SCHOOL OF CLINICAL MEDICINE

Department of Internal Medicine

29 April 2021

Professor Maria Papathanasopoulos
Chair, Faculty Graduate Studies
Wits University

Dear Professor Papathanasoupolous

Re: Dr June Fabian: Student Number 8600952/N - PhD degree. Turn-it-in Report

I am writing this letter also on behalf of Dr Fabian's supervisors (Prof Jaya George & Dr Alisha Wade). We have reviewed the full Turn-it-in report, which has a similarity index of 20%. This is largely derived from definitions, terminology, list of abbreviations, titles of papers/manuscripts, names of people and institutions.

We are therefore satisfied that there has been no plagiarism.

Yours sincerely

Saraladevi Naicker
[MB ChB(Natal), MRCP(UK), FRCP(Lond), PCP(SA), PhD(Natal)
Supervisor
Emeritus Professor of Medicine
Email: saraladevi.naicker@wits.ac.za

phd_in_v4_28.4.2021_June.docx

ORIGINALITY REPORT

20%	17%	15%	6%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	academic.oup.com Internet Source	1%
2	caremore.com Internet Source	1%
3	June Fabian, Jaya A George, Harriet R Etheredge, Manuel van Deventer et al. "Methods and reporting of kidney function: a systematic review of studies from sub-Saharan Africa", Clinical Kidney Journal, 2019 Publication	1%
4	hdl.handle.net Internet Source	1%
5	bmcnephrol.biomedcentral.com Internet Source	1%
6	eprints.gla.ac.uk Internet Source	1%
7	link.springer.com Internet Source	1%
8	researchonline.lshtm.ac.uk Internet Source	

		1 %
9	Kimberly J. Reidy, Rebecca C. Hjorten, Claire L. Simpson, Avi Z. Rosenberg et al. "Fetal—not Maternal—APOL1 Genotype Associated with Risk for Preeclampsia in Those with African Ancestry", The American Journal of Human Genetics, 2018 Publication	<1 %
10	lup.lub.lu.se Internet Source	<1 %
11	wiredspace.wits.ac.za Internet Source	<1 %
12	escholarship.org Internet Source	<1 %
13	Submitted to Chester College of Higher Education Student Paper	<1 %
14	Munin.uit.no Internet Source	<1 %
15	orbi.uliege.be Internet Source	<1 %
16	doaj.org Internet Source	<1 %
17	epubs.surrey.ac.uk Internet Source	

		<1%
18	journals.plos.org Internet Source	<1%
19	Sean Currin, Mwawi Gondwe, Nokthula Mayindi, Shingirai Chipungu et al. "Evaluating chronic kidney disease in rural South Africa: comparing estimated glomerular filtration rate using point-of-care creatinine to iohexol measured GFR", Clinical Chemistry and Laboratory Medicine (CCLM), 2021 Publication	<1%
20	www.turkbiyokimyadernegi.org.tr Internet Source	<1%
21	www.ncbi.nlm.nih.gov Internet Source	<1%
22	sajhivmed.org.za Internet Source	<1%
23	www.researchsquare.com Internet Source	<1%
24	pure.uva.nl Internet Source	<1%
25	edoc.ub.uni-muenchen.de Internet Source	<1%
26	wellcomeopenresearch.org Internet Source	

		<1 %
27	kar.kent.ac.uk Internet Source	<1 %
28	insights.ovid.com Internet Source	<1 %
29	e-sciencecentral.org Internet Source	<1 %
30	s-space.snu.ac.kr Internet Source	<1 %
31	theses.gla.ac.uk Internet Source	<1 %
32	uwe-repository.worktribe.com Internet Source	<1 %
33	www.metrologie-francaise.fr Internet Source	<1 %
34	Submitted to North West University Student Paper	<1 %
35	Submitted to Trinity College Dublin Student Paper	<1 %
36	Submitted to University of Wales, Bangor Student Paper	<1 %
37	Submitted to Samuel Merritt University Student Paper	<1 %

38	watermark.silverchair.com Internet Source	<1 %
39	Management of Chronic Kidney Disease, 2014. Publication	<1 %
40	Sean D. Currin, Mwawi S. Gondwe, Nokthula B. Mayindi, Shingirai Chipungu et al. "Diagnostic accuracy of semiquantitative point of care urine albumin to creatinine ratio and urine dipstick analysis in a primary care resource limited setting in South Africa", BMC Nephrology, 2021 Publication	<1 %
41	bmchealthservres.biomedcentral.com Internet Source	<1 %
42	bmcpublichealth.biomedcentral.com Internet Source	<1 %
43	Submitted to Monash University Student Paper	<1 %
44	mafiadoc.com Internet Source	<1 %
45	www.frontiersin.org Internet Source	<1 %
46	www.nice.org.uk Internet Source	<1 %

47	Huang, S.H.S.. "Performance of the creatinine-based and the cystatin C-based glomerular filtration rate (GFR) estimating equations in a heterogenous sample of patients referred for nuclear GFR testing", Translational Research, 201106 Publication	<1 %
48	Sikhumbuzo A Mabunda, Khuthala Sigovana, Wezile Chitha, Teke Apalata, Sibusiso Nomatshila. "Socio-demographic associations of HIV among women attending antenatal care in selected rural primary care facilities in South Africa's Eastern Cape province.", Research Square, 2020 Publication	<1 %
49	etheses.whiterose.ac.uk Internet Source	<1 %
50	Enyu Imai, Masaru Horio, Kosaku Nitta, Kunihiro Yamagata et al. "Estimation of glomerular filtration rate by the MDRD study equation modified for Japanese patients with chronic kidney disease", Clinical and Experimental Nephrology, 2007 Publication	<1 %
51	Submitted to The University of Manchester Student Paper	<1 %
52	www.science.gov Internet Source	

		<1%
53	Jaya A George, Jean-Tristan Brandenburg, June Fabian, Nigel J Crowther et al. "Kidney damage and associated risk factors in rural and urban sub-Saharan Africa (AWI-Gen): a cross-sectional population study", The Lancet Global Health, 2019 Publication	<1%
54	Submitted to Queen's University of Belfast Student Paper	<1%
55	Ahmed M. E. Elkhalifa, Rashad Abdul-Ghani, Abdelhakam G. Tamomh, Nur Eldin Eltaher et al. "Hematological Indices and Abnormalities Among Patients With Uncomplicated Falciparum Malaria in Kosti City of the White Nile State, Sudan: a Comparative Study", Research Square, 2021 Publication	<1%
56	Submitted to London School of Hygiene and Tropical Medicine Student Paper	<1%
57	bmcmicrobiol.biomedcentral.com Internet Source	<1%
58	Lesley A. Inker, Sara J. Couture, Hocine Tighiouart, Alison G. Abraham et al. "A New Panel Estimated GFR, Including β2-	<1%

Microglobulin and β-Trace Protein and Not Including Race, Developed in a Diverse Population", American Journal of Kidney Diseases, 2020
Publication

59	Submitted to University of Witwatersrand Student Paper	<1 %
60	ajlmonline.org Internet Source	<1 %
61	pdfs.semanticscholar.org Internet Source	<1 %
62	Submitted to University of St Andrews Student Paper	<1 %
63	biobanking.org Internet Source	<1 %
64	duter.unistra.fr Internet Source	<1 %
65	ghrp.biomedcentral.com Internet Source	<1 %
66	repository.up.ac.za Internet Source	<1 %
67	Ernest K Sumaili, Revital Shemer, Etty Kruzel-Davila, Eric P Cohen et al. " G1 is the major risk allele for hypertension-attributed	<1 %

nephropathy in Central Africa ", Clinical Kidney Journal, 2019
Publication

68	Gwladys N. Gbaguidi, Corine Y. Houehanou, Salimanou A. Amidou, Jacques Vigan, Dismand S. Houinato, Philippe Lacroix. "Prevalence of abnormal kidney function in a rural population of Benin and associated risk factors", BMC Nephrology, 2021 Publication	<1 %
69	Robert Kalyesubula, Jeffrey P. Hau, Gershim Asiki, Billy Ssebunya et al. "Impaired renal function in a rural Ugandan population cohort", Wellcome Open Research, 2018 Publication	<1 %
70	Tae-Dong Jeong, Eun-Jung Cho, Woochang Lee, Sail Chun, Ki-Sook Hong, Won-Ki Min. "Accuracy Assessment of Five Equations Used for Estimating the Glomerular Filtration Rate in Korean Adults", Annals of Laboratory Medicine, 2017 Publication	<1 %
71	Submitted to University of Birmingham Student Paper	<1 %
72	repositorium.sdum.uminho.pt Internet Source	<1 %
73	vbn.aau.dk Internet Source	

		<1 %
74	www.mdpi.com Internet Source	<1 %
75	www.zora.uzh.ch Internet Source	<1 %
76	samj.org.za Internet Source	<1 %
77	slidelegend.com Internet Source	<1 %
78	www.e-sc.org Internet Source	<1 %
79	Jithin Sam Varghese, John Maluccio, Solveig Cunningham, Manuel Ramirez-Zea, Aryeh Stein. "Development of a temporally harmonized asset index: Evidence from across 50 years of follow up of a birth cohort in Guatemala", Research Square, 2021 Publication	<1 %
80	Lesley A. Inker, Andrew S. Levey. "Measurement and Estimation of Kidney Function", Elsevier BV, 2019 Publication	<1 %
81	archive.org Internet Source	<1 %

82	en.wikipedia.org Internet Source	<1 %
83	gpnotebook.com Internet Source	<1 %
84	professional.heart.org Internet Source	<1 %
85	www.slideshare.net Internet Source	<1 %
86	Cindy George, Amelie Mogueo, Ikechi Okpechi, Justin B Echouffo-Tcheugui, Andre Pascal Kengne. "Chronic kidney disease in low-income to middle-income countries: the case for increased screening", BMJ Global Health, 2017 Publication	<1 %
87	Faizeen Zafar, Adam M. Lubert, David A. Katz, Garick D. Hill et al. "Long-Term Kidney Function After the Fontan Operation", Journal of the American College of Cardiology, 2020 Publication	<1 %
88	Lesley A. Inker, Christina Wyatt, Rebecca Creamer, James Hellinger et al. "Performance of Creatinine and Cystatin C GFR Estimating Equations in an HIV-Positive Population on Antiretrovirals", JAIDS Journal of Acquired Immune Deficiency Syndromes, 2012 Publication	<1 %

89	Stuart A. Ali, Cassandra Soo, Godfred Agongo, Marianne Alberts et al. "Genomic and environmental risk factors for cardiometabolic diseases in Africa: methods used for Phase 1 of the AWI-Gen population cross-sectional study", Global Health Action, 2018 Publication	<1 %
90	Teresa K. Chen, Michael J. Choi, W.H. Linda Kao, Brad C. Astor et al. " Examination of Potential Modifiers of the Association of Alleles with CKD Progression ", Clinical Journal of the American Society of Nephrology, 2015 Publication	<1 %
91	bmcmedicine.biomedcentral.com Internet Source	<1 %
92	ccsenet.org Internet Source	<1 %
93	docplayer.net Internet Source	<1 %
94	etheses.bham.ac.uk Internet Source	<1 %
95	fr.scribd.com Internet Source	<1 %
96	www.labome.org Internet Source	<1 %

97	Andrea O. Ruiz-Alejos, Ben Caplin, J. Jaime Miranda, Neil Pearce, Antonio Bernabé-Ortiz. "CKD and CKDu in Northern Peru: a cross-sectional analysis under the DEGREE protocol", Research Square, 2020 Publication	<1 %
98	Archna Bajaj, Katalin Susztak, Scott M. Damrauer. " and Cardiovascular Disease ", Arteriosclerosis, Thrombosis, and Vascular Biology, 2017 Publication	<1 %
99	Elke Schaeffner. "Determining the Glomerular Filtration Rate—An Overview", Journal of Renal Nutrition, 2017 Publication	<1 %
100	Mitku Mammo Taderegew. "Assessment of renal impairment using estimated glomerular filtration rate among type 2 diabetes mellitus patients in North-East Ethiopia: a cross-sectional study", Journal of Diabetes & Metabolic Disorders, 2020 Publication	<1 %
101	Mustafa Zafer Temiz, Ibrahim Hacibey, Ramazan Omer Yazar, Mehmet Salih Sevdi et al. "Altered kidney function and acute kidney damage markers predict survival outcomes of COVID-19 patients: A prospective pilot study", Cold Spring Harbor Laboratory, 2021	<1 %

Publication

102	Nareshni Moodley, Sudesh Hariparshad, Fozy Peer, Verena Gounden. "Evaluation of the CKD-EPI creatinine based glomerular filtration rate estimating equation in Black African and Indian adults in KwaZulu-Natal, South Africa", Clinical Biochemistry, 2018 Publication	<1 %
103	Nicola Glaser, Andreas Deckert, Sam Phiri, Dietrich Rothenbacher, Florian Neuhann. "Comparison of Various Equations for Estimating GFR in Malawi: How to Determine Renal Function in Resource Limited Settings?", PLOS ONE, 2015 Publication	<1 %
104	Salman Ahmed, Cameron T. Nutt, Nwamaka D. Eneanya, Peter P. Reese et al. "Examining the Potential Impact of Race Multiplier Utilization in Estimated Glomerular Filtration Rate Calculation on African-American Care Outcomes", Journal of General Internal Medicine, 2020 Publication	<1 %
105	Tan, Karen, and Sunil K. Sethi. "Biomarkers in cardiorenal syndromes", Translational Research, 2014. Publication	<1 %

106	Vivek Kumar, Ashok Kumar Yadav, Yoshinari Yasuda, Masaru Horio et al. "Existing creatinine-based equations overestimate glomerular filtration rate in Indians", BMC Nephrology, 2018 Publication	<1 %
107	cjasn.asnjournals.org Internet Source	<1 %
108	discovery.ucl.ac.uk Internet Source	<1 %
109	doi.org Internet Source	<1 %
110	espace.library.uq.edu.au Internet Source	<1 %
111	health-sciences.nwu.ac.za Internet Source	<1 %
112	mgend.med.kyoto-u.ac.jp Internet Source	<1 %
113	pubs.rsna.org Internet Source	<1 %
114	repository.nwu.ac.za Internet Source	<1 %
115	scholar.sun.ac.za Internet Source	<1 %

116	synapse.koreamed.org Internet Source	<1 %
117	worldwidescience.org Internet Source	<1 %
118	www.birmingham.ac.uk Internet Source	<1 %
119	www.degruyter.com Internet Source	<1 %
120	www.fsk.it Internet Source	<1 %
121	www.hindawi.com Internet Source	<1 %
122	www.hlth.gov.bc.ca Internet Source	<1 %
123	www.jcmad.com Internet Source	<1 %
124	www.jcpsp.pk Internet Source	<1 %
125	www.kdigo.org Internet Source	<1 %
126	www.nature.com Internet Source	<1 %
127	www.niddk.nih.gov Internet Source	<1 %

128	www.researchgate.net Internet Source	<1%
129	www.tandfonline.com Internet Source	<1%
130	Gagneux-Brunon, A., C. Mariat, and P. Delanaye. "Cystatin C in HIV-infected patients: promising but not yet ready for prime time", Nephrology Dialysis Transplantation, 2012. Publication	<1%
131	PIERRE DELANAYE. "Kidney function", Elsevier BV, 2020 Publication	<1%
132	Piyawan Kittiskulnam, Krittaya Tiskajornsiri, Pisut Katavetin, Tawatchai Chaiwatanarat et al. "The failure of glomerular filtration rate estimating equations among obese population", PLOS ONE, 2020 Publication	<1%
133	"AMERICAN TRANSPLANT CONGRESS 2012 ABSTRACTS", American Journal of Transplantation, 05/2012 Publication	<1%
134	"Chapter 1: Definition and classification of CKD", Kidney International Supplements, 2013. Publication	<1%

135 Anna Jonsson, Ida Viklund, Andreas Jonsson, Fredrik Valham, Ellinor Bergdahl, Krister Lindmark, Helena Norberg. "Comparison of creatinine ‐ based methods for estimating glomerular filtration rate in patients with heart failure", ESC Heart Failure, 2020 <1%
Publication

136 Felicia Leion, Josefine Hegbrant, Emil den Bakker, Magnus Jonsson et al. "Estimating glomerular filtration rate (GFR) in children. The average between a cystatin C- and a creatinine-based equation improves estimation of GFR in both children and adults and enables diagnosing Shrunken Pore Syndrome", Scandinavian Journal of Clinical and Laboratory Investigation, 2017 <1%
Publication

137 Mogamat Razeen Davids, Mogamat-Yazied Chothia, Nicola Wearne, Bianca Davidson, Mignon McCulloch. "Chapter 6 Nephrology in South Africa", Springer Science and Business Media LLC, 2021 <1%
Publication

138 R. Kalyesubula, J. Fabian, R. Newton, J. George Anna, G. Dreyer, S. Naicker, M. Crampin, L. Smeeth, L. Tomlinson. "SAT-196 HOW TO ESTIMATE GLOMERULAR FILTRATION RATE IN SUB-SAHARAN AFRICA: DESIGN AND <1%

METHODS OF THE AFRICAN RESEARCH ON KIDNEY DISEASES (ARK) STUDY", Kidney International Reports, 2019
Publication

---

139 Richard J. Glassock, David G. Warnock, Pierre Delanaye. "The global burden of chronic kidney disease: estimates, variability and pitfalls", Nature Reviews Nephrology, 2016 <1%
Publication

---

140 Robert Kalyesubula, Jeffrey P. Hau, Gershim Asiki, Billy Ssebunya et al. "Impaired renal function in a rural Ugandan population cohort", Wellcome Open Research, 2019 <1%
Publication

---

141 Seape, T., V. Gounden, H. E. van Deventer, G. Candy, and J. A. George. "ANNALS EXPRESS: Cystatin C and creatinine based equations in the assessment of renal function in HIV positive patients prior to commencing Highly Active Antiretroviral Therapy (HAART)", Annals of Clinical Biochemistry An international journal of biochemistry and laboratory medicine <1%
Publication

---

142 Vasile Cepoi, Mihai Onofriescu, Liviu Segall, Adrian Covic. "The prevalence of chronic kidney disease in the general population in <1%

Romania: a study on 60,000 persons", International Urology and Nephrology, 2011
Publication

143 Wyatt, Christina M., George J. Schwartz, Willis Owino Ong'or, Joseph Abuya, Alison G. Abraham, Charles Mboku, Loice B. M'mene, Winnie J. Koima, Mathew Hotta, Paula Maier, Paul E. Klotman, and Kara Wools-Kaloustian. "Estimating Kidney Function in HIV-Infected Adults in Kenya: Comparison to a Direct Measure of Glomerular Filtration Rate by Iohexol Clearance", PLoS ONE, 2013.
Publication
<1%

144 wrap.warwick.ac.uk
Internet Source
<1%

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		