TECHNISCHE UNIVERSITÄT MÜNCHEN

FAKULTÄT FÜR MEDIZIN

Plasma proteomic biomarkers and DNA methylation-based measures of biological aging and lifespan: a multi-omics approach to kidney function and related traits

Pamela Raquel Matías García

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademisches Grades einer

Doktorin der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitz	Prof. Dr. Radu Roland Rad
Prüfende/-r der Dissertation	Prof. Dr. Juliane Winkelmann
	apl. Prof. Dr. Jerzy Adamski

Die Dissertation wurde am 19.07.2021 bei der Technische Universität München eingereicht und durch die Fakultät für Medizin am 12.10.2021 angenommen.

In loving memory of my Dad, Rubén Matías Silva,
who taught me how to run and encouraged me to fly.
I will always be thankful for those early mornings and
the adventures we shared at the beginning of this journey.

And to my mom, Sandra García Rosales,
whose love, support and encouragement over the years
carried me from my first steps to this achievement here.
Mi agradecimiento, admiración y girasoles para ti, mamá 🌻

"In his essay entitled 'What is Semantics?', Anatol Rapoport wrote:

'There are two suffixes in our language (and similar ones in other European languages) which suggest organized knowledge. One is the venerable, academic "ology," that reminds one of university curricula and scholarship. The other is the energetic and somewhat mysterious "ics," which has a connotative flavor of magic. Where "ology" suggests academic isolation (ichthyology, philology), "ics" suggests a method of attack on life's problems. It contains a faint throwback to the ancient dreams of the philosopher's stone and of "keys" to the riddles of the universe. Ancient words ending in "ics" are mathematics and meta-physics. Of more recent origin are economics, statistics, semantics, and cybernetics'.

One might add genetics, and now, *genomics*".

(McKusick & Ruddle, 1987)

Acknowledgements

This thesis is the final product of many years of work that would not have been possible without the support of many people around me, so I would like to dedicate this section to all the persons who accompanied me on this adventure.

Foremost I would like to express my sincere gratitude to Prof. Dr. Juliane Winkelmann for the opportunity to work with her, for the freedom and trust she put in me to explore this topic and for her continued support throughout my doctoral research. My gratitude also goes to Prof. Dr. Jerzy Adamski for his interest and involvement in our discussions, for his feedback and helpful advice, and for also being part of my thesis committee.

I am grateful to Prof. Dr. Annette Peters for her mentorship during my time at the Institute of Epidemiology at Helmholtz Zentrum München, and to Dr. Christian Gieger for his support through the years. I would like to offer my special thanks to Dr. Melanie Waldenberger for her guidance through each stage of the process and the many years of collaborative work. I am deeply grateful for their mentorship, which has been instrumental in defining my research interests and developing my research skills. I would also like to thank the TUM Publishing Fund for having supported the open access publication of one of the first-author publications, and the HELENA Graduate School for the funding/ support provided to participate in external and internal academic events.

I am indebted to my collaborators and co-authors for their crucial input for both manuscripts. I would like to express gratitude to Prof. Dr. Nora Franceschini and Dr. Laura Raffield for their guidance and insightful discussions on epigenetic aging in relation to kidney function. I am grateful to Dr. Alexander Teumer for his mentorship in statistical matters, to Dr. Johannes Graumann for his involvement

in discussions on plasma proteomics, and to Dr. Cavin Ward-Caviness for his guidance regarding the subtleties of the epigenetic clocks. I would also like to extend my sincere thanks to Prof. Dr. Maciej Tomaszewski and his team for their readiness to contribute to our work with their invaluable kidney tissue resource. A special thank you to Prof. Dr. Anna Köttgen, Prof. Dr. Iris Heid and the members of the CKDGen Consortium for enabling newcomers in the field of kidney research to learn from their shared collective expertise and for the opportunity to work with them. A heartfelt thank you to all colleagues involved in the KORA study, as well as a sincere thank you to my collaborators from QMDiab, INTERVAL, HUNT, NAS, ESTHER, JHS and WHI for your efforts. Our collaborations were a truly formative experience for me, and I am thankful for your shared expertise and the fruitful work. My gratitude goes also to all participants and study personnel from the cohorts and study samples that contributed to the work presented in this dissertation.

I would also like to thank colleagues and friends at the Research Unit Molecular Epidemiology (AME) for all the nice experiences. A big thank you to Rory Wilson for all the interesting scientific (and also non-scientific) discussions through the years, for his advice on a wide range of matters and for his readiness to help out. I would also like to thank Mohamed Elhadad for the team work, insightful discussions and companionship in our time as doctoral researchers. Thank you to all colleagues who, although not mentioned here, made working in a team like ours my doctoral experience all the more rewarding and enjoyable.

I would like to thank my family for believing in me from day one, and for their support despite the long distance. My dad, who taught me to keep pushing until my last breath. My mom, for her tireless encouragement. To my siblings, Arantza and Rubén, who never fail to gift me laughter, even in the most stressful times. A wholehearted thank you also goes to my aunt Gudelia and to my uncle Ulises for their support through the years. Thanks to all the friends here and there who cheered me on and made life more colorful during my time as doctoral researcher. Finally, I am deeply grateful to my newlywed husband Aaron for the many ways in which his unconditional love and encouragement kept me afloat while navigating this adventure <3.

Abstract

The kidneys are responsible for endocrine, metabolic and excretory functions vital to maintaining a homeostatic balance. Mostly asymptomatic in early stages, chronic kidney disease (CKD) is diagnosed based on levels of estimated glomerular filtration rate (eGFR) falling below 60 ml/min/1.73 m^2 – that is, when around 50% of kidney function has been lost. CKD is an increasingly prevalent non-communicable disease in aging populations that imposes a considerable burden on individuals and health care systems. Kidney research is thus a field in which molecular epidemiological research can provide novel insights into the biological mechanisms underlying kidney function and disease by identifying molecular markers associated and/or predictive of disease progression.

Many well-designed studies have been conducted in the last decades to characterize molecular (or -omic) markers in relation to kidney disease. Whereas urine proteomic biomarkers have made their way to being used in clinical practice (e.g. to predict disease progression), the complex proteomic profile of blood remains relatively unexplored. Moreover, prior studies on proteomic markers and renal function have not distinguished causality from correlation. Likewise, multiple studies have shown epigenetic age acceleration measured in blood to be associated with aging-related diseases (e.g. frailty, cognitive function and physical fitness, among a number of other conditions) and mortality. Nevertheless, whether DNAm-based predictors of age and mortality are correlated with different parameters of kidney aging and low function has not been investigated.

The general purpose of this dissertation is to contribute to the understanding of the molecular basis of kidney function (and disease) by investigating the association between kidney traits and omic-based biomarkers, more precisely plasma proteins and DNAm-based predictors of aging and/or mortality, in multi-ethnic population-

based studies. This was achieved by conducting two large-scale epidemiological research projects with international collaboration partners with the objective of addressing two main research goals: to determine whether plasma proteins are associated with eGFR and CKD and assess whether this association might be causal, and to investigate if epigenetic age acceleration and DNAm-based mortality measures in blood are associated with kidney function as reflected in multiple traits.

The findings presented in this thesis consist of 57 plasma proteins associated with kidney function (with eGFR as a proxy thereof). Of these associations, one seems to be of causal nature: Mendelian randomization suggested that eGFR has a positive effect on the levels of testican-2. Likewise, the strong and consistent association of different kidney traits reflecting poor kidney function (i.e. low eGFR, prevalent CKD, high urinary albumin-to-creatinine ratio [uACR], prevalent albuminuria and high serum urate) with DNAm-based signatures of aging and/or lifespan highlights the pervasive nature of aging-related chronic inflammation and immune system aging. This thesis presents an initial epidemiological appraise on the correlation between kidney function, proteomic biomarkers and epigenetic age acceleration in population-based studies. Further dedicated research to assess the clinical validity of the proteins and "epigenetic clocks" identified here for disease stratification and prognosis is warranted.

Zusammenfassung

Die Nieren sind für endokrine, metabolische und exkretorische Funktionen verantwortlich, die für die Entgiftung des Körpers und die Regelung des Flüssigkeits- und Säure-Basen-Haushalts wichtig sind. Chronisches Nierenversagen (auf Englisch chronic kidney disease [CKD] gennant) ist im Frühstadium meist asymptomatisch und wird erst diagnostiziert, wenn die glomeruläre Filtrationsrate (eGFR) unter 60 ml/min/1.73 m^2 fällt - das heißt, wenn etwa 50% der Nierenfunktion bereits verloren gegangen ist. CKD ist eine weit verbreitete Krankheit der alternden Bevölkerung, die sowohl für die Patienten als auch für Gesundheitssysteme eine erhebliche Belastung darstellt. Die Nierenforschung ist daher ein Bereich, indem molekulare epidemiologische Forschung neue Einblicke in die biologischen Mechanismen der Nierenfunktion und -erkrankung liefern kann: neue Biomarker, die Informationen über den Gesundheitszustand oder den Krankheitsverlauf liefern, können dadurch identifiziert werden.

In den letzten Jahrzehnten wurden zahlreiche Studien durchgeführt um Biomarker im Zusammenhang mit Nierenerkrankungen zu charakterisieren. Proteomische Biomarker im Urin werden in der klinischen Praxis verwendet (z. B. zur Vorhersage des Krankheitsverlaufs). Das komplexe proteomische Profil des Blutes ist jedoch noch relativ unerforscht und frühere Studien haben nicht zwischen Kausalität und Korrelation unterschieden. Mehrere Studien haben gezeigt, dass die im Blut gemessene epigenetische Altersbeschleunigung mit altersbedingten Krankheiten (unter anderem Gebrechlichkeit, kognitive Funktion und körperliche Fitness) und Sterblichkeit assoziiert sind. Dennoch wurde bisher nicht untersucht, ob epigenetische (DNAm-basierte) Prädiktoren für Alter und Sterblichkeit mit Parametern der geringen Nierenfunktion in Zusammenhang stehen.

Das Ziel dieser Dissertation ist es die Assoziation zwischen Nierenfunktion und

molekulare Biomarker (Plasmaproteinen und DNAm-basierten Prädiktoren für Alterung und/oder Sterblichkeit) in multiethnischen bevölkerungsbasierten Studien zu untersuchen und damit zum Verständnis der molekularen Grundlagen von Nierenfunktion und –erkrankung beizutragen. Dies wurde durch die Durchführung von zwei epidemiologischen Forschungsprojekten mit internationalen Kooperationspartnern erreicht, die zwei Hauptforschungsziele verfolgten. Das erste Ziel war es festzustellen ob Plasmaproteine mit eGFR und CKD assoziiert sind und zu beurteilen ob diese Assoziation kausal ist, das Zweite zu beurteilen ob epigenetische Altersbeschleunigung und DNAm-basierte Mortalitätsmarkers im Blut mit Nierenfunktion assoziiert sind.

Im Rahmen dieser Arbeit wurden 57 mit Nierenfunktion assoziierte Plasmaproteine identifiziert, davon scheint eines von kausaler Natur zu sein: eGFR hat einen positiven Effekt auf den Plasmaspiegel von Testican-2. Ebenso unterstreicht die starke und konsistente Assoziation verschiedener Nierenmerkmale, die eine schlechte Nierenfunktion widerspiegeln (z. B. niedrige eGFR, prävalente CKD, hoher Albumin-Kreatinin-Quotient, Albuminurie und hohes Serumurat), mit DNAm-basierten Prädiktoren des Alterns und/oder der Lebensspanne die allgegenwärtige Natur der altersbedingten chronischen Entzündung und des Alterns des Immunsystems.

Diese Arbeit präsentiert eine erste epidemiologische Einschätzung der Korrelation zwischen Nierenfunktion, proteomischen Biomarkern und epigenetischer Altersbeschleunigung in bevölkerungsbasierten Studien. Um die klinische Validität der hier identifizierten Proteine und „epigenetischen Uhren“ für die Krankheitsstratifizierung und -prognose abschließend zu beurteilen, ist weitere Forschung notwendig.

Table of Contents

List of Abbreviations

2SMR	Two-sample Mendelian Randomization
AKI	acute kidney injury
ASN	American Society of Nephrology
ATC	Anatomical Therapeutic Chemical (ATC) Classification System
BMI	body mass index
CCL14	C-C motif chemokine 14, protein
CKD	Chronic kidney disease
CKD-EPI	Chronic Kidney Disease Epidemiology Collaboration
CKD273	urinary proteome-based classifier predicting disease progression
CVD	cardiovascular disease
DAF	Complement decay-accelerating factor, protein encoded by gene *CD55*
DNA	deoxyribonucleic acid
DNAm	DNA methylation
DNAmAA	DNAmAge acceleration, calculated as the difference between an individuals' chronological age and DNAmAge
DNAmADM	DNAm-estimated levels of adrenomedullin

DNAmAge	DNAm-based measure of aging, also known as epigenetic age
DNAmPACKYRS	DNAm-estimated smoking pack years
DNAmPAI1	DNAm-estimated levels of plasminogen activator inhibitor-1
ECM	extracellular matix
EEAA	Extrinsic Epigenetic Age Acceleration
eGFR	estimated glomerular filtration rate
eQTL	expression quantitative trait loci
ESRD	end-stage renal disease
ESTHER	Epidemiologische Studie zu Chancen der Verhütung, Früherkennung und optimierten THerapie chronischer ERkrankungen in der älteren Bevölkerung
EWAS	epigenome-wide association study
FDR	False discovery rate
FWER	Family-wise error rate
GDF15	Growth Differentiation Factor 15, protein
GrimAA	Age acceleration as measured by the difference between GrimAge and chronological age
GTEX	Genotype-Tissue Expression (GTEx) project
GWAS	Genome-wide association study
HannumAA	Age acceleration as measured by the difference between Hannum's estimated DNAmAge and chronological age
HDL	high density lipoprotein
HorvathAA	Age acceleration as measured by the difference between Horvath's estimated DNAmAge and chronological age
HUNT	Nord-Trøndelag Health Study

IEAA	Intrinsic Epigenetic Age Acceleration
IGBFP6	Insulin Like Growth Factor Binding Protein 6, protein
INTERVAL	The "INTERVAL" study
IVW MR	Inverse-variance weighted Mendelian Randomization
JHS	Jackson Heart Study
K/DOQI	National Kidney Foundation's Kidney Disease Outcomes Quality Initiative
KORA	"Kooperative Gesundheitsforschung in der Region Augsburg", in English "Cooperative health research in the Region of Augsburg"
LD	Linkage disequilibrium
MCP	matricellular proteins
MP2K2	Dual specificity mitogen-activated protein kinase kinase 2, protein
MR	Mendelian Randomization
MRS	Mortality Risk Score
MS	Mass-spectrometry
NAS	Normative Aging Study
PAI-1	Plasminogen activator inhibitor-1
PEA	Proximity Extension Assay
PhenoAA	Age acceleration as measured by the difference between PhenoAge and chronological age
PPI	Protein-protein interaction network
pQTL	protein quantitative trait loci

QMDIAB	Qatar Metabolomics Study on Diabetes
RELT	Receptor Expressed In Lymphoid Tissues, also known as Tumor Necrosis Factor Receptor Superfamily Member 19L, protein
REML	Restricted Maximum Likelihood
RLS	restless leg syndrome
RNA	Ribonucleic acid
RNAi	RNA interference
RRT	renal replacement therapy
SD	standard deviation
SNP	single nucleotide polymorphism
SPARC	secreted protein acidic and rich in cysteine, protein family
SPOCK2	SPARC (Osteonectin), Cwcv And Kazal Like Domains Proteoglycan 2, gene name
TAJ	Tumor Necrosis Factor Receptor Superfamily Member 19, protein
TIMP-1	TIMP Metallopeptidase Inhibitor 1, protein
TNF SR-I	Tumor Necrosis Factor Soluble Receptor I, protein
TNF SR-II	Tumor Necrosis Factor Soluble Receptor II, protein
uACR	Urine Albumin-to-Creatinine Ratio
WHI	Women's Health Initiative

List of Figures

List of Tables

Preface

This document is a publication-based dissertation stemming from the doctoral research conducted by the candidate between October 2017 and April 2021 under the supervision of Prof. Dr. Juliane Winkelmann and Prof. Dr. Jerzy Adamski (Technical University of Munich, TUM) and mentorship of Dr. Melanie Waldenberger and Dr. Christian Gieger at the Research Unit Molecular Epidemiology from the Institute of Epidemiology (Helmholtz Zentrum München). This dissertation fulfills the criteria listed in §6 (2) "Regulations for the Award of Doctoral Degrees" effective January 1, 2014:

1. the present dissertation is based on two first-author papers accepted for publication in international peer-reviewed journals, listed in the section dedicated to the candidate's publication record;

2. the dissertation provides scientific background and formulates the research questions in an introductory section (Chapter 1), describes the analytical strategies used to address these questions in a methodology section (Chapter 2), presents the obtained results (Chapter 3), contextualizes them in relation to relevant literature in a discussion section (Chapter 4), and provides an outlook of how these findings may motivate future research (Chapter 5);

3. the dissertation includes a short summary of each publication and the individual contributions of the candidate (Chapter 3);

4. the original papers and written consent for their inclusion in this dissertation was provided by the publishers (Appendix A to Appendix C).

The overarching goal of this doctoral project was to identify potential -omic biomarkers associated with kidney function and related traits in well characterized and established human cohorts in order to offer novel insights into the molecular

mechanisms underlying kidney function and disease. To this end, comprehensive analyses of -omics data in multiple human cohort studies were conducted to investigate the association between kidney traits, plasma proteins, and DNA methylation (DNAm)-based measures of aging. These analyses resulted in the publication of two original research papers: one addressing the identification of proteomic biomarkers associated with estimated glomerular filtration rate (eGFR) and assessing the causal nature of the identified associations using data from genome-wide association studies (GWAS) (Matías-García et al., 2021b), and the second one exploring the association between five kidney traits (eGFR, chronic kidney disease [CKD], urinary albumin-to-creatinine ratio [uACR], microalbuminuria and serum urate) and seven DNAm-based predictors of aging and/or mortality (Matías-García et al., 2021a).

Publication Record

Articles in peer-reviewed journals

First-author contributions

This publication-based cumulative dissertation consists of the following peer-reviewed, first-author original publications:

- **Matías-García PR**, Wilson R, Guo Q, Zaghlool S, Eales J, Xu X, Charchar F, Dormer J, Maalmi H, Schlosser P, Elhadad M, Nano J, Sharma S, Peters A, Fornoni A, Mook-Kanamori D, Winkelmann J, Danesh J, Di Angelantonio E, Ouwehand W, Watkins N, Roberts D, Petrera A, Graumann J, Koenig W, Hveem K, Jonasson C, Köttgen A, Butterworth A, Prunotto M, Hauck S, Herder C, Suhre K, Gieger C, Tomaszewski M, Teumer A, Waldenberger M. Plasma Proteomics of Renal Function: A Trans-ethnic Meta-analysis and Mendelian Randomization Study. J Am Soc Nephrol. 2021 Jun 16:ASN.2020071070. doi: 10.1681/ASN.2020071070. Epub ahead of print. PMID: 34135082.

- **Matías-García PR**, Ward-Caviness CK, Raffield LM, Gao X, Zhang Y, Wilson R, Gào X, Nano J, Bostom A, Colicino E, Correa A, Coull B, Eaton C, Hou L, Just AC, Kunze S, Lange L, Lange E, Lin X, Liu S, Nwanaji-Enwerem JC, Reiner A, Shen J, Schöttker B, Vokonas P, Zheng Y, Young B, Schwartz J, Horvath S, Lu A, Whitsel EA, Koenig W, Adamski J, Winkelmann J, Brenner H, Baccarelli AA, Gieger C, Peters A, Franceschini N, Waldenberger M. DNAm-based signatures of accelerated aging and mortality in blood are associated with low renal function. Clin Epigenetics. 2021 Jun 2;13(1):121. doi: 10.1186/s13148-021-01082-w. PMID: 34078457.

Other first-author publications not included in this thesis are:

- **Matias-Garcia PR**, Wilson R, Mussack V, Reischl E, Waldenberger M, Gieger C, Anton G, Peters A, Kuehn-Steven A. Impact of long-term storage and freeze-thawing on eight circulating microRNAs in plasma samples. PLoS One. 2020 Jan 14;15(1):e0227648. doi: 10.1371/journal.pone.0227648. PMID: 31935258.

- Gorski M*, Jung B*, Li Y*, **Matias-Garcia PR***, Wuttke M, Coassin S, Thio CHL, Kleber ME, Winkler TW, Wanner V, Chai JF, Chu AY, Cocca M, Feitosa MF, Ghasemi S, Hoppmann A, Horn K, Li M, Nutile T, Scholz M, Sieber KB, Teumer A, Tin A, Wang J, Tayo BO, Ahluwalia TS, Almgren P, Bakker SJL, Banas B, Bansal N, Biggs ML, Boerwinkle E, Bottinger EP, Brenner H, Carroll RJ, Chalmers J, Chee ML, Chee ML, Cheng CY, Coresh J, de Borst MH, De-genhardt F, Eckardt KU, Endlich K, Franke A, Freitag-Wolf S, Gampawar P, Gansevoort RT, Ghanbari M, Gieger C, Hamet P, Ho K, Hofer E, Holleczek B, Xian Foo VH, Hutri-Kähönen N, Hwang SJ, Ikram MA, Josyula NS, Kähönen M, Khor CC, Koenig W, Kramer H, Krämer BK, Kühnel B, Lange LA, Lehtimäki T, Lieb W; Lifelines Cohort Study; Regeneron Genetics Center, Loos RJF, Lukas MA, Lyytikäinen LP, Meisinger C, Meitinger T, Melander O, Milaneschi Y, Mishra PP, Mononen N, Mychaleckyj JC, Nadkarni GN, Nauck M, Nikus K, Ning B, Nolte IM, O'Donoghue ML, Orho-Melander M, Pendergrass SA, Penninx BWJH, Preuss MH, Psaty BM, Raffield LM, Raitakari OT, Rettig R, Rheinberger M, Rice KM, Rosenkranz AR, Rossing P, Rotter JI, Sabanayagam C, Schmidt H, Schmidt R, Schöttker B, Schulz CA, Sedaghat S, Shaffer CM, Strauch K, Szymczak S, Taylor KD, Tremblay J, Chaker L, van der Harst P, van der Most PJ, Verweij N, Völker U, Waldenberger M, Wallentin L, Water-worth DM, White HD, Wilson JG, Wong TY, Woodward M, Yang Q, Yasuda M, Yerges-Armstrong LM, Zhang Y, Snieder H, Wanner C, Böger CA, Köttgen A, Kronenberg F, Pattaro C, Heid IM. Meta-analysis uncovers genome-wide significant variants for rapid kidney function decline. Kidney Int. 2021 Apr;99(4):926-939. doi: 10.1016/j.kint.2020.09.030. Epub 2020 Oct 31. PMID: 33137338. *** - denotes equal contribution**

Co-author contributions

Additional publications as co-author, which resulted from side-projects done during my doctoral research, are listed next:

- Zaghlool SB, Sharma S, Molnar M, **Matías-García PR**, Elhadad MA, Waldenberger M, Peters A, Rathmann W, Graumann J, Gieger C, Grallert H, Suhre K. Revealing the role of the human blood plasma proteome in obesity using genetic drivers. Nat Commun. 2021 Feb 24;12(1):1279. doi: 10.1038/s41467-021-21542-4. PMID: 33627659.

- Gomez-Alonso MDC, Kretschmer A, Wilson R, Pfeiffer L, Karhunen V, Seppälä I, Zhang W, Mittelstraß K, Wahl S, **Matias-Garcia PR**, Prokisch H, Horn S, Meitinger T, Serrano-Garcia LR, Sebert S, Raitakari O, Loh M, Rathmann W, Müller-Nurasyid M, Herder C, Roden M, Hurme M, Jarvelin MR, Ala-Korpela M, Kooner JS, Peters A, Lehtimäki T, Chambers JC, Gieger C, Kettunen J, Waldenberger M. DNA methylation and lipid metabolism: an EWAS of 226 metabolic measures. Clin Epigenetics. 2021 Jan 7;13(1):7. doi: 10.1186/s13148-020-00957-8. PMID: 33413638.

- Perna L, Zhang Y, **Matias-Garcia PR**, Ladwig KH, Wiechmann T, Wild B, Wal-denberger M, Schöttker B, Mons U, Ihle A, Kliegel M, Schwettmann L, Peters A, Brenner H. Subjective mental health, incidence of depressive symptoms in later life, and the role of epigenetics: results from two longitudinal cohort studies. Transl Psychiatry. 2020 Sep 21;10(1):323. doi: 10.1038/s41398-020-00997-x. PMID: 32958748.

- Elhadad MA, Jonasson C, Huth C, Wilson R, Gieger C, **Matias P**, Grallert H, Graumann J, Gailus-Durner V, Rathmann W, von Toerne C, Hauck SM, Koenig W, Sinner MF, Oprea TI, Suhre K, Thorand B, Hveem K, Peters

A, Waldenberger M. Deciphering the Plasma Proteome of Type 2 Diabetes. Diabetes. 2020 Dec;69(12):2766-2778. doi: 10.2337/db20-0296. Epub 2020 Sep 14. PMID: 32928870.

- McCartney DL, Min JL, Richmond RC, Lu AT, Sobczyk MK, Davies G, Broer L, Guo X, Jeong A, Jung J, Kasela S, Katrinli S, Kuo P-L, **Matias-Garcia PR**, Mishra PP, Nygaard M, Palviainen T, Patki A, Raffield LM, Ratliff SM, Richardson TG, Robinson O, Soerensen M, Sun D, Tsai PC, van der Zee MD, Walker RM, Wang X, Wang Y, Xia R, Xu Z, Yao J, Zhao W, Correa A, Boerwinkle E, Dugué PA, Durda P, Elliott HR, Gieger C, de Geus EJC, Harris SE, Hemani G, Imboden M, Kähönen M, Kardia SLR, Kresovich JK, Li S, Lunetta KL, Mangino M, Mason D, McIntosh AM, Mengel-From J, Moore AZ, Murabito JM, Ollikainen M, Pankow JS, Pedersen NL, Peters A, Polidoro S, Porteous DJ, Raitakari O, Rich SS, Sandler DP, Sillanpää E, Smith AK, Southey MC, Strauch K, Tiwari H, Tanaka T, Tillin T, Uitterlinden AG, Van Den Berg DJ, van Dongen J, Wilson JG, Wright J, Yet I, Arnett D, Bandinelli S, Bell JT, Binder AM, Boomsma DI, Chen W, Christensen K, Conneely KN, Elliott P, Ferrucci L, Fornage M, Hägg S, Hayward C, Irvin M, Kaprio J, Lawlor DA, Lehtimäki T, Lohoff FW, Milani L, Milne RL, Probst-Hensch N, Reiner AP, Ritz B, Rotter JI, Smith JA, Taylor JA, van Meurs JBJ, Vineis P, Waldenberger M, Deary IJ, Relton CL, Horvath S, Marioni RE. Genome-wide association studies identify 137 loci for DNA methylation biomarkers of ageing. bioRxiv. 2020. 2020.2006.2029.133702. doi:10.1101/2020.06.29.133702 – manuscript accepted at Genome Biology

- Hawe JS, Wilson R, Schmid K, Zhou Li, Lakshmi L, Lehne BC, Kühnel B, Scott WR, Wielscher M, Yew YW, Baumbach C, Lee DP, Marouli E, Bernard M, Pfeiffer L, **Matías-García P**, Autio MI, Bourgeois S, Herder C, Karhunen V, Meitinger T, Prokisch H, Rathmann W, Roden M, Sebert S, Shin J, Strauch K, Zhang W, Tan WLW, Hauck SM, Merl-Pham J, Grallert H, Barbosa EGV, MuTHER Consortium, Illig T, Peters A, Paus T, Pausova Z, Deloukas P, Foo RSY, Jarvelin MR, Kooner JS, Loh M, Heinig M, Gieger C, Waldenberger M, Chambers J. Genetic variation influencing DNA methy-

lation provides new insights into the molecular pathways regulating genomic function. 2021. – manuscript accepted at Nature Genetics

- Marttila S, Viiri L, Mishra PP, Kühnel B, **Matias-Garcia PR**, Lyytikäinen LP, Ceder T, Mononen N, Rathmann W, Winkelmann J, Peters A, Kähönen M, Hutri-Kähönen N, Juonala M, Aalto-Setälä K, Raitakari O, Lehtimäki T, Waldenberger M, Raitoharju E. Methylation status of nc886 epiallele reflects periconceptional conditions and is associated with glucose metabolism through nc886 RNAs. 2021. – manuscript accepted at Clinical Epigenetics

Conference contributions

- **Matías-García PR**, Sharma S: Progress report on human plasma proteomic signatures of obesity and kidney disease, talk presented at the 9th Grainau Workshop of Genetic Epidemiology, Grainau, Germany, April 2019

- **Matías-García PR**, Raffield L, Gao X, Zhang Y, Horvath S, Lu A, Brenner H, Schöttker B, Batorsky A, Colicino E, Shen J, Just AC, Nwanaji-Enwerem JC, Coull B, Lin X, Vokonas P, Zheng Y, Hou L, Schwartz j, Baccarelli AA, Wilson R, Gieger C, Franceschini N, Waldenberger M: Epigenetic age acceleration and renal function, poster and 'lightning talk' presented at the Conference Epigenomics of Common Diseases, Hinxton, UK, November 2019

- **Matías-García PR**, Wilson R, Guo Q, Zaghlool SB, Sharma S, Schlosser P, Köttgen A, Peters A, Mook-Kanamori DO, Graumann J, Koenig W, Hveem K, Jonasson C, Butterworth A, Suhre K, Gieger C, Teumer A, Waldenberger M: Insights into the proteomics of renal function: a trans-ethnic meta-analysis and Mendelian randomization study, e-poster presented at the virtual conference of the European Society of Human Genetics 2020, online, June 2020

Book chapters

- **Matías-García PR**, Martinez-Hurtado JL. Kidney Smartphone Diagnostics. Methods Mol Biol. 2018;1735:487-498. doi: 10.1007/978-1-4939-7614-0_36.

PMID: 29380339.

- **Matías-García PR**, Martinez-Hurtado JL, Beckley A, Schmidmayr M, Seifert-Klauss V. Hormonal Smartphone Diagnostics. Methods Mol Biol. 2018;1735:505-515. doi: 10.1007/978-1-4939-7614-0_38. PMID: 29380341.

1 Introduction

This chapter will introduce the reader to principles of observational epidemiology and large-scale profiling of molecular -omic data with a focus on population-based studies (Figure 1.1). In order to outline the findings corresponding to the work on DNAm-based aging and mortality predictors, details on the construction and interpretation of the so-called "epigenetic clocks" are presented (Figure 1.2 to Figure 1.5). As this dissertation describes work on proteomic biomarker discovery, a state of the art analytical method to obtain proteomics data is shown in Figure 1.6. This chapter will also offer a brief review of concepts and current knowledge on the (molecular) epidemiology of chronic kidney disease (CKD) and describe the knowledge gaps this doctoral research aimed to narrow. Finally, the objectives of this thesis will be outlined with regard to the concepts and ideas covered in this chapter, and these aims will be linked to the molecular epidemiological studies described in both included first-author publications (Matías-García et al., 2021a; Matías-García et al., 2021b).

1.1 Principles of (Molecular) Epidemiology

Molecular epidemiology, defined in the 2005 edition of the Handbook of Epidemiology as "the application of the techniques of molecular biology to the study of populations, with a particular focus on the investigation of disease" (Vineis et al., 2005), can be understood as a field of epidemiology characterized by the use of biological markers – biomarkers – to better characterize exposures, susceptibility of disease and health outcomes (Khoury et al., 2008). Molecular epidemiology aims to offer knowledge useful for health research and clinical practice by investigating biomarkers involved in disease etiology, intermediate phenotypes and pathological events, as well as pre/clinical disease and prognosis (N. Rothman et al., 2011).

1.1.1 Use of biomarkers

The use of biological molecules in epidemiological studies is not a novelty on its own: classical examples are the determination of antibody titers when investigating exposure to infectious disease, or blood lipids in cardiovascular research. As once health outcomes were based on conventional biomarkers like estrogen receptor status in breast cancer, nowadays -omic biomarker research can assess specific gene expression signatures and relate them to disease prognosis (Khoury et al., 2008). Epidemiological research identifies three groups of biomarkers: biomarkers of exposure and/or dose, of a biological effect (e.g. tissue damage, molecular dysfunction, early pathophysiological events) and of susceptibility (Schulte et al., 2011). Other sources suggest this classification can be extended to a continuum of biomarkers reflecting multi-stage molecular events ranging from exposure to initiation and progression of disease, as well as individual susceptibility markers (Garcia-Closas et al., 2011).

1.1.2 Study design and aims

In general, epidemiology is conducted based on either experimental or observational studies: the first are characterized by the random designation and administration of interventions to groups of individuals in order to assess the effect on a given outcome (e.g. clinical trials), whereas the second are conducted in settings in which due to ethical or feasibility reasons such allocation of an intervention is not done, but the occurrence of exposures and outcomes is registered (Khoury et al., 2008).

Molecular epidemiology is especially well-suited to contribute to improving the definition of exposures in settings with low expected exposure level and/or multiple sources of exposure; to offer novel insights into the dynamics of gene-environment interactions; and to bring forward the identification of markers of early response to treatment (Vineis et al., 2005). To achieve these aims, research must be conducted following a clear workflow, starting with an appropriate study design that will allow for the generation of good quality -omics data and an appropriate analysis thereof (Franks & Pomares-Millan, 2020) (Figure 1.1).

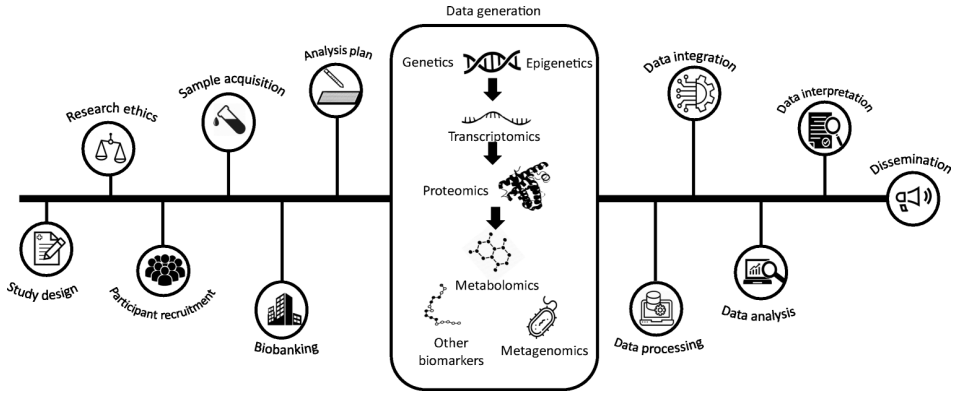

Figure 1.1: Workflow of -omics studies, reproduced from (Franks & Pomares-Millan, 2020)

An extensive description of general principles of epidemiological study design is beyond the scope of this thesis. Nevertheless, as observational epidemiology studies are of relevance to the work presented in this dissertation, the following paragraphs will provide additional details on their execution, advantages and caveats in the context of molecular epidemiology.

Observational epidemiologic studies

In this dissertation, data from individuals participating in population-based cohorts were used to analyze correlations between –omic molecular phenotypes and parameters reflecting kidney function. This section will therefore focus on study designs in which the unit of observation is the individual, unlike ecological studies in which the unit of observation is a population or a community. Rothman and colleagues distinguish four types of observational studies in their chapter dedicated to types of epidemiologic studies in Modern Epidemiology: cohort, case-control, cross-sectional and ecologic studies (K. Rothman et al., 2008).

Cohort studies identify a source population and classify their individuals based on their exposure status to then evaluate the occurrence of disease. Cohort studies aim at assessing disease prevalence and incidence, are useful to understand the natu-

ral history of disease and risk factors, and produce data on environmental (including lifestyle and other exposures) factors (K. Rothman et al., 2008). The establishment of longitudinal cohort studies, where individuals recruited to a cohort are followed over time, offers the possibility of conducting repeated sampling of health outcomes and molecular phenotypes, which may be predecessors or consequences of disease (Khoury et al., 2008).

Case-control studies select individuals based on their disease status and identify appropriate controls independently from their exposure status from one shared source population (either individuals in a hospital registry for hospital-based studies or a sample population for population-based studies) (Vineis et al., 2005). Case-control studies can be used to study rare diseases and their genetic architecture, as genetic information does not change with time, as well as gene-gene and gene-environment interactions, and their effect in disease (Khoury et al., 2008). However, some of the limitations of this study design include the potential for selection bias (where cases all stem from a similar geographic region and the distribution of co-morbidities and other diseases in control individuals may not be representative of the general population) and bias coming from differential participation from cases and controls (Garcia-Closas et al., 2011).

Cross-sectional studies select a representative sample of individuals from a source population independently from their disease or exposure status, and assess both exposure and disease at the same point in time (Vineis et al., 2005). These studies offer a "snapshot" of events in the population (Garcia-Closas et al., 2011), and are conducted to assess exposures and to explore the correlations between health outcomes and molecular phenotypes (Khoury et al., 2008). However, this study design does not distinguish between causal information on the incidence of disease and the natural history of disease, and additional considerations are necessary to use data produced by such a study to offer causal insights (Hernan, 2018; Khoury et al., 2008).

1.2 -Omics

Khoury and colleagues were right when they wrote in 2008 that "the use of biomarkers in epidemiology will reach a new level of complexity, with the simultaneous study of hundreds or even thousands of data points for each person" (Khoury et al., 2008). Such data points are collected and studied by the omic sciences, whose –omics suffix is derived from the Ancient Greek. **Genomics** was a term coined over thirty years ago to describe the then-emerging discipline merging molecular and cell biology with genetics and computational science that was to be dedicated to mapping/sequencing genes and analyzing data stemming thereof – and was used first to name the journal Genomics (McKusick & Ruddle, 1987). The term is now used to describe the study of the functions and interactions of all the genes that constitute the genome, the advent of genomic medicine and its potential applications to prevention, diagnosis and therapy (Guttmacher & Collins, 2002). The possibility to measure additional biological molecules with improved resolution and at large-scale allowed for other –omic disciplines to emerge and be applied to human studies to assess gene expression, molecular products thereof, and their interaction with the environment (Figure 1.1) (Franks & Pomares-Millan, 2020; Hasin et al., 2017). The following sections will further introduce the reader to three –omic disciplines relevant for this thesis: **genomics**, **epigenomics** and **proteomics**.

1.2.1 Genomics

The field of genomics, the most mature –omic discipline (Hasin et al., 2017), has rapidly moved forward since the completion of the mapping of the human genome to the possibility of quickly sequencing entire individual genomes nowadays, and presents its own set of ethical, technical and biological challenges going into the future (McGuire et al., 2020). Genome-wide association studies (GWAS) are conducted to assess the correlation between the frequency of certain genetic variants and a phenotype of interest (e.g. a marker of disease, progression or response to treatment). GWAS have produced knowledge on variants with a causal role either in disease initiation or involved in pathophysiological mechanisms, thus offering insights into the genetic architecture of disease susceptibility; likewise, novel pharmacological targets and disease biomarkers have also been identified as a product of these studies and integrated into clinical practice (Tam et al., 2019). GWAS

have also proven to be helpful in identifying replicable genetic variant-trait associations and insights into biological mechanisms despite the relatively small effect sizes identified in these studies (Tam et al., 2019).

Genetic variants identified in GWAS may not necessary be causal themselves, but are rather in linkage disequilibrium (LD) (i.e. alleles from different loci are associated in a non-random manner) with those with a causal role, so additional wet-lab experiments are necessary to describe precise causal mechanisms (Tam et al., 2019). Nevertheless, data on the association between genetic variants and traits produced by GWAS can be used together with other types of data and analytical methods to produce additional insights beyond gene-trait associations: SNP heritability can be estimated to outline the genetic architecture of disease, polygenic risk scores to detect pleiotropy and validate GWAS discoveries, or Mendelian randomization to assess the causal nature of associations between phenotypes (Visscher et al., 2017). An introduction to the principles, conduction and limitations of Mendelian randomization is presented in Chapter 2.

1.2.2 Epigenomics: DNA methylation

Epigenetics, initially described as the causal interactions between genes and gene-products allowing for the expression of phenotypic traits, is now better understood as the (heritable) regulation of gene expression independent to changes in DNA sequence (Tollefsbol, 2011). Three central mechanisms through which epigenetic mechanisms regulate gene expression have been defined, namely: **DNA methylation** (**DNAm**), chromatin modifications and non-coding RNA participating in RNA interference (RNAi) mechanisms (Figure 1.2).

DNAm, the most extensively studied epigenetic mechanism, is a process in which certain DNA regions, usually cytosine-guanine dinucleotides (CpG sites) undergo chemical modification by having a methyl group transferred (if removed, the process is called de-methylation). DNAm is a main player in the regulation of gene expression, it is involved in gene imprinting and X chromosome inactivation (Suzuki & Bird, 2008), as well as in cellular differentiation (Tollefsbol, 2011). Likewise, the role of DNAm in numerous conditions and diseases (e.g. cancer, cardiovascular and

metabolic diseases, autoimmune disorders) has been shown (Kulis & Esteller, 2010; Morales-Nebreda et al., 2018; Muka et al., 2016).

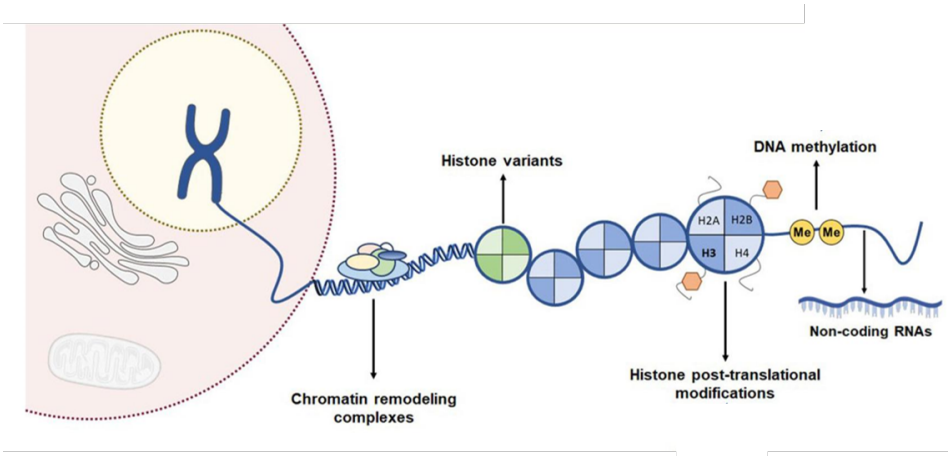

Figure 1.2: Epigenetic mechanisms, modified from (Miranda-Gonçalves et al., 2018)

A variety of methods exists to detect and quantify DNAm levels, and depending on the research question and aims, some techniques may be more appropriate than others. Although a review of the different DNAm profiling methods and considerations on profiling specific regions versus genome-wide methylation patterns is beyond the scope of this thesis, the reader may wish to refer to (Hattori & Ushijima, 2011) for a detailed account on the analytical methods available to profile genomic regions, as well as their strengths and limitations. Methods to assess genome-wide methylation patterns, mostly based on principles used in assessment of gene-specific methylation, have been coupled with microarray technology, sequencing and/or cloning principles (Rauch & Pfeifer, 2011). A relevant example of the microarray technology is the Illumina HumanMethylation 450K beadchip (Bibikova et al., 2011), a high-throughput platform offering information on DNAm levels on the single nucleotide level that has been widely used in population-based epidemiologic studies.

A relatively recent development in population-based epigenomics was the development of "**epigenetic clocks**" based on genome-wide methylation patterns

measured using the aforementioned microarray technology. The next section will introduce the reader to the concept of "epigenetic clock", present how to interpret estimates of epigenetic aging produced by such DNAm-based predictors and provide details on the most widely used predictors in research.

Epigenetic clocks and epigenetic age acceleration

The "epigenetic clocks" are algorithms meant to predict age and lifespan based on DNAm levels of specific (and largely non-overlapping) sets of CpGs (Liu et al., 2020). Of note, although the term "epigenetic clocks" and "epigenetic aging" have been widely used in the literature, "DNAm-based aging predictors" and "DNAm-Age" seem to be a more appropriate terms to describe the algorithms and produced estimates thereof, considering these measures of aging are based solely on DNAm and that by definition measures of epigenetic age could also be based on changes in chromatin modifications or RNAi.

Importantly, although different measures of DNAm-based aging have been proposed, their derivation is similar: penalized regression algorithms are trained on data from the sample donors using aging (or some measure thereof), which results in the selection of specific CpG dinucleotides. These are assigned weights based on how much they change with age, and are then combined in a linear manner using these weights to produce a numerical estimate of epigenetic age (Field et al., 2018; Horvath & Raj, 2018).

Figure 1.3 shows the elegant summary presented by Field and colleagues on how these DNAm-based aging estimates are produced (Field et al., 2018). While negative coefficients, shown in blue, indicate that methylation decreases with age, positive coefficients shown in red suggest increasing methylation with age. Furthermore, CpGs may have different rates of change, where darker colors represent faster aging rates (panel A from Figure 1.3). Panel B shows the calculation of three individuals' epigenetic age (one individual per row) based on the methylation levels of 8 CpGs (one CpG per column) included in a fictitious clock for demonstration purposes: the methylation value at each of the eight CpG sites is color coded as white-to-black circles, and the weight assigned to each CpG (i.e. the coefficient

obtained in the penalized regression algorithm) is shown in the red/blue squares; the products of the methylation value and the weight at each CpG, shown as a numerical value in the top right corner of each square, are added to obtain the predicted DNAm-based age estimate (**DNAmAge**) for each individual (Figure 1.3) (Field et al., 2018).

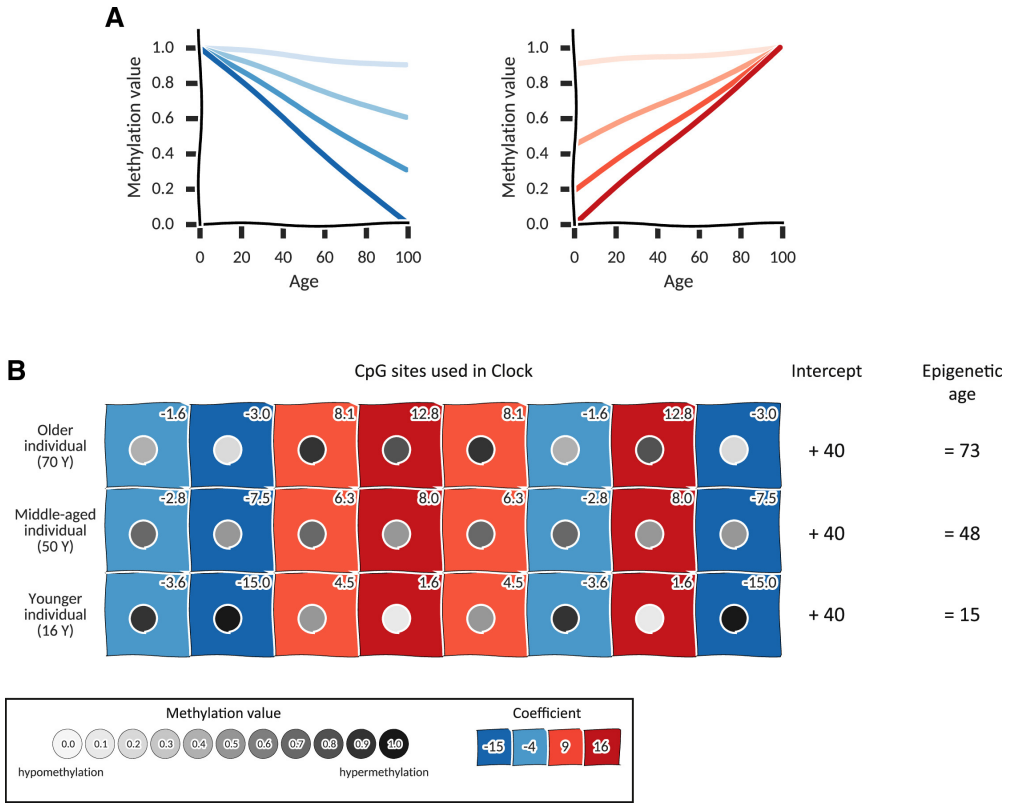

Figure 1.3: Derivation of epigenetic clocks, adapted from (Field et al., 2018)

Measures of biological aging are considered to be a proxy of an individual's physiological status and thus deviations from chronologic age (measured by the number of calendar years lived) may reflect a faster (acceleration) or slower (deceleration) aging rate (Figure 1.4) (Yu et al., 2020). In this regard, the existence of discrepancies between DNAmAge and chronological age, a phenomena known as **epigenetic age acceleration**, is interesting as it might reflect physiological dysregulation and

premature aging (Dhingra et al., 2018; Horvath & Raj, 2018). Evidence suggests that epigenetic age acceleration reflects biological aging beyond chronological age, capturing a "cumulative epigenetic drift that represents a multifactorial degenerative process across tissues and organisms" (Lim & Song, 2018). However, whether epigenetic aging plays a role in the mechanistic pathophysiology of disease remains to be elucidated (Dhingra et al., 2018; Jung & Pfeifer, 2015).

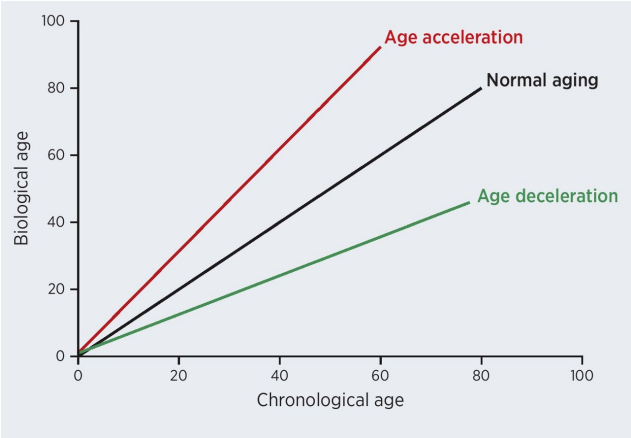

Figure 1.4: Age acceleration/deceleration as discrepancies between biological and chronological age, reproduced from (Yu et al., 2020)

Epigenetic clocks, depending on their construction, reflect information from chronological and/or biological aging processes. Figure 1.5 shows that, on the one hand, chronological clocks (i.e. clocks trained with chronological age as its outcome) are meant to capture elapsed time since birth and thus do not reflect biological aging. In that sense, the correlation between DNAmAge and chronological age from individuals experiencing accelerated biological aging (shown in orange) and those with a slower aging process (shown in blue) is expected to be perfect. On the other hand, biological clocks (i.e. clocks trained using some measure of biological aging as outcome) do identify individuals with accelerated biological aging, although at the expense of their accuracy as chronological clocks. Finally, although hybrid clocks do capture chronological age to a large extent, deviations from the expected DNAmAge estimates are thought to reflect biological age (Field et al., 2018).

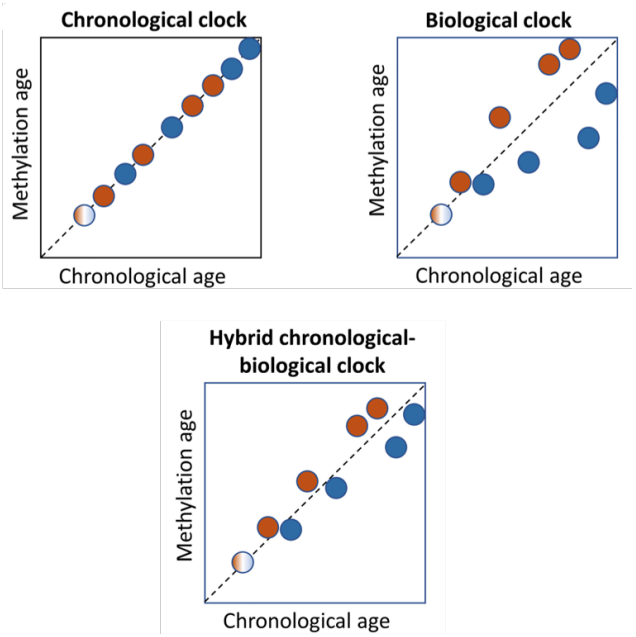

Figure 1.5: Chronological, biological and hybrid clocks, modified from (Field et al., 2018)

Two of the earliest measures of epigenetic age were proposed by **Hannum** (Hannum et al., 2013) and **Horvath** (Horvath, 2013). The first was developed using chronological age as outcome and was based solely on DNAm data from whole blood, identifying 71 CpGs (Hannum et al., 2013). Hannum's clock captures age-related changes in blood cell composition (Marioni et al., 2015) and is a good mortality predictor (Chen et al., 2016). Horvath's epigenetic clock was also derived to predict chronological age and resulted in the selection of 353 CpGs. Nevertheless, this clock extended its applicability to multiple tissues and the entire lifespan, as DNAm data from 30+ different tissues donated by both adults and children was used in its construction, and thus this estimator is sometimes referred to as a pan-tissue epigenetic clock (Horvath, 2013). The estimated DNAm-based age measure, DNAmAge or epigenetic age, greatly correlates with chronological age across tissues and samples (Horvath, 2013), although deviation from this is thought to reflect biological age (Horvath & Raj, 2018). As noted by (Field et al., 2018), both Hannum's and Horvath's clocks are likely "hybrid" clocks, that is, clocks calibrated against biological age that cannot accurately reflect biological age by design (Figure 1.5),

and thus show a poorer performance than other clocks in predicting disease and response to the environment (Field et al., 2018; Quach et al., 2017).

These clocks, often referred to as the "first-generation" DNAm-based predictors, have been refined and improved in order to come up with measures that are either independent of blood cell composition and mirror cell-intrinsic ageing-related changes (such as the **intrinsinc epigenetic age acceleration**, **IEAA**) or that leverage age-related changes in blood cell composition, hence representing a measure for immune aging (**extrinsic age epigenetic age acceleration**, **EEAA**) (Chen et al., 2016; Quach et al., 2017). Such measures and others derived from similar approaches have been used to analyze the relationship between epigenetic age and ageing-related phenotypes and mortality, producing numerous findings (Dhingra et al., 2018; Horvath & Raj, 2018).

A "second generation" of DNAm-based age predictors aimed at better capturing changes in physiological regulation by incorporating composite surrogate markers of biological age (e.g. clinical biomarkers and related traits) as outcome measures (Horvath & Raj, 2018). **PhenoAge** was developed based on a measure of "phenotypic age" – that is, a variable built considering 10 biological parameters and risk factors: chronological age, albumin, creatinine, glucose and C-reactive protein levels, lymphocyte percentage, mean cell volume, red blood cell distribution width, alkaline phosphatase levels and white blood cell count (Levine et al., 2018). This algorithm resulted in the selection of 513 CpGs, which when used to estimate DNAmAge, outperform the "first generation" of clocks in relation to predicting lifespan and cardiovascular disease (Levine et al., 2018). Moreover, this epigenetic clock reflects DNAm changes related to tobacco exposure, a significant driver of mortality-associated DNAm changes (Zhang et al., 2017), unlike those of Hannum and Horvath. However, the use of this clock is limited to blood and adult-derived samples (Horvath & Raj, 2018).

GrimAge, a mortality predictor based on plasma protein levels associated with mortality outcomes, was also recently developed following a two-stage procedure (Lu et al., 2019). First, DNAm-based surrogate biomarkers of smoking pack-years and proteins associated with mortality or morbitidy (i.e. adrenomedullin,

C-reactive protein, plasminogen activator inhibitor-1 [PAI-1]), growth differentiation factor 15 [GDF15], b2-microglobulin, leptin, tissue inhibitor metalloproteinase 1 [TIMP-1]) were constructed. This was followed by the regression of time-to-death due to all-cause mortality on these DNAm-based surrogate biomarkers. The age-adjusted version of DNAm GrimAge, a measure equivalent to epigenetic age acceleration, outperformed other clocks in relation to prediction of age-related disease, its association with poor levels of known clinical biomarkers and with adverse findings from computed tomography data (Lu et al., 2019).

Another interesting DNAm-based lifespan marker, developed solely based on mortality data, is the 10-CpG epigenetic **mortality risk score** (**MRS**) (Zhang et al., 2017). Recently validated (X. Gao, Colicino, et al., 2019), this measure can also be used to define risk levels based on the total number of "aberrantly" methylated CpG sites identified; these are defined by the cut-offs derived from the 1st quartile from the nine CpGs negatively correlated with mortality and the 4th quartile of single CpG positively correlated with mortality as defined in (Zhang et al., 2017). Individuals are then assigned to one of three mortality risk levels based on the total number of "aberrantly" methylated CpGs: low risk, MRS = 0–1; moderate risk, MRS = 2–5; and high risk, MRS >5. This measure allows for the identification of individuals with higher risk of death due to cancer and cardiovascular disease (Zhang et al., 2017).

1.2.3 Proteomics

Proteomics, the discipline dedicated to identification, quantification and analysis of proteins, has also benefited from recent technological developments that have made it possible to simultaneously profile hundreds of proteins at the population level. Proteins circulating in plasma, known as the plasma proteome, reflect the systemic physiological status of an individual, as it is a product of various tissues (Uhlén et al., 2019) and thus represents a very complex matrix of clinical interest (Ping et al., 2005). Moreover, it is also informative of certain lifestyle exposures (e.g. medication intake), as well as genetic predisposition to disease and early disease (Geyer et al., 2016). Additional to its diversity in relation to protein function,

the plasma proteome is also characterized by a wide range of protein abundance, where the reference range of known proteins covers >11 orders of magnitude (J. G. Smith & Gerszten, 2017). Moreover, the most abundant proteins (e.g. albumin and immunoglobulins) constitute up to 99% of the protein mass measured in blood, thus posing a technical challenge for its profiling (Tirumalai et al., 2003). Proteins in the lower abundance end of this spectrum may result from active secretion or tissue leakage (either due to pathological mechanisms or regular cell turnover), and may thus be of particular interest to clinical research (Uhlén et al., 2019). Given the nature of the plasma proteome and the challenges its profiling presents, assessing the proteome at a scale feasible in population-based studies has been possible only in recent years (Hasin et al., 2017; Zanini et al., 2020).

Methods of proteomic profiling

Methods to profile proteomics can be classified as belonging to **mass-spectrometry** (MS) or to **affinity-based assays** (Zanini et al., 2020). MS, either in its targeted or untargeted design, allows for the identification of post-translational modifications, but its implementation in large-scale studies is hindered by its labor-intensive nature and its limited coverage of moderate-to-abundant proteins only (Zanini et al., 2020). Affinity assays using antibodies to capture specific proteins and quantify them based on signal intensity of reporter antibodies (e.g. enzymatic antibody-labeled in the case of enzyme-linked immunosorbent assays, ELISAs) are the gold-standard in current clinical practice (J. G. Smith & Gerszten, 2017). However, these assays are limited in their detection range due to their poor sensitivity to detect low-abundance proteins (Fulwyler & McHugh, 1990), in their multiplexing because of issues with cross-reactivity between reagents (Ellington et al., 2010), and in its flexibility due to the costs associated to development of new antibodies (J. G. Smith & Gerszten, 2017).

The use of oligonucleotide-based aptamers, which are nucleotide sequences that fold and interact with proteins with high-binding affinity (Gold et al., 2010), is one of the most developed alternatives to affinity-based methods using antibodies. A recent review on emerging technologies for large scale proteomic profiling highlights

this method for its higher multiplexing possibilities and sample throughput, as well as the ease of reagent development (J. G. Smith & Gerszten, 2017). Figure 1.6 shows the workflow from SOMAscan, a commercial platform using aptamers of 40 nucleotides to conduct relative quantification of protein levels in blood. In brief, proteins interact with and are captured by the bead-fixed aptamers to be then biotin labeled, followed by a step in which the aptamers themselves are released from the beads. These protein-aptamer constructs are captured again on streptavidin beads, after which the aptamers are eluted and can be then analyzed using a microarray-based DNA quantification (J. G. Smith & Gerszten, 2017). Of note, although high specificity and reproducibility has been reported for this platform, unspecific binding has been observed in some cases and it is recommended that results are verified using other analytical methods. A more detailed description on the technical details of this assay and its use in this work is offered in the first-author publication (Matías-García et al., 2021b).

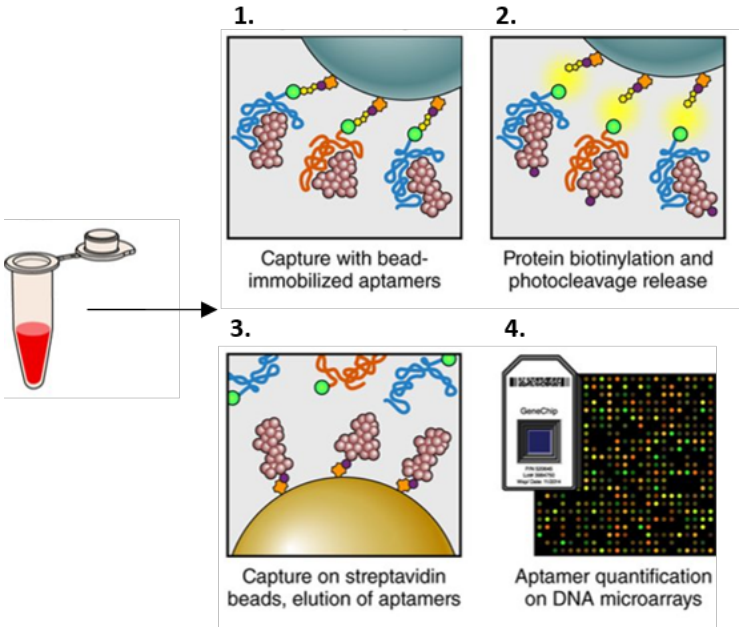

Figure 1.6: Workflow from aptamer-based proteomic platform, modified from (J. G. Smith & Gerszten, 2017)

1.3 Kidney function and disease

The kidneys are two fist-sized organs that, despite their relative small size, perform excretory, endocrine and metabolic functions that are central to maintaining homeostasis (Levin et al., 2013). Kidney damage is a term that broadly describes the identification of renal abnormalities that may happen before reduction in kidney function, but may not be informative of the disease etiology (Levin et al., 2013).

1.3.1 Measures of kidney function and damage

Kidney function, specifically their ability to filter blood, is most often assessed in clinical settings by estimating **glomerular filtration rate** (**eGFR**) based on blood levels of serum creatinine, an endogenous filtration marker (Eckardt et al., 2013; Levey et al., 2009). The estimation of GFR using alternative filtration markers, like cystatin C, has been recommended in cases where GFR estimation has to be more accurate. For example, this is recommended whenever GFR estimates will lead clinical decisions, as in the case of patients with low serum creatinine-based eGFR (45-59 ml/min/1.73 m^2) but no additional markers of kidney damage who thus require further confirmation of kidney disease to have a referral to nephrological care (Inker et al., 2014; Levin et al., 2013). Although GFR represents only one dimension of kidney function (namely, excretory function), it has been recognized as "the best overall measure of kidney function in health and disease" ("K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification", 2002; Levin et al., 2013). Decreased GFR is defined as <60 ml/min/1.73 m^2, a threshold indicating less than half of the average values observed in young adults (125 ml/min/1.73 m^2), and GFR <15 ml/min/1.73 m^2 is considered as kidney failure (Levin et al., 2013).

Beyond eGFR, markers of kidney damage include **proteinuria** due to increased glomerular permeability (albuminuria), incomplete tubular reabsorption of proteins (tubular proteinuria) or higher levels of small proteins circulating in plasma (Levin et al., 2013). Although proteinuria is associated with progression to ESRD and mortality independently from changes in eGFR (Astor et al., 2011; Matsushita et al., 2010), both eGFR and albuminuria seem to be complementary markers (Inker et al., 2014). The use of **urinary albumin-to-creatinine ratio** (**uACR**), a marker

of kidney injury determined in spot urine samples, has been more emphatically recommended: uACR has been observed to be a more specific and sensitive marker of kidney damage than albuminuria (Inker et al., 2014), and thus may be useful in identifying early kidney damage preceding eGFR decline (for example, in diabetic nephropathy) (Matsushita et al., 2010). A third biomarker of low eGFR is **serum urate**, a by-product of purine metabolism disposed of by the kidneys; higher levels of this marker are associated with cardiovascular and kidney disease (Joosten et al., 2020).

1.3.2 CKD: Staging and definition

Chronic kidney disease (**CKD**), defined by anomalies in kidney function or structure lasting more than 3 months, is either characterized by reduced eGFR (<60 ml/min/1.73 m^2) or >1 markers of kidney damage (Inker et al., 2014). A 5-stage system based on GFR levels was initially proposed in 2002 ("K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification", 2002), and it was later updated to better reflect disease prognosis by incorporating three albuminuria categories and etiology of disease (hinted by location of findings or presence of systemic disease) to these stages (Figure 1.7) (Inker et al., 2014). Figure 1.7 shows low risk categories in green, where CKD is not present in the absence of additional markers of kidney damage; categories with moderately increased risk for disease progression are shown in yellow, followed by categories in orange with high and very high risk in red (Inker et al., 2014).

1.3.3 Epidemiology of CKD

CKD has a global **prevalence** of 10-16% and is expected to be an increasingly prevalent noncommunicable disease in aging populations (Eckardt et al., 2013; Hill et al., 2016; Levey et al., 2007). Although decreased GFR correlates with the progressive reduction of other kidney functions in CKD (Levin et al., 2013), it follows widespread structural damage – changes in serum creatinine are not evident until 50% of the renal filtration function is lost (Mischak et al., 2015). The existence of a blind spot for early renal disease detection is thus obvious, making early CKD an increasingly prevalent silent disease (Sanchez-Nino et al., 2017). Moreover, there is a lack of therapeutic interventions for CKD. The only available treatment for kid-

ney disease in its final stage, end-stage renal disease (ESRD), is renal replacement therapy (RRT) either as renal transplantation or dialysis; while access to RRT is already currently not available for all patients in need of it, the number of people needing RRT will double by 2030 in comparison to the data reported in 2010 (Liyanage et al., 2015).

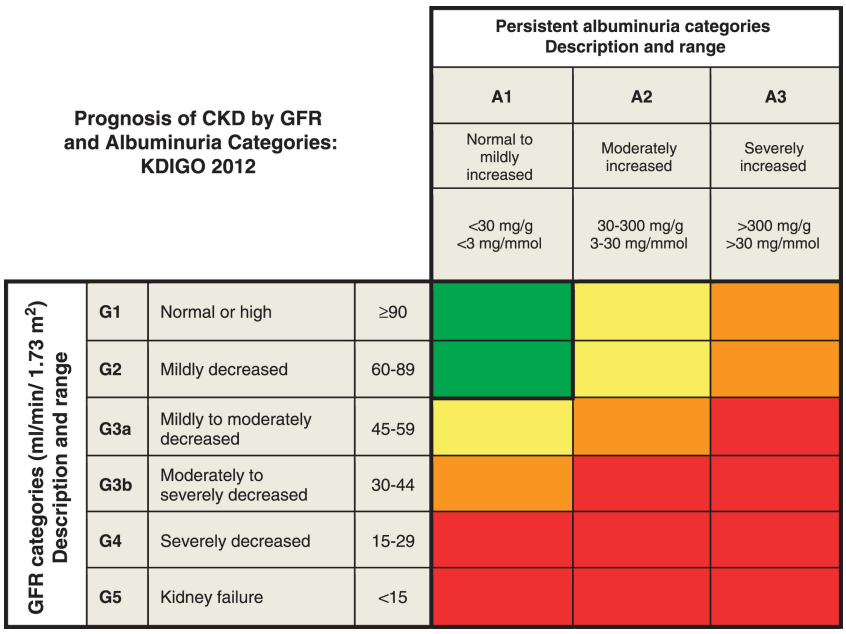

Figure 1.7: Prognosis of CKD by GFR and albuminuria categories, reproduced from (Inker et al., 2014)

The increasing CKD prevalence, over and above the lack of therapeutic interventions and the limited access to end-stage therapy, present a significant **global burden** on both individuals and governmental health budgets (Eckardt et al., 2013; Levey et al., 2007; Sanchez-Nino et al., 2017). CKD is one of the leading causes of mortality, especially in countries with limited access to treatment and highly prevalent risk factors (e.g. diabetes and hypertension), and a significant contributor to morbidity worldwide ("Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013", 2015; Rhee & Kovesdy, 2015).

1.3.4 Risk factors and outcomes

Several **risk factors** for kidney disease have been identified, including diabetes, hypertension, the presence of autoimmune diseases, systemic and urinary infections, as well as family history or exposure to certain environmental and/or pharmacological agents (Inker et al., 2014). Kidney disease is a risk factor for a number of complications that may happen independently from disease stage; these complications can be broadly classified into events derived from drug toxicity (e.g. faulty pharmacokinetics leading to drug toxicity and acute kidney injury [AKI]), metabolic and endocrine dysfunction (e.g. anemia, malnutrition and bone disorders as a result of reduced kidney function) and increased mortality risk due to **cardiovascular disease** (**CVD**) (Levin et al., 2013). Decreased GFR is associated with an increased prevalence of risk factors of CVD and the start of clinical abnormalities attributable to kidney failure (Sarnak et al., 2003). The strong link between kidney disease progression and incident CVD events is well-known ("K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification", 2002), as individuals with CKD are more likely to die due to CVD than reaching ESRD (Levey et al., 1998). Restless legs syndrome (RLS), a sleep-related movement disorder, is also associated with kidney disease and represents an important and frequent comorbidity in ESRD (Trenkwalder et al., 2016). Interestingly, a genetic contribution from RLS to ESRD has been reported (Schormair et al., 2011).

1.3.5 Kidney omics: current perspectives

Kidney research has seen many advances in the last two decades with regard to the profiling of -omic phenotypes at the population level, although a "systems epidemiology" approach to renal research questions is an idea initially proposed only around ten years ago (Haring & Wallaschofski, 2012). Although a review of the literature on -omic studies in kidney research is beyond the scope of this dissertation, the next paragraphs will introduce the reader to a few examples of relevant studies in the areas of genomics, epigenomics and proteomics of kidney function.

Genome-wide association studies (**GWAS**) in the field of renal research have identified genetic variants associated with numerous kidney traits, such as eGFR (Wuttke & Köttgen, 2016; Wuttke et al., 2019), eGFR decline (Gorski et al.,

2021), albuminuria (Teumer et al., 2019) and serum urate (Tin et al., 2019). In order to better understand the functional role of these genetic variants in kidney disease, studies thereafter have also focused on identifying genomic loci associated with gene expression and/or protein levels – such loci are called eQTL (expression quantitative trait loci) and pQTL (protein quantitative trait loci), respectively. For example, Martini and colleagues used transcriptomic and data from single nucleotide polymorphisms (SNPs) previously identified in an eGFR GWAS in kidney tissue from patients with CKD to identify eQTLs, offering insights on the opposite roles of inflammatory and metabolic processes in CKD (Martini et al., 2014).

Epigenome-wide association studies (**EWAS**), conducted in a similar manner to GWAS but with a focus on the association between methylation levels at CpG sites (or differentially methylated positions) rather than allele frequencies, have also been conducted to identify epigenetic signatures of eGFR (Breeze et al., 2021; Chu et al., 2017). Although a few studies have reported associations (or lack thereof) between a few kidney traits and some measures of epigenetic aging (Lu et al., 2019; Roshandel et al., 2020; J. A. Smith et al., 2019), these studies have been limited by the relatively small sample sizes and lack of replication samples.

In relation to **proteomic** studies of kidney function, earlier studies were focused on urinary proteins, which when combined proved to be a successful biomarker of disease progression (e.g. CKD273 classifier) (Good et al., 2010; Mischak et al., 2015). The blood proteome has been explored in more recent studies, mostly due to the existence of new technologies allowing for the profiling of complex biological samples – an example being SOMAscan, a platform using DNA aptamers to measure hundreds of plasma proteomic biomarkers (Gold et al., 2010). The reader may wish to refer to (Mischak et al., 2015) and (Sanchez-Nino et al., 2017) for comprehensive reviews on proteomics of kidney function. In brief, although this platform has been used in epidemiological studies of other health outcomes, prior assessments of the plasma proteome in relation to kidney function using aptamer-based technologies (or a similar approach) have been limited by the comparatively low number of proteins assessed (Carlsson et al., 2017) or by small sample sizes without replication (Christensson et al., 2017; Gold et al., 2010).

During peer-review of our manuscript, a similar study to ours was published (Ngo et al., 2020) – however, causality in the identified protein-eGFR associations was not addressed. **Mendelian randomization** (**MR**), an increasingly popular method in molecular epidemiology to infer causal effects between two traits, can be used to this end (Pierce & Burgess, 2013; Sekula et al., 2016).

1.4 Aims

The aim of this dissertation is to contribute to the understanding of the molecular basis of kidney function and disease by investigating the association between kidney traits and -omic biomarkers, more precisely with plasma proteins and DNA methylation (DNAm)-based predictors of aging and/or mortality. In collaboration with international research partners, two large-scale epidemiological research projects using data from multi-ethnic population-based studies were conducted with the objective of addressing two research questions:

1. Are plasma proteins associated with eGFR and CKD, and if so, is the relationship protein-eGFR causal?

   In order to identify proteins associated with eGFR and CKD, regression analyses adjusting for potential confounders were conducted in a German population-based cohort with 965 individuals in the discovery stage. In the replication stage, associations identified in the discovery stage were further tested in independent population-based studies of European and admixed ancestry with up to 1,887 individuals. Replicated eGFR-protein associations were further investigated using publicly available gene and protein expression datasets, pathway analyses and protein-protein interaction networks. Finally, replicated eGFR-protein associations were evaluated by conducting Mendelian randomization analyses, using genetic variants as instruments to infer causal effects of kidney function on plasma proteins and vice versa.

2. Are any measures of DNAm-based aging and mortality associated with the aforementioned and additional kidney traits?

   The association between DNAm-based measures of aging and mortality (or

epigenetic clocks) and multiple kidney traits was examined in up to 9,688 samples from five independent population-based cohorts with participants of European ancestry, as well as African American and Latino/Hispanic participants. Regression analyses adjusting for potential confounders were conducted in all studies, followed by transethnic and ethnic-specific meta-analyses in order to obtain combined and ethnic-specific estimates of the associations between DNAmAge acceleration and multiple kidney traits.

2 Methods

This chapter will address the study design, variable definitions and statistical approaches adopted in the two large-scale epidemiological research projects to address the aforementioned research questions, as outlined in (Lisa, 2014).

2.1 Plasma proteomics and kidney function

2.1.1 Study design

To investigate the association of plasma proteins and kidney function, a cross-sectional study was conducted. Data on the plasma levels of 993 proteins measured using an aptamer-based platform and kidney traits (i.e. eGFR and CKD) were available in up to 2,882 individuals (N = 2,548 with European ancestry, N = 334 admixed ancestry) from three population-based and one case-control studies; data was analyzed following a discovery-replication approach (Figure 2.1). Information on kidney traits (eGFR, CKD, uACR, albuminuria and serum urate) and established risk factors (i.e. age, sex, BMI, smoking, diabetes, hypertension, triglycerides, HDL and blood lipid lowering drugs) as well as blood samples, were collected at the time of interview by trained personnel following the standard operating procedures established by each cohort (Table 2.1). In brief, eGFR was estimated based on serum creatinine as per the CKD-EPI equation (Levey et al., 2009); chronic kidney disease (CKD) was defined as <60 ml/min/1.73 m^2 (Jha et al., 2013) and albuminuria was defined as >29 mg/g of creatinine (Toto, 2004).

2.1.2 Participating cohort studies

KORA (Cooperative health research in the Region of Augsburg) is a population-based cohort from Augsburg, southern Germany and two surrounding counties. The KORA F4 (2006-2008) study is a follow-up survey collecting detailed clinical

Table 2.1: Data availability across cohorts in proteomics study

Available variables	Studies included			
	KORA F4	QMDiab	INTERVAL	HUNT3
Aim #1: Association with plasma proteins	Discovery		Replication	
Aptamer-based protein levels	✓	✓	—	—
Aim #2: Association with DNAmAge acceleration	—	—	—	—
DNAm	✓	—	—	—
Kidney traits				
eGFR	✓	✓	✓	✓
CKD	✓	✓	✓	✓
uACR	✓	—	—	—
Microalbuminuria	✓	—	—	—
Serum urate	✓	—	—	—
Covariates				
Age	✓	✓	✓	✓
Sex	✓	✓	✓	✓
BMI	✓	✓	✓	✓
Smoking	✓	✓	✓	✓
Diabetes	✓	✓	✓	✓
Hypertension	✓	✓	✓	✓
Triglycerides	✓	✓	✓	✓
HDL	✓	✓	✓	✓
Lipid-lowering drugs	✓	—	✓	—

Available variables at assessment time-points in each study are marked with a ✓, "—" denotes not available.

and demographic information as well as peripheral blood for -omics analyses from participants of the KORA S4 survey (1999–2001) (Holle et al., 2005; Rathmann et al., 2009). The KORA F4 served as a discovery cohort, and three additional studies were included to test for replication of our findings: Nord-Trøndelag Health Study (**HUNT**), more precisely the third survey (HUNT3) from this population-based study from Norway with data on participants of European descent (Krokstad et al., 2012); the INTERVAL Study (**INTERVAL**), a randomized trial on blood donation intervals with participants of European descent from the UK (Moore et al., 2016); and the Qatar Metabolomics Study on Diabetes (**QMDIAB**), a cross-sectional case-control study on type 2 diabetes from participants of Arab, South Asian and Filipino descent in Qatar (Mook-Kanamori et al., 2014).

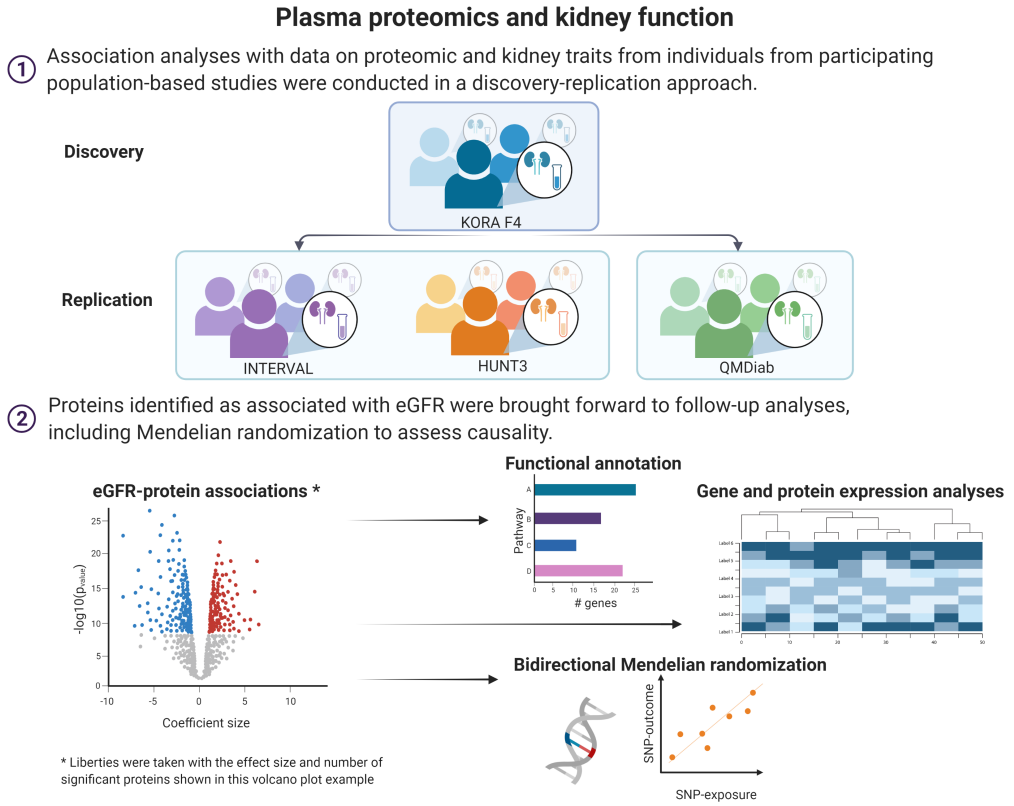

Figure 2.1: Study design of project on plasma proteomics and kidney function, based on (Matías-García et al., 2021b)

2.1.3 Assessment of proteomics

Plasma proteins were measured using SOMAScan, a multiplex aptamer-based platform allowing for high-throughput measurement with high sensitivity and specificity of secreted, extracellular and intracellular proteins in blood plasma (Rohloff et al., 2014). Additional details on the technical details of these measurements are offered in (Nayor et al., 2020; Suhre & Arnold, 2017; Sun et al., 2018) and in the first-author publication (Matías-García et al., 2021b).

2.1.4 Variable definitions

Table 2.2 displays how variables were used in the proteomics and kidney function analyses. Column 1 displays the name, column 2 provides a definition of the variable, whereas column 3 shows how each variable was coded in the analyses.

Table 2.2: Variable categorization for aim #1

Name	Description and units	Type
Outcome variables		
eGFR	Estimated glomerular filtration rate as calculated by the serum creatinine-based CKD-EPI equation, log-transformed (ml/min/1.73 m^2)	continuous
CKD	Chronic kidney disease, defined as <60 ml/min/1.73 m^2 (yes, no)	dichotomous
Exposure variables		
Protein	Protein plasma levels obtained using an aptamer-based platform, measured in relative fluorescence units (RFU)	continuous
Covariates		
Age	Age at time of examination (years)	continuous
Sex	Self-reported biological sex (male, female)	dichotomous
BMI	Body mass index (kg/m^2)	continuous
Smoking status	Self-reported smoking (current, former or never-smoker)	dichotomous
Diabetes	Fasting plasma glucose $\geq$ 126 mg/dl or treatment for diabetes (yes, no)	dichotomous

Hypertension	Systolic blood pressure $\geq$ 140 mm Hg or diastolic blood pressure $\geq$ 90 mm Hg or treatment for hypertension (yes, no)	dichotomous
Triglycerides	Log-transformed (mg/dL)	continuous
HDL	High density lipoprotein (mg/dL)	continuous
Intake of lipid lowering drugs	Self-reported intake of drugs with ATC code C10 (yes, no)	dichotomous

2.1.5 Statistical analysis

Logistic and linear multiple regression models were used to assess the relationship between kidney traits as outcomes and protein levels as predictors. This was first conducted in the discovery study, KORA F4, adjusting for the covariates listed in Subsection 2.1.4:

$$\text{kidney trait} \sim \text{protein level} + \text{covariates}$$

Replication using the same statistical models was conducted in three independent studies, though only associations with eGFR and CKD could be explored as the other traits were not available in these cohorts (Table 2.1). Pathway analyses, protein-protein interaction networks and causal inference were conducted based on the set of replicated eGFR-protein associations. All statistical and bioinformatics analyses were done with R v.3.6.0 (R Core Team, 2019). Details on the statistical principles from the conducted analyses are provided in Section 2.3 of this Chapter.

2.2 DNAmAge acceleration and kidney traits

2.2.1 Study design

To determine whether DNAmAge acceleration (DNAmAA) is associated with parameters of low renal function, a cross-sectional study was conducted. The association between DNAm-based aging and mortality predictors and kidney traits was examined in up to 9,688 individuals (N = 6,363 with European ancestry, N = 2,718 African American and N = 607 Hispanic/Latino individuals) from five population-based studies, and the estimates from each study pooled in a large meta-analysis

(Figure 2.2). Information on kidney traits (eGFR, CKD, uACR, albuminuria and serum urate) and established risk factors (i.e. age, sex, BMI, smoking, diabetes, hypertension, triglycerides and HDL) as well as blood samples, were collected at the time of interview by trained personnel following the standard operating procedures established by each cohort (Table 2.3). In brief, eGFR was estimated based on serum creatinine as per the CKD-EPI equation (Levey et al., 2009); chronic kidney disease (CKD) was defined as <60 ml/min/1.73 m^2 (Jha et al., 2013) and albuminuria was defined as >29 mg/g of creatinine (Toto, 2004).

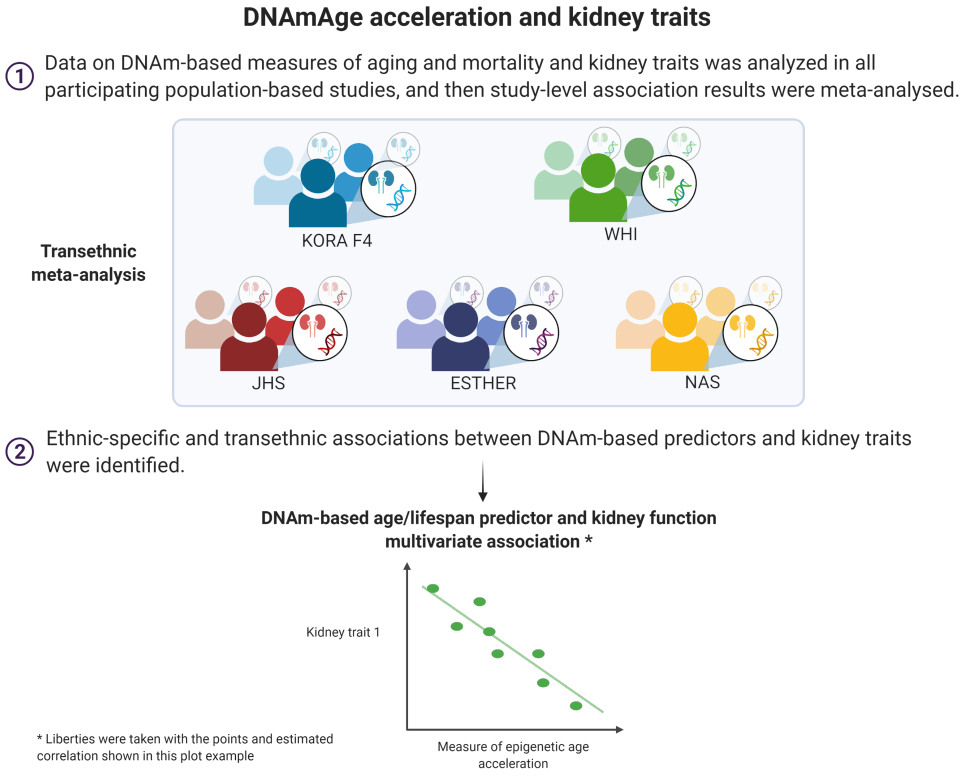

Figure 2.2: Study design of project on DNAmAge acceleration and kidney traits, based on (Matías-García et al., 2021a)

Table 2.3: Data availability across cohorts in DNAmAge study

Available variables	Studies included				
	KORA F4	ESTHER	NAS	WHI	JHS
Aim #1: Association with plasma proteins	—	—	—	—	—
Aptamer-based protein levels	✓	—	—	—	—
Aim #2: Association with DNAmAge acceleration		Meta-analysis			
DNAm	✓	✓	✓	✓	✓
Kidney traits					
eGFR	✓	✓	✓	✓	✓
CKD	✓	✓	✓	✓	✓
uACR	✓	✓	—	—	✓
Microalbuminuria	✓	✓	—	—	✓
Serum urate	✓	✓	✓	—	✓
Covariates					
Age	✓	✓	✓	✓	✓
Sex	✓	✓	✓	✓	✓
BMI	✓	✓	✓	✓	✓
Smoking	✓	✓	✓	✓	✓
Diabetes	✓	✓	✓	✓	✓
Hypertension	✓	✓	✓	✓	✓
Triglycerides	✓	✓	✓	✓	✓
HDL	✓	✓	✓	✓	✓
Lipid-lowering drugs	✓	—	—	—	—

Available variables at assessment time-points in each study are marked with a ✓, "—" denotes not available.

2.2.2 Participating cohort studies

The **KORA F4** (2006-2008) study is a follow-up survey collecting detailed clinical and demographic information as well as peripheral blood for -omics analyses from participants of the KORA S4 survey (1999–2001) (Holle et al., 2005; Rathmann et al., 2009).

The **ESTHER study** (Epidemiologische Studie zu Chancen der Verhütung, Früherkennung und optimierten THerapie chronischer ERkrankungen in der älteren Bevölkerung) is a population-based cohort study conducted in the federal state of Saarland, Germany (Raum et al., 2007); individuals' information on sociodemographic characteristics, lifestyle factors, and history of major diseases, as well as blood samples for -omics profiling, were obtained at the baseline interview (2000-2002) (Zhang et al., 2017).

The **NAS** (Normative Aging Study) is a longitudinal study established by the U.S. Department of Veterans Affairs in 1963 based in the Greater Boston area (Mehta et al., 2016); information on lifestyles, dietary habits, activity levels, and demographic factors, as well as blood samples, are collected at each visit (up to four visits between 1999 and 2013) (X. Gao, Colicino, et al., 2019; Mehta et al., 2016).

The **WHI** (Women's Health Initiative) is a study of postmenopausal women recruited from 40 U.S. clinical centers who participated in an observational study or in clinical trials during 1993-1998 (Anderson et al., 2003; Howard et al., 2006; Jackson et al., 2006; The Women's Health Initiative Study Group, 1998).

The **JHS** (Jackson Heart Study) is a prospective, community-based cohort designed to investigate risk factors for cardiovascular disease among African Americans in the Jackson, Mississippi, metropolitan tri-county area (Hinds, Madison, and Rankin); information on clinical variables, lifestyle and sociocultural factors were obtained at the baseline JHS examination (2000-2004) and in two subsequent clinic visits (2005-2008 and 2009-2013) (Carpenter et al., 2004; Taylor et al., 2005; Wilson et al., 2005).

2.2.3 Assessment of DNAm and epigenetic age

Methylation levels of approximately 480,000 or 850,000 CpG sites were measured using the 450k or EPIC methylation arrays, respectively (Bibikova et al., 2011). Methylation data was normalized and processed by an analyst in each cohort applying the pre-processing pipeline of their preference and following a standardized workflow. DNAm data was used to estimate DNAm-based predictors of aging and mortality, and whenever appropriate DNAmAA was calculated as the difference between an individual's DNAmAge and chronological age (Fransquet et al., 2019). Additional technical details of these measurements are offered in (X. Gao, Colicino, et al., 2019; Zeilinger et al., 2013) and in the first-author publication (Matías-García et al., 2021a).

2.2.4 Variable definitions

Table 2.4 displays how variables were used in the DNAmAge and kidney function analyses. Column 1 displays the name, column 2 provides a definition of the variable, whereas column 3 shows how each variable was coded in the analyses.

Table 2.4: Variable categorization for aim #2

Name	Description and units	Type
Outcome variables		
eGFR	Estimated glomerular filtration rate as calculated by the serum creatinine-based CKD-EPI equation, log-transformed (ml/min/1.73 m^2)	continuous
CKD	Chronic kidney disease, defined as <60 ml/min/1.73 m^2 (yes/no)	dichotomous
uACR	Urinary albumin-to-creatinine ratio (mg/g)	continuous
Microalbuminuria	uACR $\geq$ 30 mg/g (yes/no)	dichotomous

Serum urate	Also known as uric acid (mg/dl)	continuous
Exposure variables		
Hannum DNAmAA	DNAmAge Acceleration measured by the Hannum clock, defined as the difference between Hannum's DNAmAge and chronological age	continuous
Horvath DNAmAA	DNAmAge Acceleration measured by the Horvath clock	continuous
EEAA	Extrinsic Epigenetic Age Acceleration	continuous
IEAA	Intrinsic Epigenetic Age Acceleration	continuous
PhenoAA	PhenoAge acceleration	continuous
GrimAA	GrimAge acceleration	continuous
MRS	Mortality risk score, calculated as the linear combination of 10 CpGs	continuous
MRS, categorical	Three risk levels based on the total number of "aberrantly" methylated CpGs: low risk if 0–1 CpGs; moderate risk if 2–5 CpGs; high risk if >5 CpGs	categorical
Covariates		
Age	Age at time of examination (years)	continuous
Sex	Self-reported biological sex (male, female)	dichotomous
BMI	Body mass index (kg/m^2)	continuous
Smoking status	Self-reported smoking (current, former or never-smoker)	dichotomous
Diabetes	Fasting plasma glucose $\geq$ 126 mg/dl or treatment for diabetes (yes, no)	dichotomous
Hypertension	Systolic blood pressure $\geq$ 140 mm Hg or diastolic blood pressure $\geq$ 90 mm Hg or treatment for hypertension (yes/no)	dichotomous
Triglycerides	Log-transformed levels (mg/dL)	continuous
HDL	High density lipoprotein (mg/dL)	continuous

Houseman variables	Blood cell proportions estimated using DNAm, as described in (Houseman et al., 2012)	continuous

2.2.5 Statistical analysis

Logistic and linear regression models were used to assess the relationship between five kidney traits (eGFR, CKD, UACR, microalbuminuria and serum urate) as outcomes and seven DNAm-based measures of aging and/or lifespan as predictors, adjusting for the covariates defined in Subsection 2.2.4:

$$\text{kidney trait} \sim \text{DNAm-based aging/lifespan measure} + \text{covariates}$$

These analyses were conducted in all participating studies by a designated analyst. Study-levels results were meta-analysed using fixed- and random-effects models using the metafor package v2.0 (Viechtbauer, 2010) in R version 3.5.3 (R Core Team, 2019). Additional details on the statistical principles are provided in the next section of this Chapter.

2.3 Statistical approaches

In order to evaluate the association between eGFR and/or other kidney traits (as outcomes) and plasma protein levels (exposure in aim #1) or DNAm-based measures of aging (exposure in aim #2), several statistical analyses were conducted (Table 2.5).

The following sections will provide an overview of the statistical approaches used in this dissertation.

Table 2.5: Data analysis strategies

Data analysis plan	
Aim #1: Association with proteomics	
Univariate analysis	Mean and SD of continuous variables, number and percent of categorical variables
Bivariate analysis	Assessment of unadjusted relationship between plasma proteins and eGFR using correlation in discovery cohort
Multivariate analysis	Linear and logistic regression models to model the association between plasma proteins and eGFR, adjusting for known confounders, calculating odds ratios and/or beta coefficients, as well as their corresponding standard errors
Causal inference	Two-sample Mendelian randomization using inverse-variance weighted (IVW) MR, MR-Egger, weighted median MR, mode-based MR, and Wald's ratio
Aim #2: Association with DNAmAge acceleration	
Univariate analysis	Mean and SD of continuous variables, number and percent of categorical variables
Multivariate analysis	Linear and logistic regression models to model the association between DNAmAA and kidney traits, adjusting for known confounders, calculating odds ratios and/or beta coefficients, as well as their corresponding standard errors
Meta-analysis	Trans-ethnic and ethnic-specific meta-analyses of study-level results using fixed-effect and random-effects models

2.3.1 Association analyses

Foulkes divides the data produced from population-based genetic association studies into three components (Foulkes, 2009), although this conceptualization can be extended to most -omic analyses. The first component is the *genotype* of the individual, which when may be exchanged with the term *molecular phenotype* when extending this terminology to other types of -omics; the second component is the *phenotype*, or trait(s) of interest known to be correlated with the health outcome under scrutiny; and the third component, *covariates*, are additional variables capturing biological, environmental and socioeconomic information relevant to the analyses at hand (Foulkes, 2009). In this sense, the goal of association studies in molecu-

lar epidemiology is first to determine whether a relationship between a molecular phenotype and a trait exists, and second, to describe this relationship. The next subsections will describe the statistical tests applied to explore such associations, strategies used to take into account the issues derived from multiple testing, and introduce the reader to a causal inference method using genetic variation to assess causality in trait-trait associations.

Regression modeling

The linear relationship between two variables x and y can be modeled as a straight line, under the assumption that the outcome variable y_i is linearly related to the predictor variables x_i with a normally distributed random component. This is modeled in the following equation:

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \text{ for } i = 1, \ldots, n \tag{2.1}$$

with

$$\epsilon_i \sim N(0, \sigma^2) \tag{2.2}$$

Here, ϵ_i is a term representing the residual error, which follows the normal distribution with a constant mean of 0 and variance σ^2. The unknown regression coefficients (β_0 and β_1) and the population variance σ^2 are to be estimated from the data at hand. One of the most common methods to estimate the regression coefficients is the method of least squares, which produces a fitted line that satisfies the condition that the sum of the square of the differences from the fitted line to the observed points should be minimized. Under the normality assumption of the residuals (Equation 2.2), the estimates produced by least squares are equal to the estimates produced by maximum likelihood methods (Heiberger & Holland, 2015). Multiple linear regression extends these principles to cases in which two or more predictors are included (Heiberger & Holland, 2015). The calculated β coefficients can be interpreted as the estimated mean difference in the outcome per unit change in the exposure given all other variables included remain constant – for example, the mean difference in eGFR per increase in protein levels after adjusting for potential confounders.

Of note, as in $\mathrm{E}(Y \mid X = x)$, Y is the "regressand" (also known as dependent variable or outcome variable), and X the "regressor" (also termed independent variable, predictor or covariate). Despite the confusion that the traditional "in/dependent variable" terminology may cause, the regression of Y on X, $\mathrm{E}(Y \mid X = x)$, or X on Y, $\mathrm{E}(X \mid Y = y)$, can be considered for any pair of variables and does not necessarily imply a causal direction of effect or any temporal relation between variables defined as regressor and regressand (Greenland, 2005).

Logistic regression is useful to assess data where a dependent variable is coded as a dichotomous variable (i.e. it can only adopt two possible values, like sick/healthy and treatment/control) or it represents a sample proportion. Logistic regression is in principle applied following the same principles from linear regression, although it introduces the use of a link function to allow for the analysis of a dichotomous dependent variable. The logarithm of the odds, also known as the logit transformation, allows for this transformation from having a closed interval [0,1] to the set of all real numbers:

$$y = logit(p) = \ln(\frac{p}{1-p}) \tag{2.3}$$

So that the model for logistic regression with one predictor is expressed as:

$$logit(p) = \beta_0 + \beta_1 x + \epsilon \tag{2.4}$$

where p is the outcome variable, either coded as a binary variable or a proportion, x is the predictor variable, and ϵ is the residual that is assumed to follow a binomial distribution. Additional details on the estimation of the coefficients by maximum likelihood methods can be found in (Heiberger & Holland, 2015).

Meta-analysis

The following section is based on the guide by (Harrer et al., 2021). Meta-analysis involves analyzing multiple studies and applying statistical methods to pool and explore heterogeneity in results from the different studies. There are two main meta-analysis models, namely fixed-effect and random-effects models.

The **fixed-effect model** works under the assumption that all obtained estimates are derived from one homogeneous population, and therefore builds on the existence of one true estimate shared by all studies (θ):

$$\hat{\theta}_k = \theta + \epsilon_k \tag{2.5}$$

where $\hat{\theta}_k$ is the observed estimate from the study k, obtained as the true effect of the population θ modified by the sampling error in the study ϵ_k. The sampling error is also represented by the standard error of the estimate. Under this model, studies with smaller standard errors (therefore offering more precise estimates) are assigned a greater weight prior to pooling the estimates. In the **inverse-variance meta-analysis**, the weight of each study (ω_k) is then given by the inverse of the variance of each estimate ($\frac{1}{s_k^2}$).

Irrespective of the choice of weight for the studies, the combined estimate is then given by the weighted average of the estimates across a set of k studies:

$$\hat{\theta} = \frac{\sum_{k=1}^{K} \omega_k \hat{\theta}_k}{\sum_{k=1}^{K} \omega_k} \tag{2.6}$$

However, in a perhaps more realistic scenario, between-study heterogeneity may affect the obtained effects and should be taken into account in the meta-analysis models. The **random-effects model** assumes that the estimated effect ($\hat{\theta}_k$) results from a study-specific true effect size (θ_k) considering the sampling error (ϵ_k), as follows:

$$\hat{\theta}_k = \theta_k + \epsilon_k \tag{2.7}$$

The inclusion of the term from the study's true effect size (θ_k) shows that there is no assumption of a true universal effect, but a study-specific effect θ_k derived from the mean of a distribution of true effect sizes (μ) and its own error estimation (ζ_k):

$$\theta_k = \mu + \zeta_k \tag{2.8}$$

The observed effect is then obtained under the random-effects model as follows:

$$\hat{\theta}_k = \mu + \zeta_k + \epsilon_k \tag{2.9}$$

To take the error (ζ_k) (and thus between-study heterogeneity) into account, the variance of the distribution of the true effect sizes (τ^2) is calculated and included in the given random-effects weights to each effect (ω_k^*):

$$\omega_k^* = \frac{1}{s_k^2 + \tau^2} \tag{2.10}$$

Of note, there are many different methods to estimate τ^2, of which the most often applied is that of DerSimonian and Laird; another option is its estimation by Restricted Maximum Likelihood (REML).

The random-effects pooled size is calculated using the random-effects weights ω_k^* instead of ω_k in Equation 2.6, as shown next:

$$\hat{\theta} = \frac{\sum_{k=1}^{K} \omega_k^* \hat{\theta}_k}{\sum_{k=1}^{K} \omega_k^*} \tag{2.11}$$

An alternative meta-analysis model that can be used to pool p-values in circumstances where raw data and the effect estimates derived thereof (e.g. regression coefficient) cannot be combined is the "inverse normal" or **weighted Z-test method**. This test, a version of Stouffler's inverse normal method, transforms and combine the p-values from the k study (p_k) by using the inverse normal transformation and then weights using the square root of the sample size as a study-specific weights (ω_k). The combined p-value is obtained based on the distribution of the statistic from sum of the weighted p-values across studies (Zaykin, 2011).

Multiple testing correction

In regression analyses, hypothesis testing can be done on single regression coefficients, where the null hypothesis of no effect of the predictor on the outcome (H_0: $\beta_i = 0$) is tested against the alternative hypothesis of H_1: $\beta_i \neq 0$. A test

statistic is calculated and its p-value, or the probability of observing an estimate under the assumption of H_0 being true that is at least as extreme as the estimated from the sample, is used to either reject H_0 (if $\alpha >$ p value) or retain H_0 (if otherwise) (Heiberger & Holland, 2015).

Type I errors are produced when the null hypothesis H_0 of is rejected when H_0 is true. Type II errors are when H_0 is not rejected when the alternative hypothesis H_1 is true (Table 2.6). The probability of making a type I error, denoted as α, is pre-specified to a certain accepted level (most commonly 0.05). β, on the other side, is the probability of making a Type II error; the probability of correctly rejecting a false H_0, or the power of a test, is therefore given by $1 - \beta$ (Heiberger & Holland, 2015).

Table 2.6: Hypothesis testing

		Decision based on test	
		Retain H_0	Reject H_0
Reality	H_0 is true	True negative	Type I error
	H_1 is true	Type II error	True positive

Based on (Heiberger & Holland, 2015) and (Foulkes, 2009)

Molecular epidemiology studies also have the particular feature of including hundreds or thousands of molecular phenotypes for their analysis (e.g. genes, proteins, CpG sites or gene transcripts). If multiple testing in the statistical analyses is not accounted for, spurious positive associations arising from Type I and Type II errors may be reported (Khoury et al., 2008). Testing multiple hypotheses results in the inflation of the error rate, which can be controlled through the adjustment for multiple comparisons in two ways: controlling the family-wise error rate (FWER: probability of making at least one type-I error) and the false discovery rate (FDR: proportion of true H_0 from those declared significant). A straightforward (but conservative) approach to control the FWER for multiple testing is apply the Bonferroni adjustment, a single step adjustment method done by dividing the overall level $\alpha = 0.05$ by the number of independent tests conducted, and setting this as

the level α' for each test (Foulkes, 2009).

2.3.2 Mendelian Randomization

Mendelian Randomization (MR) is a causal inference method based on instrumental variable analyses. Since gene variants (more specifically polymorphisms at the single nucleotide level, SNPs) are naturally randomly assorted during gamete formation (and this prior to any exposures/outcomes) and do not change with time, they can be used as "fixed" variables to assess exposures to lifelong traits on health outcomes (Davey Smith & Hemani, 2014). The nature of these variables, also called instruments, minimizes the risk of reverse causation and confounding common to effect estimates from observational epidemiology (Davey Smith & Hemani, 2014).

In a nutshell, if genetic variation has an impact on health outcomes similar to that of environmental exposures, it can be said that genetic variation itself is related to disease risk by modifying the exposure (Davey Smith & Hemani, 2014; G. D. Smith & Ebrahim, 2003). For MR to produce robust causal estimates of "genetically" determined exposures, SNPs used must meet its three underlying assumptions (Figure 2.3, panel A): genetic variants (Z_A) must be strongly associated with the exposure (marked as (1)), and not associated with the outcome but through the exposure (2) or with potential confounders (3) (Davey Smith & Hemani, 2014; Lawlor et al., 2008).

Two-sample Mendelian randomization (2SMR) can be used to explore causality when summary statistics from gene-exposure and gene-outcome associations are estimated in different population samples (Z_A and Z_B in Figure 2.3, panel A) (Davey Smith & Hemani, 2014; Haycock et al., 2016; Pierce & Burgess, 2013). This presents an ideal tool in an era when GWAS on multiple different traits, like protein levels (Emilsson et al., 2018; Sun et al., 2018) and kidney traits (Wuttke et al., 2019), have been made publicly available. These datasets can be used by researchers to explore the causal nature of observational associations in the field of kidney research (Sekula et al., 2016). MR can also be conducted in a bidirectional manner, meaning causality in both directions (i.e. from exposure to outcome and vice versa) is explored (panel B in Figure 2.3). Comprehensive details on the

different methods, sensitivity analyses and interpretation of MR results has been provided in the main text and supplemental notes from the first-author publication (Matías-García et al., 2021b).

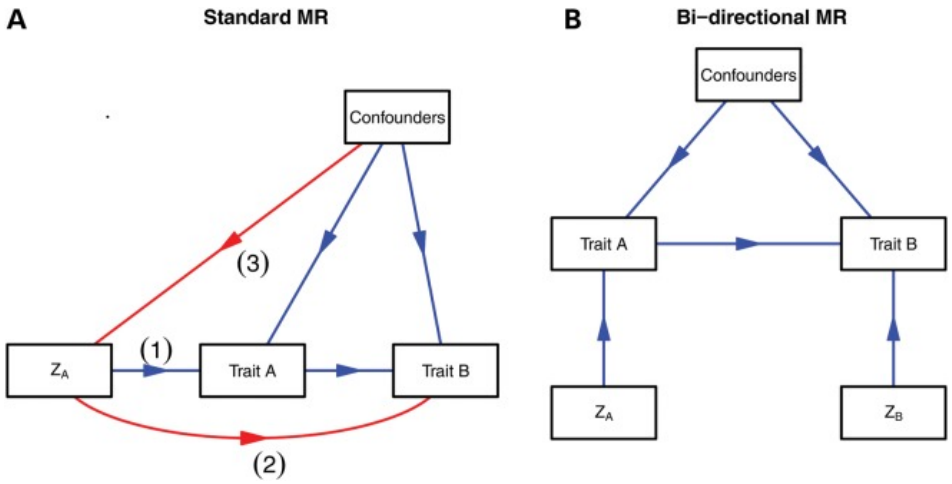

Figure 2.3: Representation of causality inference and assumptions in Mendelian Randomization, reproduced from (Davey Smith & Hemani, 2014)

3 First-author Publications

3.1 Plasma proteomics and kidney function

3.1.1 Publication

The original research paper was published in the Journal of the American Society of Nephrology (Matías-García et al., 2021b), and is included in this dissertation in Appendix A.

3.1.2 Aim of this study

A cross-sectional study using data from population-based and one case-control studies was conducted to identify potential novel plasma proteomic biomarkers of kidney function, specifically of estimated glomerular filtration rate (eGFR) and chronic kidney disease (CKD, defined as reduced eGFR < 60 ml/min/1.73 m^2).

3.1.3 Summary of results

In this epidemiological study of aptamer-based plasma proteins and kidney disease, we were able to identify 80 plasma proteins associated with eGFR (of which 34 were also associated with CKD) after adjusting for potential confounders in the discovery study. Collaboration with other population samples having available data on aptamer-based proteomics allowed us to replicate the identified associations by running the same regression models in independent studies, thus serving as an external validation step. Replication was conducted in three stages: protein-eGFR and -CKD associations were replicated first in European ancestry studies, then in an admixed-ancestry study, and the set of overlapping associations between both aforementioned replication sets was considered the final set of replicated trans-ethnic associations: 57 proteins were consistently associated with eGFR, of which 23 were additionally correlated with CKD. Cystatin C, b2-microglobulin, IGBFP6

and testican-2 were among the top proteins identified. Many of the proteins identified in this association study have been reported in earlier studies using the same aptamer-based proteomics platform, thus our study provided confirmatory evidence of their association. Moreover, one novel protein, contactin-4, was identified in our study. Additionally, sensitivity analyses using eGFR calculated based on cystatin-C in the discovery sample showed consistent results for all 57 replicated proteins.

The set of 57 replicated proteins associated with eGFR were further investigated in enrichment analyses. Gene/protein expression across all tissues was also investigated in publicly available data sources (GTEx and ProteomeDB). No enrichment of pathways, biological processes or molecular functions were identified, likely due to the targeted nature of the aptamer-based platform. Moreover, the expression of most of the studied proteins were detected across many tissues. Protein-protein interaction network analysis (PPI) showed that the set of 57 proteins shared more interactions within the network built than expected, suggesting potential shared functions and pathways.

Causal effects using publicly available GWAS data on gene-eGFR and gene-protein associations were estimated by two-sample bidirectional Mendelian randomization. Robust evidence of a positive causal effect across multiple MR methods was observed for the effect of kidney function (with eGFR as a proxy thereof) on testican-2. This suggests interventions (either pharmacological or lifestyle) designed to improve eGFR are genetically predicted to also have an increasing effect in the plasma levels of testican-2. Evidence of causal effects of three proteins (melanoma inhibitory activity [MIA], carbonic anhydrase III and cystatin M) on eGFR was also observed, although these estimates were based on fewer SNPs and could not be thus tested in sensitivity analyses.

Gene expression analyses in kidney tissue conducted with publicly available datasets from *Nephroseq* showed a positive bivariate correlation between eGFR and the expression of the protein-coding gene for testican-2, *SPOCK2*. Multivariable analyses conducted with RNA-sequencing transcriptome data from human kidney showed no significant association with eGFR, but a negative association with tubular atrophy and interstitial fibrosis.

Our study provides insights into novel potential proteomic biomarkers of eGFR and CKD using a novel multiplex technology in at the population sample level. We were able to identify and replicate a number of plasma proteins associated with kidney function as measured by clinically relevant parameters, where these associations seem to extend across the ethnicities included in our study. Results from the association study, causal inference analyses and gene expression in kidney tissue further suggest testican-2 and its protein-coding gene *SPOCK2* may have an important role in kidney function. Low plasma levels of testican-2 may be indicative of poor renal function, suggesting this protein may be a physiological marker of kidney disease progression with potential clinical relevance.

3.1.4 Contributions

I contributed to this project and its resulting publication by planning the study design, creating an analysis plan and establishing a multidisciplinary collaboration with researchers from independent studies. I was also responsible for conducting the association and sensitivity analyses in the discovery population-based study, as well as the causal inference analyses with GWAS data. All the main figures, tables and references were created and/or curated by me. Regarding the paper writing, I wrote the first full draft of the manuscript and integrated feedback from co-authors. I also was the main responsible for addressing the referees' comments during the two rounds of peer-review and for the communication with the editor.

3.2 DNAmAge acceleration and kidney traits

3.2.1 Publication

The original research paper was published in the Journal of Clinical Epigenetics (Matías-García et al., 2021a), and is included in this dissertation in Appendix B.

3.2.2 Aim of this study

A cross-sectional study with five multi-ethnic, population-based studies was conducted to investigate whether established epigenetic age acceleration, as measured

by multiple DNAm-based predictors of aging and mortality, is associated with kidney function and related traits (i.e. eGFR, CKD, uACR, albuminuria, and serum urate).

3.2.3 Summary of results

In this epidemiological study of DNAm-based predictors of aging and mortality and kidney disease, we were able to identify associations between epigenetic age acceleration and poor kidney function. Collaboration with other population samples having available data on DNAm and kidney traits allowed us to adjust for the same set of confounders and run the same regression models in all studies.

The study-level results were then pooled in a large-scale, trans-ethnic meta-analysis of up to 9,688 individuals. We identified 23 significant DNAm-based predictor-kidney trait associations ($p < 1.43E-03$ and concordant direction of effect across studies) in the transethnic fixed-effect meta-analysis. These associations can be divided into three interesting groups. In the first, all included parameters of poor kidney function (i.e. lower eGFR, prevalent CKD, higher uACR or microalbuminuria and higher serum urate) were associated with a positive epigenetic age acceleration as measured by PhenoAge, extrinsic epigenetic age acceleration (EEAA) and mortality risk score (MRS). The second group, composed by the "first-generation" epigenetic clocks HannumAge and HorvathAge, showed a different pattern of association: age acceleration in the Horvath estimate (HorvathAA) was only negatively associated with eGFR (i.e. epigenetic age acceleration correlated with lowering levels of eGFR). HannumAA, on the other hand, was also negatively associated with all traits but serum urate. Finally, the third group of associations featured those identified with the mortality predictor GrimAge, where a measure of age acceleration calculated with this DNAm-based predictor was associated with uACR, albuminuria and serum urate but not with serum creatinine-based traits. Of note, low heterogeneity between-studies was identified in most associations between DNAm-based predictors and kidney traits. In cases where heterogeneity was high, evidence from a random-effects model was interpreted and included in the description of the main results.

Six associations between DNAm-based predictors and kidney traits were replicated across studies with participants of European ancestry and African Americans. The 10 CpG MRS stands out for its replication across all subgroups (including the small Hispanic/Latino cohort) in relation to its association with low eGFR, CKD and high levels of serum urate. Likewise, the associations between low eGFR and extrinsic measure of aging (EEAA), as well as those between PhenoAA and CKD and urate, were also independently observed in studies with African Americans and individuals with European ancestry. Ethnic-specific replication was observed for 16 associations in total, where studies of European ancestry mostly were driving the associations identified for uACR and eGFR with PhenoAA, EEAA and HorvathAA.

Secondary analyses to investigate associations with the eight DNAm-based components of GrimAge and the categorical risk variables of the MRS (high risk and moderate risk) were also conducted. These analyses showed that DNAm-estimated adrenomedullin (DNAmADM), plasminogen activator inhibitor-1 (DNAmPAI1) and pack years (DNAmPACKYRS) were positively associated with higher uACR, higher serum urate levels and microalbuminuria. Likewise, an increase in uACR was identified for individuals with > 5 "aberrantly" methylated CpGs compared to those with 0-1 aberrantly methylated loci (high risk MRS vs low risk MRS). Similar associations were observed for microalbuminuria.

Our study, the first study of this nature on this topic, offers evidence on the correlation between multiple kidney traits and DNAm-based aging and lifespan predictors measured in whole blood, as well as with some secondary DNAm-estimated markers. The observed associations seem to reflect the impact of mechanisms such as immunosenescence, inflammaging and oxidative stress on kidney disease. Our study also contributes to a better understanding of the DNAm-based predictors themselves, as it seems like the changes captured by the CpGs included in these predictors also mirror pathological changes common to systemic inflammation and renal disease. Nevertheless, future research in clinical studies is required to assess whether these DNAm-based predictors may be useful in disease prognosis and stratification. Functional studies to pinpoint the molecular mechanisms underlying the physiological interplay between epigenetic mechanisms and biological aging are also warranted.

3.2.4 Contributions

I contributed to this project and its resulting publication by planning the study design, creating an analysis plan and establishing a collaboration with scientists from independent studies. I was also responsible for conducting the association analyses in the KORA F4 study, as well as for combining the data from all studies in the meta-analyses. All the main figures, tables and references were created and/or curated by me. Regarding the paper writing, I wrote the first full draft of the manuscript and integrated feed-back from co-authors. I also was responsible for addressing the referees' comments during peer-review and for the communication with the managing editor.

4 Discussion

The aim of this thesis is to contribute to the understanding of the molecular mechanisms underlying kidney disease by identifying novel potential plasma proteomic biomarkers and epigenetic measures of aging associated with kidney traits. This aim was achieved by conducting two large-scale epidemiological studies with information from multi-ethnic population-based cohorts, assessing our findings with regard to the available body of scientific and medical literature and by making our results publicly available as open-access original research articles (Matías-García et al., 2021a; Matías-García et al., 2021b). This chapter will guide the reader through a brief recapitulation of relevant literature to interpret the main findings, strengths and limitations from both studies, and present future opportunities to further explore these topics.

4.1 Plasma proteomics and kidney function

4.1.1 About our findings in relation to current literature

We identified 57 proteins associated with eGFR (23 of them were also associated with CKD) in one of the first two transethnic meta-analyses of renal function proteomics (Matías-García et al., 2021b). Several well-known biomarkers of renal function were included in our findings (e.g. cystatin C, b2-microglobulin, IGBFP6) (Carlsson et al., 2017; Christensson et al., 2017; Gold et al., 2010; Jovanovic et al., 2003; Niewczas et al., 2019), thus supporting the validity of our aptamer-based proteomic analyses.

Our results are also in line with those reported by other renal studies using the same aptamer-based platform or similar proteomic profiling technologies (Carlsson et al., 2017; Christensson et al., 2017; Gold et al., 2010; Ngo et al., 2020): we replicate 15 of the proteins identified in the pioneer SOMAScan study of plasma

samples from 42 CKD patients (Gold et al., 2010), five of the proteins associated with lower baseline eGFR and 5-year eGFR decline in a study examining 80 circulating proteins in 1,000 participants (Carlsson et al., 2017), and a large number of proteins reported in a recent SOMAScan study of 2,893 plasma proteins in 389 Swedish individuals (Christensson et al., 2017). Moreover, five of our proteins (TNF SR-I and –II, TAJ, RELT, DAF and CCL14) were included in a signature capturing the inflammatory process underlying end-stage renal disease in diabetic cohorts (Niewczas et al., 2019), and another five proteins (b2-microglobulin, cystatin C, DAF, MP2K2 and testican-2) were included in a set of proteins meant to reflect renal health in a "stand-alone" blood test (Williams et al., 2019). Interestingly, 40% of our proteins (k = 23) were identified in podocyte exosome-enriched urine, suggesting their involvement in cellular functional processes underlying glomerular filter permeability (Prunotto et al., 2013). At the time of revision of this paper, a study with a similar design and aims was published (Ngo et al., 2020) – conducted in two population-based studies with indiviuals of European ancestry and African American individuals, this study found 126 proteins associated with baseline eGFR. There is an overlap of 43 proteins independently reported in both ours and the work by (Ngo et al., 2020), including well-known proteomic biomarkers and novel candidates like testican-2. This set of replicated proteins is listed in (Matías-García et al., 2021b).

All in all, our study serves as both a replication of the aforementioned studies, as well represents one of the first two aptamer-based studies of eGFR with internal replication and large sample size. Likewise, sensitivity analyses conducted in the discovery sample showed that the direction of the association and statistical significance in the cystatin C-based eGFR analyses were congruent with those from serum creatinine-based eGFR. This offered evidence on the relevance of these biomarkers to kidney function beyond the potential biases caused by the use of serum creatinine to estimate GFR.

To investigate whether genetic susceptibility to renal function (using eGFR as a proxy thereof) or plasma protein levels may have a causal effect on the other, Mendelian randomization (MR) was conducted with publicly available results on GWAS from kidney function and protein levels. To date, only one study using

MR to assess causality of kidney function and proteomic biomarkers has been reported (Mohammadi-Shemirani et al., 2019); however, its focus, studied population and biomarker selection are markedly different to ours. Our MR analyses provided strong evidence of a causal effect of renal function on the plasma levels of testican-2, as this effect was identified across sensitivity analyses that relax the assumptions upon which MR relies (Folkersen et al., 2020). Additional gene expression analyses using transcriptomic data from kidney tissue (Jiang et al., 2020; Rowland et al., 2019; Xu et al., 2018) showed *SPOCK2* expression was negatively associated with tubular atrophy and interstitial fibrosis, both measures of structural damage.

Testican-2, a secreted protein of the SPARC protein family (Clark & Sage, 2008), belongs to a group of matricellular proteins (MCPs) involved in extracellular matrix (ECM)-cell interactions and ECM processing (Feng et al., 2019). This protein has been detected in urine (Marimuthu et al., 2011), whereas its renal release into the bloodstream (arterial-to-renal venous gradients, V/A $>$ 1) (Ngo et al., 2020) suggests changes in its plasma levels may reflect glomerular filtration alterations (Schenk et al., 2008). *SPOCK2*, the gene coding for this protein, has been described in relation to both glomerular remodeling and maintenance of tissue integrity and wound healing mechanisms (Francki & Sage, 2001). Interestingly, and in line with the observations we made based on cross-sectional data, higher testican-2 plasma levels have also been associated with less eGFR loss over time and reduced odds of incident CKD in the aforementioned recent aptamer-based study (Ngo et al., 2020). Enriched in human glomeruli in comparison to tubuli samples (Lindenmeyer et al., 2010; Woroniecka et al., 2011) and other non-renal tissues (Nystrom et al., 2009), *SPOCK2* has been reported as a glomerular and podocyte-specific gene (Ju et al., 2013; Lindenmeyer et al., 2010), whereas evidence from immunohistochemistry and immunofluorescence of human kidney tissue show glomerular expression and podocyte-specific expression in adult human kidney samples at single-cell resolution (Ngo et al., 2020).

We also found suggestive evidence of a causal effect of three proteins (MIA, cystatin M and carbonic anhydrase III) on eGFR. Importantly, these effects were inferred using one single SNP as genetic instrument and thus no sensitivity analyses could be conducted, meaning the evidence of a causal effect is weak in these

cases. Although available information on the roles these proteins play in kidney function is limited, there is evidence these proteins are involved in murine kidney morphogenesis (Schwab et al., 2003), legumain regulation in extracellular matrix remodelling and fibronectin deposition (Morita et al., 2007; Van Vliet et al., 2001), and oxidative damage in proximal tubule dysfunction (Gailly et al., 2008). Thus, these causal effects identified by MR may be biologically plausible. However, the molecular mechanisms by which these proteins may be exerting effects on eGFR, if existing, are to be explored in appropriate experimental models following up on these initial findings.

4.1.2 About our contribution to the field

The evidence from cross-sectional population-based studies, the causal effects estimated by MR and the associations with histologic measures all support the notion that testican-2 (and *SPOCK2*) may be interesting novel biomarkers of kidney function and disease. As observed in our analyses, low plasma levels of testican-2 are associated with poor renal function. An independent study reported low plasma testican-2 levels to be predictive of incident CKD (Ngo et al., 2020). Together, the evidence at hand suggests testican-2 could be used as a biomarker of early kidney disease and/or progression (Christensson et al., 2017; Ngo et al., 2020). However, the utility of testican-2 as a biomarker with regard to its tissue of origin, the mechanisms influencing its blood levels and its potential clinical utility require further study.

4.2 DNAmAge acceleration and kidney traits

In the study of DNAmAge acceleration, we identified associations with multiple kidney traits and different measures of epigenetic age acceleration (Matías-García et al., 2021a).

4.2.1 About our findings in relation to current literature

We identified 23 associations between kidney traits and DNAm-based aging/mortality predictors in a large meta-analysis from up to five multi-ethnic population-based cohorts. These associations can be grouped into three different groups. The first

group consists of those DNAm-based predictors that were associated with all studied parameters of poor kidney health: PhenoAA, MRS and EEAA. The second group, consisting of the "first-generation" clocks (Hannum and Horvath), showed distinct patterns of association: HorvathAA was only associated with lower eGFR, while HannumAA was associated with all kidney traits but serum urate. Finally, the third group consisted of the associations with GrimAge acceleration (GrimAA), which was associated with higher uACR and serum urate, as well as prevalent microalbuminuria.

Our findings may be explained in relation to premature systemic and kidney aging (Rowland et al., 2018; Shiels et al., 2017). Premature aging is driven, among other lifestyle and environmental factors, by an increased allostatic load – to which immunosenescence, systemic inflammation ('inflammaging') and oxidative stress contribute (Franceschi & Campisi, 2014; Kooman et al., 2014; Mueller et al., 2020; Tecklenborg et al., 2018). Some of the molecular mechanisms involved in oxidative stress and chronic inflammation specific to renal aging include mitochondrial dysfunction, ureamic-induced epigenetic changes, as well as the production of reactive oxygen species in the glomeruli by pro-inflammatory factors leading to barrier function impairment and albuminuria (Joosten et al., 2020; Kooman et al., 2014; McCarthy et al., 1998; Young & Wu, 2012).

Recent evidence suggests the 10-CpG MRS captures mortality risks mediated by oxidative stress (X. Gao, Gào, et al., 2019) and inflammation-driven changes in immune cell counts (Ward-Caviness et al., 2020). Likewise, both EEAA and PhenoAA are also extrinsic aging measures (i.e. tracking changes in blood cell composition) that may nevertheless mirror intrinsic (i.e. cell composition independent) aging-related physiological dysregulation (Horvath & Raj, 2018; Levine et al., 2018). Moreover, these three DNAm-based predictors have shown a stronger predictive association with time to death than the "first-generation" predictors and intrinsic measures of age acceleration, suggesting they better reflect biological aging (Chen et al., 2016; Zhang et al., 2017) - in line with the associations observed with all kidney traits included in this study.

The "first-generation" clocks are thought to reflect distinct aspects of aging, in

line with the distinct association patterns we observed. On the one hand, Horvath was developed as a "pan-tissue" clock, capturing changes in ubiquitous cell-intrinsic pathways independent of changes in blood cell composition (Horvath, 2013). Hannum, on the other hand, is strongly correlated with changes in immune cell counts and imunosenescence (Chen et al., 2016; Marioni et al., 2015). Nevertheless, GWAS on these clocks suggest both estimators include genes involved in metabolic and immune system pathways (Gibson et al., 2019; Lu et al., 2018). Our findings may be thus explained in relation to both DNAm-based predictors being (at least partially) indicators of immune system aging. The associations between GrimAge and serum urate, uACR and albuminuria indicate this DNAm-based predictor may reflect systemic inflammation and earlier signs of renal damage. Our findings, in line with prior reports on albumin excretion in diabetic and non-diabetic subjects (Lu et al., 2019; Roshandel et al., 2020), may be explained by the role high serum urate levels play in immune system aging (Joosten et al., 2020; Kooman et al., 2014) and kidney disease progression (Astor et al., 2011; Matsushita et al., 2010).

Finally, the results observed in relation to the secondary DNAm-estimated markers (i.e. adrenomedullin [DNAmADM], plasminogen activator inhibitor-1 [DNAmPA1] and smoking pack years [DNAmPACKYRS]) are in line with their known biological roles. Individuals with cardiorenal diseases have higher blood levels of ADM and PAI-1; likewise, ADM has been suggested to predict disease progression (Dieplinger et al., 2009; Kronenberg, 2009) and PAI-1 has been proposed as a risk factor for cardiorenal complications (Vaughan, 2005; Yamamoto et al., 2005). Moreover, tobacco smoking is associated with both renal function decline and inflammation (Hall et al., 2016), and its effects in disease progression may be mediated by mechanisms like oxidative stress and endothelial dysfunction (Orth & Hallan, 2008).

All in all, it will be necessary to disentangle the biological and chronological age components of the existing clocks to better understand these and other reported association in the literature. Likewise, better defining disease-specific biological clocks and more precise chronological clocks is a challenge that needs to be overcome to allow for their future use as specialized tools (Bell et al., 2019).

4.2.2 About our contribution to the field

DNAm-based lifespan and aging predictors capture the effects of systemic inflammation and oxidative stress, mechanisms shared by numerous chronic diseases. In kidney disease in particular, immune function dysregulation and higher allostatic load are tightly correlated to changes in the epigenetic landscape and physiological renal homeostasis. Our findings on kidney traits and DNAm-based predictors of aging thus suggest DNAmAge is a promising marker linking immune system decline and inflammageing to chronic kidney disease and renal functional decline. Future studies disentangling the precise molecular mechanisms underlying the observed associations and the potential clinical value of the DNAm-based measures are warranted.

4.3 Strengths and limitations

The strengths of the work included in this dissertation feature the careful adjusting for potential confounders, the inclusion of well-established independent studies encompassing different ethnicities and the large sample size, the assessment of multiple phenotypes (e.g. almost a thousand proteins, up to five kidney traits and several different DNAm-based biomarkers) and the overall novelty of both studies. By conducting several analyses, it was possible to explore the influence of adjusting for confounders in both studies. The inclusion of many independent and well-established population-based studies allowed us to replicate and thus assess the external validity of our results. All biological samples producing the analyzed data were collected following standard operating procedures in well-described population-based cohorts. Likewise, -omic biomarkers were measured using commercially available platforms (i.e. SomaLogic, Illumina DNAm arrays) and their data was processed also according to well-defined workflow schemes. In this sense, and regarding the analytical validity of the aptamer-based and epigenetic predictors of age, our findings seem to be robust. Moreover, the findings presented in both publications show effects with the same direction (e.g. positive epigenetic age acceleration and higher levels of certain proteins associated with low renal function) across studies, thus proving that the identified associations are robust despite potential measurement error, inter- and intra-population variability. We were also

able to comprehensively address biological aging as predicted by different DNAm-based predictors, renal health as evidenced by multiple kidney traits and examine a large number of plasma proteins. This dissertation presents the first reports of confounder-adjusted eGFR-protein and kidney trait-DNAm predictor associations replicated across multi-ethnic independent studies, and as such, offers an initial systematic view of omics and kidney function from a population-based point of view.

Limitations include the phenotype definitions, type of biological samples studied, the observational nature of the data, and issues derived from the analytical approaches followed. Regarding the phenotype definition, although GFR is the best measure of kidney function, it is also an incomplete indicator of kidney disease: it does not reflect earlier changes in other sections of the kidneys (e.g. tubular dysfunction or interstitial damage) and does not fully mirror its pathophysiology (Glassock & Winearls, 2008). Nevertheless, serum creatinine is a marker that is easily determined in blood and is thus often measured in population-based epidemiological cohorts, whereas more specific measures of kidney damage may not be available in such settings. Another potential limitation of this work is that the reported associations are based mostly on data derived from blood rather than kidney tissue. Given the tissue (and even cell type) specificity of DNAm patterns and gene expression, future -omics studies in kidney tissue will be able to address better whether the associations reported are relevant to *in situ* changes and their potential clinical use. Some methodologic issues related to the assessment of oservational data include the possibility of unobserved confounding and presence of biases that cannot be adjusted for, phenotype and exposure misclassification, and the uncertainty associated with the biological meaning of the proteomic and DNAm-based biomarkers.

Regarding the choice of analytical platform used in both studies, a few limitations are to be noted. First, the use of an aptamer-based analytical platform to profile hundreds of plasma proteins, although advantageous because of its multiplexing possibilities and sample throughput, also presents issues in the specificity and cross-reactivity of the probes that have to be explored on a case-by-case basis for all proteins. Alternative methods like ultrasensitive immunoassays, proximity extension assays (PEA), and mass spectrometry (MS) methods can be used

to gain insights into the specificity of reagents used in affinity methods. Furthermore, this platform does not reflect the entire plasma proteome and does not cover post-translational modifications. Secondly, the inclusion of CpGs in the epigenetic clocks is dependent on the coverage offered by the microarray technologies used, which in its latest version with >850,000 CpGs offers a low genomic coverage of up to 3% of the 28 million CpG sites in the human genome. More informative biological clocks might be derived from methods with a better genomic resolution, as well as from the study of other epigenetic features like other DNA modifications, chromatin marks and RNAi mechanisms.

Finally, we assessed the associations between kidney traits and omics-based data one -omic layer at a time, whereas an integrative approach to this topic may offer different additional insights. Moreover, given the complex nature of kidney disease, research of mechanistic insights will require -omics data collected at multiple timepoints (Hasin et al., 2017) and cell/animal models to disentangle these mechanisms.

5 Conclusion and Outlook

By conducting two large-scale epidemiological studies in multi-ethnic population-based cohorts, we were able to identify novel associations between plasma proteomic biomarkers and DNAm-based measures of aging and kidney function (Matías-García et al., 2021a; Matías-García et al., 2021b). To further contribute to the ever-growing literature on population-based -omics studies, our findings have been made publicly available as original research articles following open-access principles. Our contributions, though subject to some limitations, contribute to narrowing the knowledge gaps outlined in Chapter 1 and suggest potential future research directions.

The coverage of the -omics markers and kidney traits assessed could be extended by using other proteomic and DNAm technologies, studying larger sample sizes and patient samples with individuals showing varying severity degrees of kidney disease. Likewise, an integrative approach to the study of -omics in kidney disease may offer results that are more readily applicable in clinical practice. The use of open-source information to identify whether either proteins or the CpGs in the clocks have been targeted in clinical trials, and if so whether their indications reflect renal biology or if they may be potential targets for related (drug) indications may also offer interesting insights. The identification of specific causal mechanisms will, however, necessarily involve additional studies in cellular and animal models, as well as dedicated longitudinal studies in human populations.

As single-cell technology becomes more readily available for their application in large-scale studies, it will be interesting to explore whether the different proteins and the epigenetic mechanisms described in this dissertation can be related to other -omic layers and gene expression signatures at the single-cell level to identify cell types driving these reported associations. In the meantime, the use of cell-type-

deconvolution algorithms to assess the cell type specificity of the DNAm-based signatures measured in bulk assays in blood (or even in kidney tissue) may also offer specific insights into their role in disease progression.

Emerging consortia dedicated to the assessment of specific molecular phenotypes in relation to health outcomes and their significant contributions to their respective fields will further push the current frontiers of knowledge. Moreover, the establishment of new biobanks and population-based cohorts at the national level across countries will enable future large-scale epidemiological research and novel collaborations. Our findings are product of academic collegiality and the collaboration with epidemiologic cohorts, as outlined throughout this work. Standing on the shoulders of giants, this dissertation presents insights into DNAm-based measures of aging and proteomics in relation to kidney function, and paves the way for future research.

Appendix A

Publication: Plasma proteomics of renal function: a trans-ethnic meta-analysis and Mendelian randomization study

CLINICAL EPIDEMIOLOGY www.jasn.org

Plasma Proteomics of Renal Function: A Transethnic Meta-Analysis and Mendelian Randomization Study

Pamela R. Matías-García,[1,2,3,4] Rory Wilson,[1,2] Qi Guo,[5] Shaza B. Zaghlool,[6] James M. Eales,[7] Xiaoguang Xu,[7] Fadi J. Charchar,[8,9,10] John Dormer,[11] Haifa Maalmi,[12,13] Pascal Schlosser,[14] Mohamed A. Elhadad,[1,2,4] Jana Nano,[2,13] Sapna Sharma,[1,2] Annette Peters,[2,4,13] Alessia Fornoni,[15] Dennis O. Mook-Kanamori,[16] Juliane Winkelmann,[17,18] John Danesh,[19,20,21,22,23,24] Emanuele Di Angelantonio,[19,20,21,22,23] Willem H. Ouwehand,[20,25,26,27] Nicholas A. Watkins,[26] David J. Roberts,[21,28,29] Agnese Petrera,[30] Johannes Graumann,[31,32] Wolfgang Koenig,[4,33,34] Kristian Hveem,[35,36] Christian Jonasson,[35,36] Anna Köttgen,[14,37] Adam Butterworth,[5] Marco Prunotto,[38] Stefanie M. Hauck,[30] Christian Herder,[12,13,39] Karsten Suhre,[6] Christian Gieger,[1,2,4] Maciej Tomaszewski,[7,40] Alexander Teumer,[41,42] Melanie Waldenberger,[1,2,4] and Human Kidney Tissue Resource*

Due to the number of contributing authors, the affiliations are listed at the end of this article.

ABSTRACT

Background Studies on the relationship between renal function and the human plasma proteome have identified several potential biomarkers. However, investigations have been conducted largely in European populations, and causality of the associations between plasma proteins and kidney function has never been addressed.

Methods A cross-sectional study of 993 plasma proteins among 2882 participants in four studies of European and admixed ancestries (KORA, INTERVAL, HUNT, QMDiab) identified transethnic associations between eGFR/CKD and proteomic biomarkers. For the replicated associations, two-sample bidirectional Mendelian randomization (MR) was used to investigate potential causal relationships. Publicly available datasets and transcriptomic data from independent studies were used to examine the association between gene expression in kidney tissue and eGFR.

Results In total, 57 plasma proteins were associated with eGFR, including one novel protein. Of these, 23 were additionally associated with CKD. The strongest inferred causal effect was the positive effect of eGFR on testican-2, in line with the known biological role of this protein and the expression of its protein-coding gene (*SPOCK2*) in renal tissue. We also observed suggestive evidence of an effect of melanoma inhibitory activity (MIA), carbonic anhydrase III, and cystatin-M on eGFR.

Conclusions In a discovery-replication setting, we identified 57 proteins transethnically associated with eGFR. The revealed causal relationships are an important stepping stone in establishing testican-2 as a clinically relevant physiological marker of kidney disease progression, and point to additional proteins warranting further investigation.

JASN 32: 1747–1763, 2021. doi: https://doi.org/10.1681/ASN.2020071070

The kidneys' ability to filter blood and maintain homeostasis is reflected in the GFR.[2] Blood levels of serum creatinine, a filtration marker, can be used for eGFR.[3,4] CKD, characterized by reduced eGFR (<60 ml/min per m^2) and proteinuria, has a global prevalence of 10% to 16%[5,6] and is expected to be increasingly common in aging populations.[2] Increased serum creatinine is not evident until approximately 50% of the renal filtration function is lost,[7] making CKD a silent disease and creating a blind spot for early kidney disease detection.[8] Its rising prevalence, in addition to the lack of

therapeutic options,[8] imposes a significant burden on health systems and individuals worldwide.[2,6]

A number of biomarker research studies have been conducted in regard to early detection, diagnosis, and/or progression prediction of kidney diseases.[7,8] Early efforts in proteome research focused on urine biomarkers, where combining multiple urinary biomarkers was successful (*e.g.*, CKD273 classifier).[7,9] More recent studies have focused on blood, an easily accessible tissue mirroring the metabolic status of multiple organs, with a complex profile requiring sensitive techniques for its study. SOMAscan, a platform using DNA aptamers to measure hundreds of plasma proteomic biomarkers,[10] has been successfully used in different epidemiological settings.[11-15] However, kidney disease has not been sufficiently investigated: prior studies have tested a limited number of proteins,[16] relied on small samples without replication,[10,17] or have not investigated causality.[18] Mendelian randomization (MR), a causal inference method relying on the random allocation of alleles at conception to estimate causal effects on outcomes, is an increasingly popular method used in genetic epidemiology studies to address causality.[19,20]

We present a cross-sectional study using a multiplexed aptamer-based proteomics platform to investigate associations between 1095 plasma proteins and eGFR/CKD, and other renal parameters in a discovery cohort (Cooperative Health Research in the Region of Augsburg S4 prospective cohort follow-up; KORA F4), with replication in three independent studies of European and admixed ancestry (INTERVAL, Nord-Trøndelag Health Study [HUNT], and Qatar Metabolomics Study on Diabetes [QMDiab]). To better understand the biological significance of the identified proteins, we conducted enrichment, protein, and transcriptome analyses across tissues, and investigated their interconnection using protein-protein interaction (PPI) network analysis. We also investigated causal effects between eGFR and the replicated proteins using two-sample bidirectional MR. We further examined the correlation between their gene expression in kidney tissue, eGFR, and histological parameters using both publicly available datasets and transcriptomic data from a kidney resource.

Received July 24, 2020. Accepted March 22, 2021.

*Investigators who supported/contributed to recruitment and/or phenotyping of human kidney studies reported in reference 1 are Wojciech Wystrychowski, Monika Szulinska, Andrzej Antczak, Maciej Glyda, Robert Krol, Joanna Zywiec, Ewa Zukowska-Szczechowska, Pawel Bogdanski, and Bernard Keavney.

Published online ahead of print. Publication date available at www.jasn.org.

Correspondence: Pamela R. Matías-García and Melanie Waldenberger, Research Unit of Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München, Ingolstädter Landstraße 1, 85764 Munich/Neuherberg, Germany. Email: pamela.matias@helmholtz-muenchen.de and waldenberger@helmholtz-muenchen.de

Significance Statement

Studies on the plasma proteome of renal function have identified several biomarkers, but have lacked replication, were limited to European populations, and/or did not investigate causality with eGFR. Among four cohorts in a transethnic cross-sectional study, 57 plasma proteins were associated with eGFR, 23 of them also with CKD. Furthermore, Mendelian randomization and gene expression analyses in kidney tissue highlighted testican-2 as a physiological marker of kidney disease progression with potential clinical relevance, and identified a few additional proteins warranting further investigation.

METHODS

Study Populations

The KORA study is a population-based sample from the general population living in the region of Augsburg, Southern Germany. The KORA F4 survey, a follow-up of the KORA S4 prospective cohort (1999–2001), was conducted from 2006 to 2008, and included a total of 3080 participants. Clinical and demographic information, and peripheral blood for "omics" analyses, were collected; details on the standardized examinations, interviews, and tests conducted in the KORA study have been previously described.[21,22] This study acted as discovery cohort in the cross-sectional association study of plasma proteins and renal function (Figure 1A).

Included in the replication phase were HUNT, namely the third survey (HUNT3) from this population-based study from Norway with data on participants of European descent;[23] the INTERVAL study, a randomized trial assessing blood donation practices across the United Kingdom with extensive phenotyping available for 50,000 participants of European descent;[24] and the QMDiab, a cross-sectional case-control study on type 2 diabetes from participants of Arab, South Asian, and Filipino descent in Qatar.[25] Population characteristics from the four studies are shown in Table 1. Information on data availability is given in Supplemental Note 1.

Sample Collection and Proteomic Profiling

EDTA plasma samples collected by the studies following standardized procedures were centrifuged, aliquoted, and stored at -80°C.[26-28] Samples for proteomic profiling and GFR estimation were taken at the same time.

Proteomic profiling in all participating studies was done using SOMAscan (SomaLogic, Inc.), an aptamer-based, affinity proteomics platform.[10,29-31] Plasma samples from KORA F4, HUNT3, and INTERVAL were shipped on dry ice to SomaLogic (Boulder, CO), and proteomic profiling was performed using a SOMAscan panel with 1129 protein-specific SOMAmer probes for KORA,[26] 3622 for INTERVAL,[27] and 5000 for HUNT3.[28] In the QMDiab cohort, the kit-based SOMAscan platform was run by the Weill Cornell Medicine Qatar proteomics core following protocols and instrumentation provided by SomaLogic Inc., under supervision of

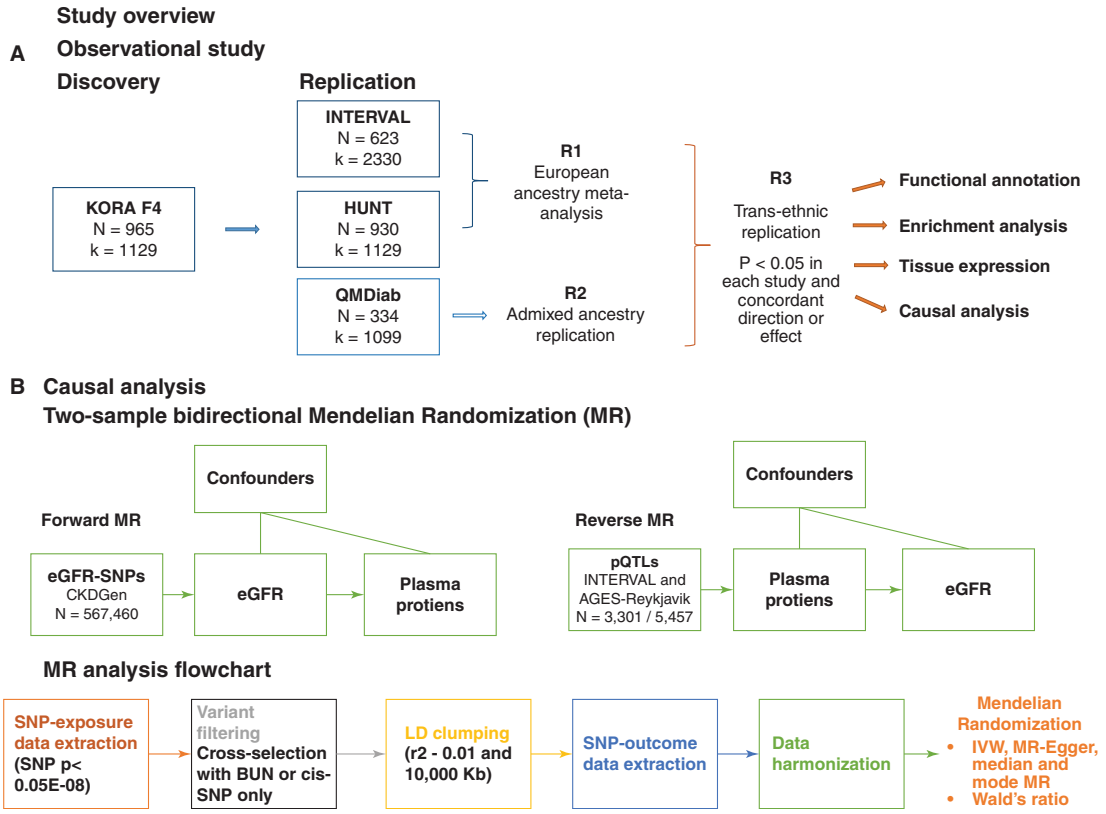

Figure 1. (A) Cross-sectional association study. Data from 995 participants and 1129 proteins from KORA F4 was used in the discovery phase of a proteome-wide association study of renal function using confounder-adjusted regression models. The replication studies were INTERVAL, HUNT, and QMDiab. Three rounds of replication are shown: R1, replication based on the meta-analysis of *P* values from the linear regression results of the studies with European ancestry; R2, replication based on the results of linear regression models performed in the Arab, South Asian, and Filipino descent sample QMDiab; R3, identification of proteins consistently associated with eGFR across samples and ethnicities. The set of proteins identified in R3 was then functionally annotated and brought forward to the causal analysis phase. (B) Causal analysis. Two-sample bidirectional MR using data on participants from European ancestry studies in the Chronic Kidney Disease (CKDGen) Consortium to instrument the forward analysis (eGFR causal to protein level) and data from INTERVAL and AGES-Reykjavik to instrument the reverse analysis (protein level causal to eGFR). Details on the data processing workflow for MR are shown.

SomaLogic personnel, to measure 1129 proteins in plasma samples.[26] The samples were all measured by individuals blinded to the identities corresponding to the samples. In summary, fluorescently labeled single-stranded synthetic nucleotides (Slow Off-rate Modified Aptamers, SOMAmers) immobilized on streptavidin-coated beads are incubated with plasma samples to capture proteins and generate SOMAmer-protein complexes. Washing steps eliminate unbound SOMAmers and unbound/nonspecifically bound proteins. The next steps are biotin labeling and photocleavage to liberate SOMAmer-protein complexes from the beads. This is followed by incubation in a buffer disrupting nonspecific interactions, recapturing the biotin-labeled protein/aptamer complexes in streptavidin-coated beads, and additional washing steps to remove nonspecific SOMAmers. These are then eluted from the target proteins and quantified on custom DNA microarrays using deposited SOMAmer-complementary oligonucleotides, which produces measurements in relative fluorescence units as proxies to protein concentrations. Quality control (QC) at the sample and SOMAmer levels using control aptamers and calibrator samples was conducted by the manufacturer. In brief, based on standard samples included on each plate, the resulting raw intensities were processed using a workflow including hybridization normalization, median signal normalization, and signal calibration to control for interplate differences. In the discovery cohort (KORA F4), QC resulted in

Table 1. Population characteristics of association studies

Trait	KORA	HUNT3	INTERVAL	QMDiab
N	995	930	623	334
Age, yr	59.31 (7.81)	68.94 (10.29)	47.36 (13.35)	47.10 (12.57)
Male	480 (48.2)	688 (74.0)	343 (55.1)	169 (50.6)
BMI, kg/m^2	27.78 (4.58)	28.38 (3.97)	27.16 (10.05)	29.66 (5.95)
Smokers	572 (57.5)	699 (75.16)	99 (15.89)	60 (18.0)
Serum creatinine (mg/dl)	0.85 (0.18)	0.92 (0.32)	0.70 (0.14)	0.85 (0.22)
eGFR, ml/min per 1.73m^2	85.98 (14.06)	80.25 (18.75)	108.27 (16.21)	96.00 (18.36)
CKD	38 (3.8)	138 (14.8)	1 (0.16)	19 (5.7)
UACR, mg/dl	5.64 (3.61, 9.94)	NA	NA	NA
MA	58 (5.9)	NA	NA	NA
HDL cholesterol, mg/dl	57.32 (15.20)	45.12 (11.24)	74.33 (24.42)	47.58 (13.75)
Triglycerides, mg/dl	107 (75, 155.5)	141.27 (106.28, 194.85)	132.75 (97.35,194.70)	169 (99.20, 215.23)
Lipid lowering medication use	142 (14.3)	NA	33 (5.30)	NA
Hypertension	397 (39.9)	355 (38.2)	48 (7.70)	103 (30.8)
Diabetes	68 (6.8)	128 (13.8)	2 (0.32)	172 (51.5)

Measurement units are shown in parentheses in the trait column, where the absence of units means it is a categorical trait. The mean and (SD) are presented for non-skewed continuous variables, whereas for skewed continuous variables median (first and third quartile) are presented. Count and % are shown for categorical variables.

the exclusion of 29 proteins and one sample, and five proteins were further excluded due to crossreactivity (Assay Change Log SSM-064_Rev_0_DCN_16-263 issued by SomaLogic, available at the integrated web-server at http://proteomics.gwas.eu),[26] producing data on $k=1095$ proteins in 999 participants (Supplemental Table 1). The same QC conducted in each study resulted in the inclusion of 3301 participants in INTERVAL,[27] 2432 individuals in HUNT3,[12] and 352 participants in QMDiab,[26] and a set of 993 proteins passing QC in all studies (column "R" in Supplemental Table 1); further details on the proteomics profiling from the samples included in this study are described elsewhere.[12,26,27] Protein mapping to several identifiers was provided by the manufacturer (Supplemental Table 1).

Outcome Definitions

Our first analysis is a proteome wide association study: we investigated associations between proteins and renal traits as outcomes, using linear regression models with adjustment for potential confounders. The primary outcomes studied in this analysis were eGFR from serum creatinine and CKD, given their availability in all included studies.

eGFR was calculated using the CKD Epidemiology Collaboration equation with serum creatinine[4] with the R package *nephro* v1.2.[32] Serum creatinine was measured using the modified kinetic Jaffé reaction in KORA, HUNT, and QMDiab (and calibrated by multiplying by 0.95),[33] and a nuclear magnetic resonance platform (Nightingale Health) in the INTERVAL study. Pearson's correlation between serum creatinine–based eGFR and this nuclear magnetic resonance–based eGFR variable was estimated in KORA (Supplemental Figure 1). CKD was defined as eGFR <60 ml/min per 1.73 m^2.[34]

We performed some analyses for outcomes available only in the discovery study. Urinary albumin and urinary creatinine were used to calculate urinary albumin-creatinine ratio (uACR) and its derived parameter microalbuminuria (MA, defined as uACR >30 mg/g). eGFR decline was defined as $\log(\text{eGFR})_{\text{follow-up}} - \log(\text{eGFR})_{\text{baseline}}$ divided by the follow-up time, where KORA F4 (2006–2008) was used as baseline and KORA FF4 (2013–2014) as its follow-up survey. Sensitivity analyses were also run using eGFR_{cys} (derived from the CKD Epidemiology Collaboration equation using cystatin C).[13]

Definition of Covariates

Covariates used in the regression analyses were age at the time of examination, sex, body mass index (BMI), smoking status, diabetes (yes/no), hypertension (yes/no), log-transformed triglycerides, HDL, and intake of lipid-lowering drugs (yes/no). See Supplemental Note 2 for precise cohort-specific definitions of covariates used.

Statistical Analysis

Data preprocessing and statistical analyses were conducted using the R language for statistical computing v.3.6.0.[35] Before statistical analysis, proteomic data were log transformed and standardized. Linear regression was used to examine the association between protein levels and continuous kidney traits (log-transformed eGFR, uACR, eGFR change), whereas logistic regression was used for binary kidney traits (CKD, MA). Multiple testing was accounted for using a Bonferroni correction considering the total number of investigated proteins at each stage ($k=1095$ in discovery).

Sensitivity analyses in the discovery sample included regression models with serum creatinine–based eGFR as an outcome and no adjustment for BMI or diabetes, and models including cystatin C–based eGFR as outcome and the same set of covariates from the main model. Pearson's correlations between the regression coefficients resulting from the sensitivity and the main analyses were calculated. Interaction analyses were also conducted for the proteins identified at discovery by adding an interaction term (each of age, sex, and smoking

status individually) to the fully adjusted model (Supplemental Note 3).

For those protein-outcome pairs significantly associated in the discovery, two replications were conducted: a European replication (R1) and a replication in an admixed population (R2) (Figure 1A). Replication was defined as $P<0.05$ and consistent direction of effect as in the discovery study. The European replication for eGFR consisted of the meta-analysis of results from HUNT and INTERVAL using Stouffer's method, a *P* value combination method especially useful when raw data cannot be pooled across studies, which is the case with aptamer-based measurements, where data in relative fluorescence units is not directly comparable across studies. Also known as "inverse normal" or weighted *Z*-test, this method takes the *P* values for the i-th study (pi), transforms them by using the inverse normal transformation, and weights them according to the square root of the sample sizes (wi). The sum is then computed, and the combined *P* value is obtained using the distribution of the resulting statistic, $T=\Sigma wiH(pi)$.[36]

For CKD, only the HUNT study was used in the European replication (INTERVAL had only one patient), and the admixed population replication was based on the results of QMDiab. Our final set of transethnic associations (R3) were those pairs of proteins-outcomes that were replicated in both R1 and R2. Replicated eGFR-associated proteins were taken to the next stages of the analysis: proteomic target validation, enrichment analyses, and MR.

Validation of Proteomic Targets

We examined the plasma levels of proteins measured using Proximity Extension Assay (PEA) technology (Olink) in a subgroup of randomly selected participants from the KORA F4 study ($n=173$).[37,38] In brief, protein abundance was quantified using real-time PCR in the PEA proteomic technology (Olink), producing relative quantification data in NPX units (normalized protein expression levels, on log2 scale); NPX values were intensity normalized with the plate median for each assay as the normalization factor, and samples and proteins that did not pass QC were excluded.[38] Eight of the most relevant proteins (cystatin C, RELT, IGFBP-6, myoglobin, TNF sR-I, RGMB, FSTL3, contactin-4), and three of the proteins identified in the causal inference analysis (carbonic anhydrase 3, melanoma inhibitory activity [MIA], cystatin M) were available in this subset of proteomic measurements. Of note, testican-2 was not available for measurement using this technology. Scatterplots of the aptamer-based and PEA measurements, annotated with their Pearson's correlations and *P* values, are shown in Supplemental Figure 2. The lack of immunoassays fully validated for specificity, linearity, and possible interference for most of the measured analytes in SOMAScan (including testican-2) limits the investigation into the concordance between these two methods.[39]

Information on specificity and crossreactivity of the aptamers was available from three independent studies[27,40,41] for 54 of the 57 proteins identified to be transethnically associated with eGFR. Target specificity issues (*i.e.*, comparable binding observed to a target that is not the product of the same gene) were observed in four proteins (ephrin-A5, IGFBP-5, hemojuvelin, and cystatin SA)[27,40] (Supplemental Tables 2–4). Moreover, in previous studies, 23 of the 57 proteins were directly validated via mass spectrometry in blood plasma/serum, and other biological matrices[41] (Supplemental Table 2), and 49 using solution affinity measurements[27, 40] (Supplemental Tables 3–4).

Functional Annotation, Enrichment, and Expression Analyses

Annotation was conducted using the R package *InterMineR* v1.6.1,[42] a tool facilitating access to data from the HumanMine release 6.0 (May 2019). DAVID v.6.8[43] was used to look for annotations for Gene Ontology Terms (molecular function, biological process) and pathway information, and to identify publications relevant to the set of 57 replicated proteins. Gene information was retrieved from the human assembly GRCh37 (hg19) using BioMart v.4,[44] and shown in Supplemental Table 5.

To investigate the expression patterns of the 57 eGFR-associated proteins and their corresponding protein coding genes across tissues, we used proteomics and RNA-seq expression data from the ProteomicsDB[45,46] and the Genotype-Tissue Expression (GTEx) database.[47] The data presented and described in this manuscript were generated on October 2, 2020 through a multigene query on the ProteomicsDB Analytics Toolbox portal from: https://www.proteomicsdb.org/proteomicsdb/#analytics/expressionHeatmap and GTEx portal https://www.gtexportal.org/home/multiGeneQueryPage.

PPI Network Analysis

We queried STRING,[48] the PPI server, to examine the relationship between the proteins that were identified as robustly associated with eGFR across studies and ethnicities ($k=57$ transethnically eGFR-associated proteins). We used the set of SOMAscan proteins available across studies as background ($k=993$), adding no additional interactors (proteins) to the network during the analyses, and considered a minimum required interaction score for a medium confidence (0.400) (Supplemental Note 4).

MR

MR, an instrumental variable method used to infer causality, leverages the natural randomization inherent in the (random) assortment of genes during gamete formation to assess the effect of lifelong exposures on health outcomes.[49] Single-nucleotide polymorphisms (SNPs) are used as instrumental variables (IV; or instruments), given their alleles are randomly assigned to individuals before any exposures/outcome and they are nonmodifiable, thus minimizing the risk of reverse causation and confounding.[49] The idea behind MR is that if

genetic variation produces differences mirroring the biological effects of environmental exposures that alter disease risk, then genetic variation itself should be related to disease risk by having an influence on the exposure.[49,50] MR uses SNPs as surrogates for an exposure of interest, allowing the estimation of the effects of life-long, genetically determined "exposures" on health outcomes.[49] MR produces robust causal inference estimates if the SNPs used are valid instruments—that is, if they meet the three assumptions on which MR relies: SNPs must be strongly associated with the exposure, and not associated with either (measured or unmeasured) confounders or with the outcome except potentially through the exposure.[51] Causality in MR is thus defined as the modification of an exposure leading to a change in the outcome, where the inferred causal effects by MR do not necessarily imply the existence of a straightforward interpretation with respect to direct causal factors,[50] nor do they offer information on the time interval (*e.g.*, during development) or target tissue in which such modification of the exposure or intervention would need to be delivered.[52]

To investigate whether a genetic liability to lower or higher eGFR causally alters plasma protein levels or *vice versa*, MR was conducted in the set of 57 proteins whose associations with eGFR showed transethnic replication. Two-sample bidirectional MR[19] was used to infer the causal effect of renal function (eGFR as proxy thereof) on plasma protein levels (forward MR) and *vice versa* (reverse MR, Figure 1B). Results from publicly available genome-wide association studies (GWAS) for (*1*) eGFR from the CKDGen consortium (meta-analysis of European-ancestry populations),[53] and (*2*) plasma proteins from INTERVAL[27] and Age, Gene/Environment Susceptibility-Reykjavik Study (AGES-Reykjavik)[41] were used to perform MR using MRBase.[54] A detailed account on the MR methods, data sources, and analyses conducted is available in Supplemental Note 5.

Instrument Selection

In the forward MR (*i.e.*, assessing the effect of renal filtration on protein levels), 256 SNPs associated with eGFR at genome-wide significance in the CKDGen results were selected as candidate IV. These SNPs were then filtered based on their relevance to renal function (associated with BUN, a complementary renal trait, with an opposite direction of effect, $N_{IV}=47$) and clumped based on linkage disequilibrium ($r2=0.01$ and $Kb=10{,}000$) to identify independent variants ($N_{IV}=41$). Summary statistics on 41 SNP-eGFR associations were extracted from the CKDGen results, and corresponding SNP-protein associations were extracted from the INTERVAL results for 47 proteins. For investigating the causal effects of eGFR on proteins, 47 eGFR-protein relationships were instrumented by 41 SNPs.

For the reverse MR (*i.e.*, interrogating the causal effect of proteins on renal filtration), gene positions (GChr37, Supplemental Table 5) were used to identify genome-wide significant cis-SNPs for 28 proteins in the INTERVAL results as candidate IV and LD clumped (same criteria as forward MR). Summary statistics on SNP-protein associations for 28 proteins were extracted from the INTERVAL results, and its corresponding SNP-eGFR associations were extracted from the CKDGen results. The same strategy was followed to identify instruments in the AGES-Reykjavik results; SNP-protein results were extracted from this dataset for 29 proteins, and SNP-eGFR results were extracted for 26 proteins from the CKDGen data. Further details of the genetic instrument selection and data harmonization process are shown in Supplemental Figure 3 and Supplemental Table 6. Thus, for investigating the causal effect of proteins on eGFR, 35 protein-eGFR relationships were instrumented by 1–5 SNPs, of which 17 proteins were examined using data from both INTERVAL and AGES-Reykjavik (Supplemental Figure 4).

Data Harmonization, Phenotypic Variance Explained, and Instrument Specificity

Details on data harmonization, the handling of palindromic SNPs, and calculating the phenotypic variance explained by the SNPs are given in Supplemental Note 5. Harmonized datasets used in the MR analyses are available in Supplemental Table 7.

To look for further evidence of horizontal pleiotropy, association between our SNPs and other traits were searched for in the GWAS Catalog[55] (Supplemental Table 8).

MR and Sensitivity Analyses

The primary MR analysis used inverse variance weighted (IVW) regression. In this method, the coefficient of the gene-outcome association is regressed on the coefficient of the gene-exposure association with the intercept constrained to zero, assuming no directional pleiotropy.[56,57] Because IVW requires two or more SNPs, in patients where only one SNP instrumented the analysis, Wald's ratio (coefficient of the gene-outcome association divided by the gene-exposure association) was calculated whenever only one SNP instrumented the analyses, as IVW MR requires two or more SNPs.[57]

For MR analyses instrumented by more than two SNPs, three further MR methods were used as sensitivity analyses.[58] MR-Egger regression was used to assess pleiotropy, because this method allows for horizontal pleiotropy and provides an estimate of the unbalanced horizontal pleiotropic effects in its intercept.[59] Weighted median[60] and weighted mode MR,[61] methods less sensitive to the presence of invalid instruments and to pleiotropic SNPs behaving as outliers, were also used. A number of additional analyses were run to check for outliers, directional pleiotropy and heterogeneity, as recommended.[58,62] Details are given in Supplemental Note 5.

Causal estimates were assessed at a Bonferroni-corrected significance level, namely, 0.05 divided by the number of proteins assessed in each MR direction (47 in forward and 51 in reverse MR). Causal effects were considered robust if they were significant at Bonferroni $P<0.05$ in the IVW or Wald estimator, and results from the pleiotropy-robust sensitivity

MR analyses examined to test for violations to MR assumptions.

Expression Analyses in Human Kidney Tissue

The correlation between expression of *SPOCK2,* one of the genes coding for proteins showing evidence for a causal relationship with eGFR, was calculated with data from microdissected tubulointerstitial components of human renal biopsies from 26 individuals with CKD at different disease stages (I–IV)[63] (GEO accession: GSE69438). Gene expression of the protein-coding genes identified in MR (*SPOCK2*, *CA3*, *CST6*, *MIA*) and renal traits was further assessed in (*1*) data from *Nephroseq* v5 (n=458), a platform of comprehensive kidney disease gene expression datasets,[64] and (*2*) a human kidney tissue resource based on RNA sequencing (n=427, see Supplemental Note 6).[1]

Within *Nephroseq*, univariate correlation analyses between eGFR and gene expression were conducted separately in study-defined histological compartments of the human kidney (*i.e.*, glomerular and tubulointerstitial) in 458 available kidney samples from three datasets of patients with kidney disease (Ju *et al.*, Sampson *et al.*, and Reich *et al.*),[63,65,66] and one dataset of "apparently" healthy renal tissue (Rodwell *et al.*).[67] The correlations were meta-analyzed using inverse variance weighted random effects models,[68] and heterogeneity was assessed using Cochran's Q test.

Multivariable regression analyses were conducted in the human kidney resource (n=427).[1] In brief, we constructed linear regression models with renal expression of each candidate as the response variable; whereas eGFR and histologically confirmed measures of structural kidney damage were used as independent variables together with age, sex, BMI, three genetic principal components, diabetes, and a variable number of surrogate variables (29 for eGFR and 26 for all histology phenotypes).[1,69] eGFR was based on circulating levels of creatinine, as reported before.[1] Histologic measures of structural integrity (glomerular sclerosis, glomerular Bowman's capsule thickening, tubular atrophy, interstitial fibrosis, interstitial inflammation, and vascular lesions) were assessed microscopically and scored on a semiquantitative scale (whereby 0 indicates no or minimal damage and 3 is consistent with the highest degree of structural injury), as reported previously.[70]

RESULTS

Figure 1 illustrates the design of this study. First, a cross-sectional association study was performed to identify proteins associated with renal function parameters in a discovery-replication setting: KORA F4 acted as the discovery, and INTERVAL, HUNT3, and QMDiab as replication studies (Figure 1A). Replicated transethnic protein associations were then assessed for causality using two-sample MR, using data from the largest GWAS available for the traits of interest (CKDGen, INTERVAL, and AGES-Reykjavik) (Figure 1B).

CROSS-SECTIONAL ASSOCIATION OF PLASMA PROTEINS AND RENAL FUNCTION

Population characteristics of the four cohorts included in the cross-sectional association study are shown in Table 1.

Results from the Discovery Study

The association between 1095 plasma proteins and eGFR/CKD was assessed in the KORA F4 study (n=995). A total of 80 proteins were significantly associated with eGFR (P<0.05/1095) (Supplemental Figure 5A). The top three negative associations (*i.e.*, higher eGFR associated with lower plasma protein levels) were observed with cystatin C (β=-0.068; 95% confidence interval [95% CI], -0.078 to -0.059] change in log-transformed eGR per standard deviation increase in protein level, P=2.63E-40), TNF receptor superfamily member 19L (RELT; β=-0.063; 95% CI, -0.073 to -0.053; P=7.82E-33) and β2-microglobulin (β=-0.059; 95% CI, -0.070 to -0.050, P= 6.16E-30), and the strongest positive association (*i.e.*, higher eGFR associated with higher plasma protein levels) was that of testican-2 (β=0.036; 95% CI, 0.026 to 0.045, P=2.066E-13) (Supplemental Table 1). Of note, 34 of these 80 proteins were also associated with CKD (Supplemental Table 1).

Sensitivity analyses showed that 71 of the 80 eGFR-associated-proteins identified in the main analysis were consistently associated with cystatin C–based eGFR, with a high correlation between regression coefficients (r=0.841, P<0.001). Models with no adjustment for BMI or diabetes produced highly similar estimates to those obtained in the main analysis (r=0.99, P<0.001 for both; Supplemental Table 9). Likewise, the exclusion of individuals with CKD (n=38) did not significantly affect the correlation between the plasma levels of proteins and eGFR (Supplemental Figures 6 and 7). Interaction analyses suggested associations with five plasma proteins were accentuated with age (Supplemental Table 10).

Additional renal outcomes were assessed in the discovery cohort: eGFR change was associated with five proteins, three proteins were associated with uACR, and no proteins were associated with MA (Supplemental Table 11).

Results from Replication Studies

Serum creatinine was the only available trait across all replication cohorts (Figure 1A), thus only associations with eGFR/CKD were further explored.

The European replication (R1) was conducted using HUNT3 and INTERVAL; results from this analysis confirmed the association of 62 of the 76 proteins available across studies (Supplemental Figure 5). The second replication round (R2) was performed in QMDiab, a population of admixed ancestry;

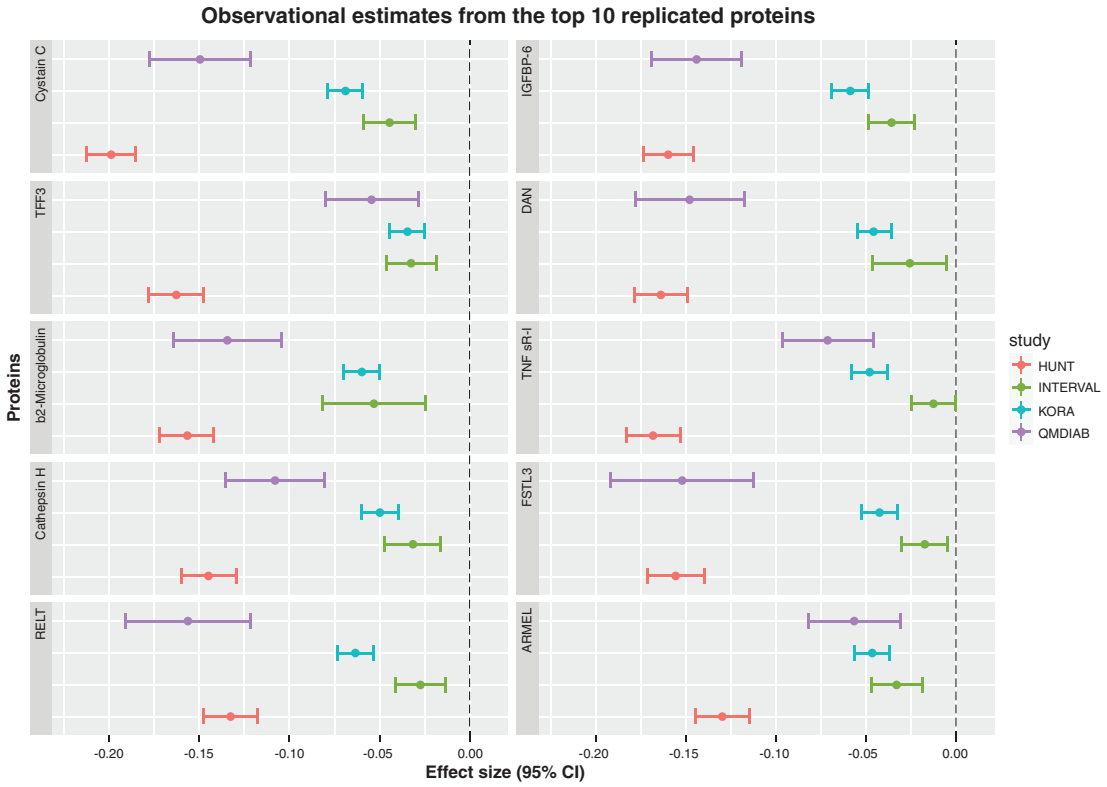

Figure 2. Regression coefficient estimates from the top 10 proteins identified in the cross-sectional association transethnic study on eGFR. The *x* axis shows the estimates and 95% CI for the regression coefficients (*i.e.*, change in log-transformed eGFR per standard deviation increase in protein level), and each panel corresponds to one protein. Estimates are color coded according to the specific study: HUNT3 in red, INTERVAL in green, KORA in blue, and QMDiab in purple. TFF3, Trefoil factor 3; RELT, TNF receptor superfamily member 19L; IGFBP-6, Insulin-like growth factor-binding protein 6; DAN, Neuroblastoma suppressor of tumorigenicity 1; TNF sR-I: TNF receptor superfamily member 1A; FSTL3, Follistatin-related protein 3; ARMEL, Cerebral dopamine neurotrophic factor.

this confirmed 65 eGFR-protein associations (Supplemental Figure 5). High correlations between *z* values from the discovery study and replication studies were observed (correlations ranging from 0.66 with INTERVAL to 0.93 with HUNT3, Supplemental Figure 8). The overlap of the proteins replicated in R1 and R2 produced the set of 57 robustly replicated transethnic eGFR-protein associations (Supplemental Table 12). Figure 2 shows the cross-sectional effect estimates for the top 10 protein-eGFR associations across the four cohorts; INTERVAL, a largely healthy and younger population, showed the smallest effect sizes, whereas the strongest effects were observed in HUNT3, a cohort of older individuals with lower mean eGFR and higher CKD prevalence. One novel protein, contactin-4, was identified. All 57 proteins were replicated in the cystatin C–based eGFR sensitivity analysis (Supplemental Table 9).

All 34 CKD-protein associations from the discovery phase were replicated in HUNT3 and 23 in QMDiab; these 23 were thus considered transethnically robust (Supplemental Table 13). Figure 3 shows the overlap between proteins associated with eGFR/CKD.

Functional Annotation, Enrichment and Expression Analyses

Several pathways, biological processes and molecular functions were represented in the set of replicated proteins (Supplemental Table 14). No enrichment was observed, perhaps due to the coverage of the analytical platform.[10,30,71] Peptides for most of the replicated proteins were detected in multiple tissues and body fluids, including kidney tissue (Supplemental Figure 9). Most genes showed ubiquitous

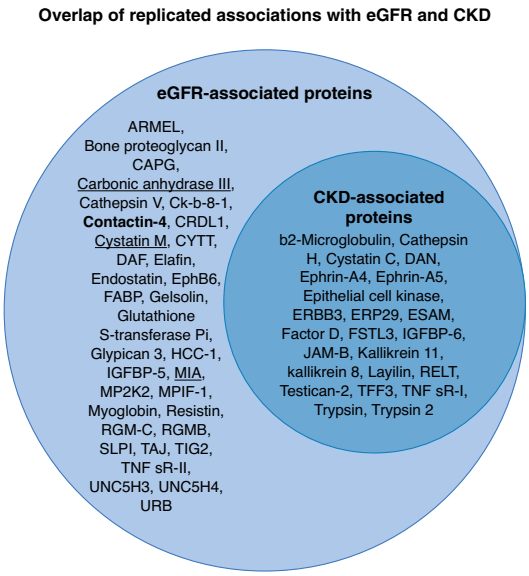

Figure 3. Results from the transethnic discovery-replication observational study. Depicted in the left circle are the 57 proteins associated with eGFR, the continuous measurement of renal function; the 34 eGFR-specific proteins reflect associations along the full range of renal function, whereas the 23 proteins also associated with CKD reflect a direct association with a clinically relevant low eGFR (<60 ml/min per 1.73m^2). Shown in bold is contactin-4, a novel protein identified by this study, and underlined are the four proteins for which evidence on causal effects was identified by MR.

expression across the human tissues represented in the ProteomeDB (Supplemental Figure 10) and GTEx datasets (Supplemental Figure 11).

PPI Network Analysis

We queried STRING to examine PPIs in the set of 57 replicated proteins. More interactions than expected were observed in the resulting network (P=1.12E-04, Supplemental Figure 12). Because the inclusion of proteins in the main PPI analysis was conditional on the platform's coverage and study design, less stringent sensitivity analyses allowing for the inclusion of additional proteins connected additional nodes (*e.g.*, *SPOCK2* or *MIA*) not connected in the main analysis to the network (Supplemental Note 4).

MR

To assess whether genetically determined higher or lower plasma levels of the 57 proteins identified as transethnically associated with eGFR may affect this renal trait, and whether genetically determined eGFR causally alters circulating levels of plasma proteins, two-sample bidirectional MR was conducted (Figure 1B).

Forward MR: eGFR Has an Effect on Testican-2

In the forward direction of the MR (*i.e.*, inferring the effect of eGFR on levels of 47 proteins), 40 SNPs explaining 1.59% of the variance of eGFR were used as instruments (Supplemental Table 15).

Plasma levels of seven proteins were identified as causally affected by eGFR according to the IVW MR model (Supplemental Table 16 and Supplemental Figures 13–19). Although no evidence of directional pleiotropy, influential SNPs, or instrument heterogeneity was observed (with the exception of IGFB6, Supplemental Table 17–19), pleiotropy-robust sensitivity MR analyses did not provide further evidence of causality for six of them, suggesting IVW findings may be driven by undetected horizontal pleiotropy.[72] In contrast, a positive causal effect of eGFR on testican-2 was identified by multiple MR methods (weighted median P=2.84E-04, Figure 4). Assuming this robust evidence reflects a true relationship, the results suggest that if eGFR is altered by an intervention mimicking the effect of the SNP on eGFR, plasma levels of testican-2 will also increase.

In total, 11 of the SNPs instrumenting the forward MR analysis were identified as potentially pleiotropic (associations at P<5E-08 with other traits, Supplemental Table 8). Restrictive MR was conducted excluding these SNPS; although not significant, perhaps due to reduced statistical power derived from using fewer SNPs as instruments and/or the exclusion of SNPs that might be on the actual causal pathway of interest, these results were in agreement with those from the main analysis (same direction and size of effect, Supplemental Table 20).

Reverse MR: MIA, Cystatin M, and Carbonic Anhydrase III Affect eGFR

In the reverse direction of the MR (*i.e.*, assessing the effect of 35 proteins on eGFR), one to five cis-SNPs explaining 0.91%–29.33% of phenotypic variance were used as genetic instruments (Supplemental Table 15).

A negative effect of MIA on eGFR was identified by multiple MR models (Table 2), results that suggest if plasma protein levels are lowered by means of an intervention mimicking the effect of the SNP on MIA, eGFR will increase. No evidence of influential SNPs was observed (Supplemental Tables 17–19), yet the funnel plot suggested directional pleiotropy (Supplemental Figure 20). One SNP instrumenting this analysis was identified as potentially pleiotropic (Supplemental Table 8).

Positive effects of carbonic anhydrase III and cystatin M on eGFR were also identified (Wald's ratio P=5.04E-04 and 8.41E-05, respectively) (Table 2). Although further sensitivity analyses could not be conducted given the availability of one

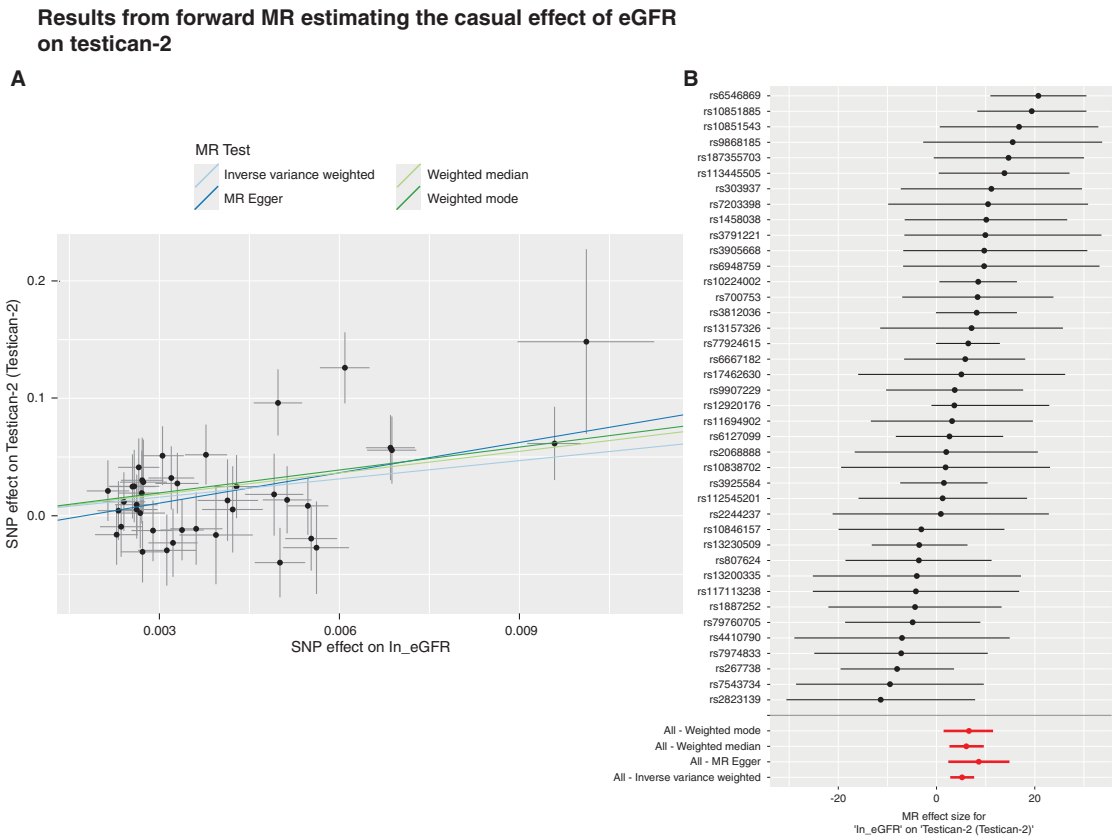

Figure 4. (A) Scatter plot showing the individual genetic effects of the selected IVs on log transformed eGFR (coefficient of the SNP-exposure association) on the *x* axis and on testican-2 plasma levels (coefficient of the SNP outcome) on the *y* axis, along with their 95% CI. Each data point corresponds to an individual SNP. The lines correspond to the slopes of the different MR methods, which can be interpreted as the change in testican-2 levels per unit increase in log-transformed eGFR, and are color coded as follows: IVW-MR in light blue, MR-Egger in dark blue, weighted median in light green, weighted mode in dark green. (B) Forest plot showing the individual causal estimates of each of the 40 genetic instruments. The red points show the pooled estimates using all SNPs in the four methods. 95% CI are shown.

cis-SNP for each protein, no gene-trait associations were found for these SNPs in the GWAS Catalog,[55] indicating a lack of evidence for pleiotropic effects. MR estimates obtained with different pGWAS data (see Methods, Supplemental Figure 4) had the same direction of effect (Supplemental Table 16). Of note, no statistically significant effect of testican-2 on eGFR was identified (Supplemental Table 16).

Gene Expression in Kidney Tissue

No correlation between eGFR and expression of *SPOCK2* (protein-coding gene for testican-2) in tubulointerstitial components of human renal biopsies from 26 individuals with CKD[63] was found (Supplemental Figure 21). Further univariate analyses conducted with *Nephroseq* data showed a statistically significant correlation between eGFR and *SPOCK2* gene expression in glomerular compartment/kidney cortex (r=0.242, P=0.033) (Supplemental Table 21). We then conducted multivariable analyses using RNA-sequencing–characterized human kidney transcriptomes from up to 427 individuals[1,69,70] (Supplemental Note 6, Supplemental Table 22). Whereas no association between gene expression and eGFR was observed (Supplemental Table 23), *SPOCK2* expression was negatively associated with tubular atrophy and interstitial fibrosis (P=0.03 for both), and *CST6* expression was negatively associated with glomerular sclerosis (P=0.02, Supplemental Table 23).

Table 2. Causal estimates across MR methods

Association	IVW	MR-Egger	Weighted Median	Weighted Mode	Wald Ratio
eGFR → *SPOCK2*					
β	5.21	8.61	6.10	6.51	
CI	2.78 to 7.66	2.41 to 14.8	2.81 to 9.40	1.19 to 11.8	–
P	2.95E-05	0.01	2.84E-04	0.021	
N_{IV}	40	40	40	40	
CA3 → eGFR					
B					0.007
CI	–	–	–	–	0.003 to 0.01
P					5.04E-04
N_{IV}					1
CST6 → eGFR					
B					0.007
CI	–	–	–	–	0.004 to 0.01
P					8.41E-05
N_{IV}					1
MIA → eGFR					
β	-0.002	-0.001	-0.002	-0.001	
CI	-0.003 to -0.001	-0.002 to 0.000	-0.003 to -0.001	-0.002 to -0.001	–
P	8.79E-04	0.299	2.00E-04	0.081	
N_{IV}	3	3	3	3	

Results from the forward MR (effect of eGFR on protein levels, *i.e.* eGFR → protein) are based on the 40 instruments retrieved from Wuttke *et al.*, 2019,[53] whereas the reverse MR (protein → eGFR) are based on the one to three instruments retrieved from the INTERVAL pGWAS reported in Sun *et al.*, 2018.[27] In bold are significant *P* values at a Bonferroni-corrected level (0.05/47 for the forward analysis, 0.05/28 for the reverse analysis). *SPOCK2*, testican-2; *CA3*, carbonic anhydrase III; *CST6*, cystatin-M; *MIA*, melanoma-derived growth regulatory protein; *β*, causal estimate.

DISCUSSION

We conducted a cross-sectional association study of plasma proteomics and eGFR/CKD in four independent cohorts, identifying known and potential novel biomarkers. Two-sample bidirectional MR suggested the existence of causal effects in four eGFR-protein associations.

A total of 80 proteins were associated with eGFR in our discovery analysis, with transethnic replication confirming 57 of these; 23 were also found to be associated with CKD. Although our analyses use serum creatinine–based eGFR due to its availability across studies and its utility in clinical practice, models using cystatin C–based eGFR show the results are robust to the GFR estimation method. Likewise, further sensitivity analyses indicate the associations are largely independent of adjustment for BMI or diabetes.

Additional analyses in the discovery cohort produced associations between eGFR decline and DAN, TNF sR-1 and FSTL3, in line with those previously reported.[18] No overlap between the set of proteins associated with uACR and the proteins previously reported[73] was observed, which may be explained by the different time points of the eGFR and albuminuria measurements.[73] No proteins were significantly associated with MA, possibly due to its low prevalence (5.9%) in KORA.

We identify several well-known biomarkers of renal function,[10,16,17,74,75] supporting the validity of our eGFR analyses. Our results are also in line with previous proteomic studies on eGFR,[10,16-18] inflammation in ESKD,[75] and a "standalone" renal health test[40] (Supplemental Table 12). Contactin-4, involved in neuronal network development, was identified as a novel eGFR-associated protein. Its plasma (but not urine) detection suggests it is either not filtered at the glomerular capillaries, or filtered but later reabsorbed into blood from the tubules, so that variations in its plasma levels may reflect changes in glomerular and tubule function.[76,77] The age-interaction effects identified for four proteins may be a consequence of their age-varying trajectories,[78] and require future investigation. The ubiquitous expression of our proteins across tissues, and our PPI network results, suggests the proteins are involved in cellular functions relevant to multiple tissues (Supplemental Table 14),[79] potentially mirroring the systemic nature of kidney disease.[80] Interestingly, podocyte-exosome enrichment in urine identified 23 of our proteins, pointing to their participation in processes underlying glomerular filter permeability.[81]

To investigate whether genetically determined renal function (using eGFR as a proxy thereof) or plasma protein levels may have a causal effect on the other, bidirectional MR with publicly available GWAS data was conducted. Our findings in the forward MR direction, supported by pleiotropy-robust sensitivity MR methods, identified a causal effect of eGFR on plasma levels of testican-2. Considering the MR definition of causality, these results suggest lifestyle or pharmacological interventions designed to improve eGFR (as a proxy of renal function) have the potential to increase plasma testican-2 levels.

Testican-2 is a secreted protein of the SPARC family,[82] a group of matricellular proteins regulating extracellular matrix–cell interactions and extracellular matrix processing,[83]

and is involved in a number of biological processes (Supplemental Table 24). In line with our results, higher testican-2 plasma levels have also been associated with less eGFR loss over time and reduced odds of incident CKD.[18] Its protein-coding gene, *SPOCK2,* is associated with both normal maintenance of organ and tissue integrity, and with wound healing and other responses to injury.[85] Despite the ubiquitous expression of this gene, its enriched expression in human glomeruli in comparison to other nonrenal tissues[18,86-89] and its renal downregulation in diabetic kidney disease[86] suggest it has a particularly relevant role in renal cellular mechanisms. Furthermore, recent evidence from arteriovenous sampling demonstrated renal release of this protein into the bloodstream,[18] which in addition to its urine detection,[84] suggests changes in its plasma levels may be indicative of kidney function.[18,76] However, the contribution other organs might have in its plasma levels cannot be ruled out.

Although the association between eGFR and *SPOCK2* renal expression did not reach statistical significance in some of our analyses (which may be partly explained by low statistical power in the dataset with 26 patients with CKD), the directionality of the coefficient (*i.e.*, positive association) was consistent across datasets. Moreover, multivariable analyses showed higher scores of histologic measures of renal structural damage to be negatively associated with *SPOCK2* renal expression, consistent with prior evidence.[86]

All in all, the agreement between the cross-sectional results reported by us and by others,[17,18,40] and our MR findings and the associations with histologic measures, indicate testican-2 and its protein-coding gene *SPOCK2* may have an important role in kidney function. Low plasma levels of testican-2 may be indicative of poor renal function, meaning this protein may be a physiological biomarker of kidney health and disease progression.[17,18] Although the reverse direction of this causal association did not reach statistical significance in our MR analyses, a reverse causal effect (*i.e.*, testican-2 on renal function) is biologically plausible given the role matricellular proteins play in extracellular matrix repair[83,90] and its *in vitro* effects on human glomerular endothelial cells.[18] The utility of testican-2 as a biomarker with regard to its potential functional effects, tissue of origin, or the mechanisms influencing its blood levels, requires further study.

Three proteins (MIA, cystatin M, and carbonic anhydrase III) were identified as potentially having a causal effect on eGFR, effects biologically plausible given their known roles (Supplemental Table 24). Nevertheless, the precise mechanisms through which these proteins could be exerting effects on eGFR remain to be elucidated. Discordant directions of effect from the observational and the causal estimates (cystatin M, carbonic anhydrase III) could be explained due to differences in sample size/characteristics, reverse causation, or confounding in the case of the observational estimates, or due to limitations inherent to the MR methods.[91,92] A further explanation might be that they represent different effects: MR examines lifelong exposures to higher or lower protein levels, whereas results from observational studies could be reflective of acute or short-term effects.[92]

The strengths of our study include the use of a multiplex proteomics platform and large sample size. This is the first report of eGFR-protein associations adjusted for multiple potential confounders, replicated in independent samples of diverse ancestries, and assessed using causal inference. MR was conducted with the largest available GWAS results from nonoverlapping European ancestry populations, thus avoiding issues derived from population stratification and sample overlap. We reduced the possibility of horizontal pleiotropy by using GWAS summary statistics from a complementary renal trait (blood urea nitrogen) to improve the specificity of the genetic instruments for eGFR, by focusing on cis-SNPs, and by using multiple pleiotropy-robust sensitivity analyses.

Our study also has several limitations. Aptamer-based proteomic methods may be affected by probe cross-reactivity and nonspecific binding,[79] although the aptamer-based measurements of most of the reported proteins have been validated in multiple independent studies.[27,40,41] This platform does not produce absolute plasma concentrations or cover post-translational modifications, limiting the interpretability of the regression coefficients and the scope of the studied plasma proteome.[79] Our findings are based on cross-sectional data, so studies examining their longitudinal changes are warranted. Despite the multiethnic nature of our study, our results may not extend to ethnic groups not represented in our analyses. We avoided weak instrument bias in MR, but cannot discount the possibility of having incurred selection bias in the case of the SNP-protein data. The sample size in which genetic associations with protein levels were calculated was significantly smaller than the sample used to identify genetic associations with eGFR, which likely resulted in differences in power. Finally, knowledge of the biological role of the proteins identified is insufficient for our findings to suggest mechanistic insights. A follow-up of our findings in appropriate experimental models would provide additional evidence on the inferred causal associations reported here and help to unravel the molecular mechanisms underlying our findings. Likewise, future validation studies using validated absolute quantitative assays with increased sensitivity for the detection of testican-2, and other proteins identified here, are warranted to establish reference ranges and to explore their suitability as prognostic and diagnostic biomarkers in clinical settings.

In summary, our transethnic population-based study of plasma proteomics and renal function identified multiple markers of kidney function. Our MR findings are a stepping stone in establishing testican-2 as a physiological marker of kidney disease progression, and further identify proteins warranting additional investigation. Our results may serve as the starting point for future translational work on the utility of these proteins as diagnostic or prognostic biomarkers of disease, and for research on mechanistic insights at the tissue and single-cell levels.

DISCLOSURES

A. Butterworth reports receiving research funding from AstraZeneca, Bayer, Biogen, BioMarin, Bioverativ, Merck, Novartis, and Sanofi; and reports receiving honoraria from Novartis. A. Fornoni reports having consultancy agreements with Dimerix, Gilead, Janssen, Novartis, ONO Pharmaceutical, and Zyversa Therapeutics; reports having an ownership interest as CSO and Vice-President of L&F Health LLC., being a shareholder in River 3 Renal Corp and Zyversa Therapeutics; reports receiving research funding from Boheringer Ingelheim and Roche; reports having patents and inventions for the use of cyclodextrin for the treatment of kidney diseases, a patent for use of small molecule inducers of cholesterol efflux; reports being a scientific advisor or member of the *Journal of Clinical Investigation* and *Kidney International*; and reports having other interests/relationships as an inventor on five pending US patents and one published patent. A. Kottgen reports receiving honoraria from Sanofi Genzyme; repost being a scientific advisor or member of the *American Journal of Kidney Diseases*, American Kidney Fund, *Journal of the American Society of Nephrology*, Kidney International, and Nature Reviews Nephrology. A. Teumer reports being a scientific advisor or member of the Editorial Board of the *Journal Endocrine Connections*; reports having other interests/relationships as a member of the American Society of Human Genetics. C. Herder reports receiving research funding from Sanofi-Aventis; reports receiving honoraria from Lilly, Sanofi-Aventis; and reports being a scientific advisor or member of the Editorial Boards for Diabetologia, Diabetic Medicine, and Diabetes/Metabolism Research and Reviews (DMRR). C. Jonasson reports receiving personal fees for research consultancy work from Bayer and Pfizer outside of the submitted work. E. Di Angelantonio reports receiving research funding from a British Heart Foundation research grant, a National Health Service Blood and Transplant research grant, National Institute for Health Research grant, and a United Kingdom Medical Research Council research grant; reports being a scientific advisor or member as Chair of the working group for the European Society of Cardiology Cardiovascular Risk Collaboration, Member of the World Obesity Federation and World Heart Federation Expert Group on Obesity and CVD, Member of the National Health Service Blood and Transplant Clinical Trial Unit Steering Committee, European Society of Cardiology/European Atherosclerosis Society Task Force for Guidelines on Primary Prevention of Cardiovascular Disease and Management of Dyslipidaemia, and Member of the World Health Organization Risk Chart Working Group. J. Dormer reports consultancy agreements with Royal College of Pathologists; reports being on the AstraZeneca Genomics Advisory Board (2018), International Cardiovascular and Metabolic Advisory Board for Novartis (since 2010), the International Cardiovascular and Metabolism Research and Development Portfolio Committee for Novartis, the Medical Research Council International Advisory Group (ING) member, London (since 2013), the MRC High Throughput Science 'Omics Panel Member, London (since 2013); the Scientific Advisory Committee for Sanofi (since 2013), and the Steering Committee of UK Biobank (since 2011). M. Prunotto is employed by Galapagos Ltd. M. Waldenberger reports being a scientific advisor or member of the Editorial board of Genes and *International Journal of Molecular Science.* Q. Guo is employed by BenevolentAI. S. Zaghlool reports receiving research funding from the Qatar Foundation. W. Koenig reports having consultancy agreements with Amgen, AstraZeneca, Corvidia, DalCor, Daiichi-Sankyo, Genentech, Kowa, Novartis, Pfizer, and The Medicines Company; reports receiving research funding from Abbott, Beckmann, Roche Diagnostics, Singulex; reports receiving honoraria from Amgen, AstraZeneca, Berlin-Chemie, Bristol-Myers Squibb, Novartis, and Sanofi; reports being a scientific advisor or member of the Advisory Boards for Amgen, AstraZeneca, Corvidia, Daiichi-Sankyo, DalCor, Esperion, Genentech, Kowa, Novartis, Pfizer, The Medicines Company; and reports being an Editorial Board member for Cardiovascular Drugs and Therapy and Clinical Chemistry. J. Danesh reports Scientific Advisor or Membership with the Scientific Advisory Board of Oxford BHF Centre for Research Excellence (2020), Scientific Advisory Board for Nightingale Health (2020), Genomics Advisory Board for AstraZeneca (2018), 2015 International Cardiovascular and Metabolism Research and Development Portfolio Committee for Novartis, MRC Global Health Group member in London (2014-16), Scientific Advisory Committee for Sanofi (2013), MRC High Throughput Science Omics panel member in London (2013), and MRC International Advisory Gro (2013). D. Roberts reports Honoraria from Wiley. All remaining authors have nothing to disclose. The HUNT part of the project re-used protein data that were originally analyzed and paid for by Somalogic Inc., CO, USA. Somalogic had no role in the design and conduct of the study; collection of phenotypic data, statistical analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication.

FUNDING

The KORA study was initiated and financed by the Helmholtz Zentrum München German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research and the State of Bavaria. This work was also supported by Weill Cornell Medicine in Qatar Biomedical Research Program, the Qatar Foundation, and Qatar National Research Fund grant NPRPC11-0115-180010 (to K. Suhre). The HUNT Study is a collaboration between HUNT Research Centre (Norwegian University of Science and Technology), Nord-Trøndelag County Council, Central Norway Health Authority, and the Norwegian Institute of Public Health. J. Danesh is supported by the National Institute for Health Research Senior Investigator Award. RNA-sequencing experiments and kidney gene expression studies were supported by British Heart Foundation project grants PG/17/35/33001 and PG/19/16/34270, and Kidney Research UK grants RP_017_20180302 and RP_013_20190305 (to M. Tomaszewski). The German Diabetes Center is funded by the German Federal Ministry of Health (Berlin, Germany), the Ministry of Culture and Science of the state North Rhine-Westphalia (Düsseldorf, Germany), and grants from the German Federal Ministry of Education and Research (Berlin, Germany) to the German Center for Diabetes Research. The work of A. Köttgen is supported by he Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 431984000 – SFB 1453.

ACKNOWLEDGMENTS

C. Geiger, P. Matías-García, M. Waldenberger, R. Wilson, and J. Winkelmann conceptualized and designed the study; J. Graumann, D. Mook-Kanamori, H. Maalmi, and A. Petrera performed experiments; F. Charchar, J. Eales, J. Graumann, Q. Guo, P. Matías-García, M. Prunotto, M. Tomaszewski, M. Waldenberger, R. Wilson, X. Xu, and S. Zaghlool analyzed data or interpreted results; A. Butterworth, F. Charchar, J. Danesh, E. Di Angelantonio, J. Dormer, J. Eales, M. Elhadad, A. Fornoni, C. Geiger, J. Graumann, Q. Guo, S. Hauck, C. Herder, Human Kidney Tissue Resource, K. Hveem, C. Jonasson, W. Koenig, A. Köttgen, H. Maalmi, D. Mook-Kanamori, W. Ouwehand, A. Peters, A. Petrera, M. Prunotto, D. Roberts, P. Schlosser, S. Sharma, K. Suhre, A. Teumer, M. Tomaszewski, M. Waldenberger, N. Watkins, J. Winkelmann, X. Xu, and S. Zaghlool contributed reagents/materials/analysis tools; P. Matías-García wrote the paper; F. Charchar, J. Danesh, J. Eales, J. Graumann, Q. Guo, S. Hauck, C. Herder, W. Koenig, A. Köttgen, J. Nano, A. Petrera, P. Schlosser, A. Teumer, M. Tomaszewski, M. Waldenberger, R. Wilson, X. Xu, and S. Zaghlool edited and improved the clarity of the manuscript; all authors discussed the results and reviewed and approved the final manuscript. We are grateful to all study participants of KORA, HUNT, INTERVAL, and QMDiab for their invaluable contributions to these studies, and all members of field staff conducting the studies. An early version of the abstract was sent to the 53rd European Society of Human Genetics Conference, and is listed online as P03.24.C at the European Society of Human Genetics Conference 2020.2 Abstract Library (https://2020.eshg.org/index.php/abstract-library/). The visual abstract was created using templates and images modified from Servier Medical Art (https://smart.servier.com/), which are licensed under a

ETHICS APPROVAL AND INFORMED CONSENT

All participants provided written consent. Details on institutions providing ethics approval from each study are: the KORA cohort ethical approval was granted by the ethics committee of the Bavarian Medical Association (REC reference number F4: 06068); HUNT: Regiona Committee for Medical Research Ethics (REK) and Data Inspectorate; INTERVAL: National Research Ethics Service Committee East of England - Cambridge East (Research Ethics Committee (REC) reference 11/EE/0538); and QMDiab: Institutional Review Boards of HMC and Weill Cornell Medicine Qatar under research protocol number 11131/11).

SUPPLEMENTAL MATERIAL

This article contains the following supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.13039/100007273/-/DCSupplemental.

Supplemental Note 1. Data availability.

Supplemental Note 2. Covariate definition.

Supplemental Note 3. Interaction analyses.

Supplemental Note 4. PPI network analysis.

Supplemental Note 5. MR analysis.

Supplemental Note 6. Expression analyses in human kidney samples.

Supplemental Table 1. Results from cross-sectional analysis of eGFR and CKD in KORA F4.

Supplemental Table 2. Validation of proteomic targets from Emilsson *et al.*, 2018.

Supplemental Table 3. Validation of proteomic targets from Sun *et al.*, 2018.

Supplemental Table 4. Validation of proteomic targets from Williams *et al.*, 2019.

Supplemental Table 5. Gene information from proteins included in MR.

Supplemental Table 6. Genetic instrument selection and data harmonization (INTERVAL and AGES- Reykjavik).

Supplemental Table 7. Harmonized summary statistics used in MR.

Supplemental Table 8. Gene-trait information retrieved from GWAS Catalog for pleiotropic SNPs.

Supplemental Table 9. Sensitivity analyses with CysC-based eGFR, no adjustment for BMI and no adjustment for T2D.

Supplemental Table 10. Sensitivity analyses (interaction with age, sex, and smoking) in KORA F4.

Supplemental Table 11. Results from observational analysis of supplementary renal phenotypes (eGFR decline, log(uACR) and MA) in KORA F4.

Supplemental Table 12. Replication of cross-sectional eGFR-protein associations.

Supplemental Table 13. Replication of cross-sectional CKD-protein associations.

Supplemental Table 14. Extended annotation file (DAVID).

Supplemental Table 15. Phenotypic variance explained by MR instruments.

Supplemental Table 16. Results from MR (IVW, MBE, weighted median, and MR-Egger).

Supplemental Table 17. Sensitivity analyses: Heterogeneity (Cochran's Q test).

Supplemental Table 18. Sensitivity analyses: Pleiotropy in MR-Egger.

Supplemental Table 19. Sensitivity analyses: Leave-one-out analyses.

Supplemental Table 20. Sensitivity analyses: Results from restrictive MR.

Supplemental Table 21. Results from correlation analyses between gene expression and eGFR from Nephroseq datasets.

Supplemental Table 22. Clinical characteristics of studies included in gene expression analyses.

Supplemental Table 23. Results from multivariate regression analyses on gene expression, eGFR and histological characteristic scoring from human kidney resource.

Supplemental Table 24. Description and biological roles of selected proteins.

Supplemental Figure 1. Correlation of serum creatinine variables in KORA F4.

Supplemental Figure 2. Correlation between aptamer-based and other measurements for proteins in KORA F4.

Supplemental Figure 3. Genetic instrument selection and data harmonization.

Supplemental Figure 4. Protein overlap in pGWAS datasets used in reverse direction of MR.

Supplemental Figure 5. Proteins and log(eGFR) distribution in discovery dataset.

Supplemental Figure 6. Cross sectional results for eGFR-protein associations across studies.

Supplemental Figure 7. Proteins and log(eGFR) distribution in discovery dataset after CKD exclusion.

Supplemental Figure 8. Correlation between Z-values for eGFR-protein associations across studies.

Supplemental Figure 9. Tissue expression of 57 eGFR-associated proteins (ProteomeDB).

Supplemental Figure 10. Tissue expression of 56 eGFR-associated protein coding genes (ProteomeDB).

Supplemental Figure 11. Expression of 56 eGFR-associated protein coding genes across tissues (GTEx).

Supplemental Figure 12. PPI network of 57 replicated eGFR-associated proteins.

Supplemental Figures 13–19. Forward MR results for effects of eGFR on proteins.

Supplemental Figure 20. Reverse MR analysis for MIA-eGFR.

Supplemental Figure 21. *SPOCK2* gene expression in renal tissue from 26 patients with CKD.

REFERENCES

1. Jiang X, Eales JM, Scannali D, Nazgiewicz A, Prestes P, Maier M, et al: Hypertension and renin-angiotensin system blockers are not associated with expression of angiotensin-converting enzyme 2 (ACE2) in the kidney. *Eur Heart J* 41: 4580–4588, 2020
2. Eckardt KU, Coresh J, Devuyst O, Johnson RJ, Kottgen A, Levey AS, et al: Evolving importance of kidney disease: from subspecialty to global health burden. *Lancet* 382: 158–169, 2013
3. Levey AS, Inker LA, Coresh J: GFR estimation: From physiology to public health. *Am J Kidney Dis* 63: 820–834, 2014
4. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, et al: A new equation to estimate glomerular filtration rate. *Ann Intern Med* 150: 604–612, 2009
5. Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al: Global prevalence of chronic kidney disease: A systematic review and meta-analysis. *PLoS One* 11: e0158765, 2016
6. Levey AS, Atkins R, Coresh J, Cohen EP, Collins AJ, Eckardt KU, et al: Chronic kidney disease as a global public health problem: Approaches and initiatives - a position statement from Kidney Disease Improving Global Outcomes. *Kidney Int* 72: 247–259, 2007
7. Mischak H, Delles C, Vlahou A, Vanholder R: Proteomic biomarkers in kidney disease: Issues in development and implementation. *Nat Rev Nephrol* 11: 221–232, 2015
8. Sanchez-Nino MD, Sanz AB, Ramos AM, Fernandez-Fernandez B, Ortiz A: Clinical proteomics in kidney disease as an exponential technology: Heading towards the disruptive phase. *Clin Kidney J* 10: 188–191, 2017

9. Good DM, Zurbig P, Argiles A, Bauer HW, Behrens G, Coon JJ, et al: Naturally occurring human urinary peptides for use in diagnosis of chronic kidney disease. *Molecular Cellular Proteomics* 9: 2424–2437, 2010
10. Gold L, Ayers D, Bertino J, Bock C, Bock A, Brody EN, et al: Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One* 5: e15004, 2010
11. Ngo D, Sinha S, Shen D, Kuhn EW, Keyes MJ, Shi X, et al: Aptamer-based proteomic profiling reveals novel candidate biomarkers and pathways in cardiovascular disease. *Circulation* 134: 270–285, 2016
12. Ganz P, Heidecker B, Hveem K, Jonasson C, Kato S, Segal MR, et al: Development and validation of a protein-based risk score for cardiovascular outcomes among patients with stable coronary heart disease. *JAMA* 315: 2532–2541, 2016
13. Zaghlool SB, Sharma S, Molnar M, Matías-García PR, Elhadad MA, Waldenberger M, et al.: Revealing the role of the human blood plasma proteome in obesity using genetic drivers. *Nature communications*: 12: 1279, 2021
14. Pena MJ, Mischak H, Heerspink HJ: Proteomics for prediction of disease progression and response to therapy in diabetic kidney disease. *Diabetologia* 59: 1819–1831, 2016
15. Elhadad MA, Jonasson C, Huth C, Wilson R, Gieger C, Matias P, et al.: Deciphering the Plasma Proteome of Type 2 Diabetes. *Diabetes* 69: 2766–2778, 2020
16. Carlsson AC, Ingelsson E, Sundstrom J, Carrero JJ, Gustafsson S, Feldreich T, et al: Use of proteomics to investigate kidney function decline over 5 years. *Clin J Am Soc Nephrol* 12: 1226–1235, 2017
17. Christensson A, Ash JA, DeLisle RK, Gaspar FW, Ostroff R, Grubb A, et al: The impact of the glomerular filtration rate on the human plasma proteome. *Proteomics Clin Appl* 12: e1700067, 2018
18. Ngo D, Wen D, Gao Y, Keyes MJ, Drury ER, Katz DH, et al: Circulating testican-2 is a podocyte-derived marker of kidney health. *Proc Natl Acad Sci USA* 117: 25026–25035, 2020
19. Pierce BL, Burgess S: Efficient design for Mendelian randomization studies: Subsample and 2-sample instrumental variable estimators. *Am J Epidemiol* 178: 1177–1184, 2013
20. Sekula P, Del Greco M: F, Pattaro, C, Köttgen, A: Mendelian randomization as an approach to assess causality using observational data. *J Am Soc Nephrol* 27: 3253–3265, 2016
21. Holle, R, Happich, M, Lowel, H, Wichmann, HE: KORA: A research platform for population based health research. *Gesundheitswesen* 67: S19–S25, 2005
22. Wichmann, HE, Gieger, C, Illig, T: KORA-gen–resource for population genetics, controls and a broad spectrum of disease phenotypes. *Das Gesundheitswesen* 67: S26–S30, 2005
23. Krokstad S, Langhammer A, Hveem K, Holmen T, Midthjell K, Stene T, et al: Cohort profile: The HUNT Study, Norway. *Int J Epidemiol* 42: 968–977, 2013
24. Moore C, Bolton T, Walker M, Kaptoge S, Allen D, Daynes M, et al: Recruitment and representativeness of blood donors in the INTERVAL randomised trial assessing varying inter-donation intervals. *Trials* 17: 458, 2016
25. Mook-Kanamori DO, Selim MM, Takiddin AH, Al-Homsi H, Al-Mahmoud KA, Al-Obaidli A, et al: 1,5-Anhydroglucitol in saliva is a noninvasive marker of short-term glycemic control. *J Clin Endocrinol Metab* 99: E479–E483, 2014
26. Suhre K, Arnold M: Connecting genetic risk to disease end points through the human blood plasma proteome. *Nat Commun* 8: 14357, 2017
27. Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al: Genomic atlas of the human plasma proteome. *Nature* 558: 73–79, 2018
28. Nayor M, Short MI, Rasheed H, Lin H, Jonasson C, Yang Q, et al: Aptamer-based proteomic platform identifies novel protein predictors of incident heart failure and echocardiographic traits. *Circ Heart Fail* 13: e006749, 2020
29. Gold L, Walker JJ, Wilcox SK, Williams S: Advances in human proteomics at high scale with the SOMAscan proteomics platform. *N Biotechnol* 29: 543–549, 2012
30. Rohloff JC, Gelinas AD, Jarvis TC, Ochsner UA, Schneider DJ, Gold L, et al: Nucleic acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol Ther Nucleic Acids* 3: e201, 2014
31. Kraemer S, Vaught JD, Bock C, Gold L, Katilius E, Keeney TR, et al: From SOMAmer-based biomarker discovery to diagnostic and clinical applications: A SOMAmer-based, streamlined multiplex proteomic assay. *PLoS One* 6: e26332–e26332, 2011
32. Pattaro C, Riegler P, Stifter G, Modenese M, Minelli C, Pramstaller PP: Estimating the glomerular filtration rate in the general population using different equations: Effects on classification and association. *Nephron Clin Pract* 123: 102–111, 2013
33. Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, et al: Decline in estimated glomerular filtration rate and subsequent risk of end-stage renal disease and mortality. *JAMA* 311: 2518–2531, 2014
34. National Kidney Foundation: K/DOQI clinical practice guidelines for chronic kidney disease: Evaluation, classification, and stratification. *Am J Kidney Dis* 39: S1–S266, 2002
35. R Core Team: *R: A language and environment for statistical computing*, Vienna, Austria, R Foundation for Statistical Computing, 2019
36. Zaykin DV: Optimally weighted Z-test is a powerful method for combining probabilities in meta-analysis. *J Evol Biol* 24: 1836–1841, 2011
37. Huth C, von Toerne C, Schederecker F, de Las Heras Gala T, Herder C, Kronenberg F, et al: Protein markers and risk of type 2 diabetes and prediabetes: A targeted proteomics approach in the KORA F4/FF4 study. *Eur J Epidemiol* 34: 409–422, 2019
38. Petrera A, von Toerne C, Behler J, Huth C, Thorand B, Hilgendorff A, et al: Multiplatform approach for plasma proteomics: Complementarity of olink proximity extension assay technology to mass spectrometry-based protein profiling. *J Proteome Res* 20: 751–762, 2021
39. Raffield LM, Dang H, Pratte KA, Jacobson S, Gillenwater LA, Ampleford E, et al: Comparison of proteomic assessment methods in multiple cohort studies. *Proteomics* 20: e1900278, 2020
40. Williams SA, Kivimaki M, Langenberg C, Hingorani AD, Casas JP, Bouchard C, et al: Plasma protein patterns as comprehensive indicators of health. *Nat Med* 25: 1851–1857, 2019
41. Emilsson V, Ilkov M, Lamb JR, Finkel N, Gudmundsson EF, Pitts R, et al: Co-regulatory networks of human serum proteins link genetics to disease. *Science* 361: 769–773, 2018
42. Kyritsis KA, Wang B, Sullivan J, Lyne R, Micklem G: InterMineR: An R package for InterMine databases. *Bioinformatics* 35: 3206–3207, 2019
43. Huang W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44–57, 2009
44. Durinck S, Spellman PT, Birney E, Huber W: Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc* 4: 1184–1191, 2009
45. Schmidt T, Samaras P, Frejno M, Gessulat S, Barnert M, Kienegger H, et al: ProteomicsDB. *Nucleic Acids Res* 46: D1271–D1281, 2018
46. Samaras P, Schmidt T, Frejno M, Gessulat S, Reinecke M, Jarzab A, et al: ProteomicsDB: A multi-omics and multi-organism resource for life science research. *Nucleic Acids Res* 48: D1153–D1163, 2020
47. GTEx Consortium: The Genotype-Tissue Expression (GTEx) project. *Nat Genet* 45: 580–585, 2013
48. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al: STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 47: D607–D613, 2019
49. Davey Smith G, Hemani G: Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet* 23: R89–R98, 2014

50. Smith GD, Ebrahim S: 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 32: 1–22, 2003
51. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G: Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med* 27: 1133–1163, 2008
52. Bretherick AD, Canela-Xandri O, Joshi PK, Clark DW, Rawlik K, Boutin TS, et al: Linking protein to phenotype with Mendelian randomization detects 38 proteins with causal roles in human diseases and traits. *PLoS Genet* 16: e1008785, 2020
53. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al: A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet* 51: 957–972, 2019
54. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al: The MR-Base platform supports systematic causal inference across the human phenome. *eLife* 7: e34408, 2018
55. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, et al: The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res* 47: D1005–D1012, 2019
56. Burgess S, Butterworth A, Thompson SG: Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol* 37: 658–665, 2013
57. Teumer A: Common methods for performing Mendelian randomization. *Front Cardiovasc Med* 5: 51, 2018
58. Burgess S, Bowden J, Fall T, Ingelsson E, Thompson SG: Sensitivity analyses for robust causal inference from Mendelian randomization analyses with multiple genetic variants. *Epidemiology* 28: 30–42, 2017
59. Bowden J, Davey Smith G, Burgess S: Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol* 44: 512–525, 2015
60. Bowden J, Davey Smith G, Haycock PC, Burgess S: Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol* 40: 304–314, 2016
61. Hartwig FP, Davey Smith G, Bowden J: Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol* 46: 1985–1998, 2017
62. Verbanck M, Chen C-Y, Neale B, Do R: Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet* 50: 693–698, 2018
63. Ju W, Nair V, Smith S, Zhu L, Shedden K, Song PXK, et al: Tissue transcriptome-driven identification of epidermal growth factor as a chronic kidney disease biomarker. *Sci Transl Med* 7: 316ra193, 2015
64. Martini S, Eichinger F, Nair V, Kretzler M: Defining human diabetic nephropathy on the molecular level: Integration of transcriptomic profiles with biological knowledge. *Rev Endocr Metab Disord* 9: 267–274, 2008
65. Sampson MG, Robertson CC, Martini S, Mariani LH, Lemley KV, Gillies CE, et al: Integrative genomics identifies novel associations with APOL1 risk genotypes in Black NEPTUNE subjects. *J Am Soc Nephrol* 27: 814–823, 2016
66. Reich HN, Tritchler D, Cattran DC, Herzenberg AM, Eichinger F, Boucherot A, et al: A molecular signature of proteinuria in glomerulonephritis. *PLoS One* 5: e13451, 2010
67. Rodwell GE, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, et al: A transcriptional profile of aging in the human kidney. *PLoS Biol* 2: e427, 2004
68. Balduzzi S, Rücker G, Schwarzer G: How to perform a meta-analysis with R: A practical tutorial. *Evid Based Ment Health* 22: 153–160, 2019
69. Xu X, Eales JM, Akbarov A, Guo H, Becker L, Talavera D, et al: Molecular insights into genome-wide association studies of chronic kidney disease-defining traits. *Nat Commun* 9: 4800, 2018
70. Rowland J, Akbarov A, Eales J, Xu X, Dormer JP, Guo H, et al: Uncovering genetic mechanisms of kidney aging through transcriptomics, genomics, and epigenomics. *Kidney Int* 95: 624–635, 2019
71. Billing AM, Ben Hamidane H, Bhagwat AM, Cotton RJ, Dib SS, Kumar P, et al: Complementarity of SOMAscan to LC-MS/MS and RNA-seq for quantitative profiling of human embryonic and mesenchymal stem cells. *J Proteomics* 150: 86–97, 2017
72. Richmond RC, Davey Smith G: Commentary: Orienting causal relationships between two phenotypes using bidirectional Mendelian randomization. *Int J Epidemiol* 48: 907–911, 2019
73. Carlsson AC, Sundström J, Carrero JJ, Gustafsson S, Stenemo M, Larsson A, et al: Use of a proximity extension assay proteomics chip to discover new biomarkers associated with albuminuria. *Eur J Prev Cardiol* 24: 340–348, 2017
74. Jovanovic D, Krstivojevic P, Obradovic I, Durdevic V, Dukanovic L: Serum cystatin C and beta2-microglobulin as markers of glomerular filtration rate. *Ren Fail* 25: 123–133, 2003
75. Niewczas MA, Pavkov ME, Skupien J, Smiles A, Md Dom ZI, Wilson JM, et al: A signature of circulating inflammatory proteins and development of end-stage renal disease in diabetes. *Nat Med* 25: 805–813, 2019
76. Schenk S, Schoenhals GJ, de Souza G, Mann M: A high confidence, manually validated human blood plasma protein reference set. *BMC Med Genomics* 1: 41, 2008
77. Jia L, Zhang L, Shao C, Song E, Sun W, Li M, et al: An attempt to understand kidney's protein handling function by comparing plasma and urine proteomes. *PLoS One* 4: e5146, 2009
78. Lehallier B, Gate D, Schaum N, Nanasi T, Lee SE, Yousef H, et al: Undulating changes in human plasma proteome profiles across the lifespan. *Nat Med* 25: 1843–1850, 2019
79. Suhre K, McCarthy MI, Schwenk JM: Genetics meets proteomics: Perspectives for large population-based studies. *Nat Rev Genet* 22: 19–37, 2021
80. Zoccali C, Vanholder R, Massy ZA, Ortiz A, Sarafidis P, Dekker FW, et al: The systemic nature of CKD. *Nat Rev Nephrol* 13: 344–358, 2017
81. Prunotto M, Farina A, Lane L, Pernin A, Schifferli J, Hochstrasser DF, et al: Proteomic analysis of podocyte exosome-enriched fraction from normal human urine. *J Proteomics* 82: 193–229, 2013
82. Clark CJ, Sage EH: A prototypic matricellular protein in the tumor microenvironment: Where there's SPARC, there's fire. *J Cell Biochem* 104: 721–732, 2008
83. Feng D, Ngov C, Henley N, Boufaied N, Gerarduzzi C: Characterization of matricellular protein expression signatures in mechanistically diverse mouse models of kidney injury. *Sci Rep* 9: 16736, 2019
84. Marimuthu A, O'Meally RN, Chaerkady R, Subbannayya Y, Nanjappa V, Kumar P, et al: A comprehensive map of the human urinary proteome. *J Proteome Res* 10: 2734–2743, 2011
85. Francki A, Sage EH: SPARC and the kidney glomerulus: matricellular proteins exhibit diverse functions under normal and pathological conditions. *Trends Cardiovasc Med* 11: 32–37, 2001
86. Woroniecka KI, Park AS, Mohtat D, Thomas DB, Pullman JM, Susztak K: Transcriptome analysis of human diabetic kidney disease. *Diabetes* 60: 2354–2369, 2011
87. Lindenmeyer MT, Eichinger F, Sen K, Anders HJ, Edenhofer I, Mattinzoli D, et al: Systematic analysis of a novel human renal glomerulus-enriched gene expression dataset. *PLoS One* 5: e11545, 2010
88. Nystrom J, Fierlbeck W, Granqvist A, Kulak SC, Ballermann BJ: A human glomerular SAGE transcriptome database. *BMC Nephrol* 10: 13, 2009
89. Ju W, Greene CS, Eichinger F, Nair V, Hodgin JB, Bitzer M, et al: Defining cell-type specificity at the transcriptional level in human disease. *Genome Res* 23: 1862–1873, 2013
90. Wynn TA, Ramalingam TR: Mechanisms of fibrosis: Therapeutic translation for fibrotic disease. *Nat Med* 18: 1028–1040, 2012
91. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G: Best (but oft-forgotten) practices: The design, analysis, and interpretation of Mendelian randomization studies. *Am J Clin Nutr* 103: 965–978, 2016
92. Davies NM, Holmes MV, Davey Smith G: Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ* 362: k601, 2018

AFFILIATIONS

[1]Research Unit Molecular Epidemiology, German Research Center for Environmental Health, Neuherberg, Germany
[2]Institute of Epidemiology, German Research Center for Environmental Health, Neuherberg, Germany
[3]TUM School of Medicine, Technical University of Munich, Munich, Germany
[4]German Center for Cardiovascular Research, Munich, Germany
[5]Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom
[6]Department of Physiology and Biophysics, Weill Cornell Medicine-Qatar, Doha, Qatar
[7]Division of Cardiovascular Sciences, University of Manchester, Manchester, United Kingdom
[8]School of Health and Life Sciences, Federation University Australia, Ballarat, Australia
[9]Department of Cardiovascular Sciences, University of Leicester, Leicester, United Kingdom
[10]Department of Physiology, University of Melbourne, Melbourne, Australia
[11]Department of Cellular Pathology, University Hospitals of Leicester National Health Service Trust, Leicester, United Kingdom
[12]Institute for Clinical Diabetology, Leibniz Center for Diabetes Research at Heinrich Heine University Düsseldorf, Düsseldorf, Germany
[13]German Center for Diabetes Research, München-Neuherberg, Germany
[14]Department of Data-Driven Medicine, Institute of Genetic Epidemiology, Faculty of Medicine and Medical Center-University of Freiburg, Freiburg, Germany
[15]Department of Medicine, Katz Family Division of Nephrology and Hypertension, University of Miami Miller School of Medicine, Miami, Florida
[16]Department of Clinical Epidemiology, Leiden University Medical Centre, Leiden, The Netherlands
[17]Institute of Neurogenomics, German Research Center for Environmental Health, Neuherberg, Germany
[18]Department of Neurogenetics and Institute of Human Genetics, Technical University of Munich, Munich, Germany
[19]Department of Public Health and Primary Care, University of Cambridge, Cambridge, United Kingdom
[20]British Heart Foundation Centre of Research Excellence, University of Cambridge, Cambridge, United Kingdom
[21]National Institute for Health Research Blood and Transplant Research Unit in Donor Health and Genomics, University of Cambridge, Cambridge, United Kingdom
[22]National Institute for Health Research Cambridge Biomedical Research Centre, University of Cambridge and Cambridge University Hospitals, Cambridge, United Kingdom
[23]Health Data Research UK Cambridge, Wellcome Genome Campus and University of Cambridge, Cambridge, United Kingdom
[24]Department of Human Genetics, Wellcome Sanger Institute, Hinxton, United Kingdom
[25]Department of Haematology, University of Cambridge, Cambridge, United Kingdom
[26]National Health Service Blood and Transplant, Cambridge Biomedical Campus, Long Road, Cambridge, United Kingdom
[27]Wellcome Sanger Institute, Hinxton, United Kingdom
[28]National Health Service Blood and Transplant Oxford Centre, Oxford, United Kingdom
[29]Radcliffe Department of Medicine, University of Oxford, Oxford, United Kingdom
[30]Research Unit Protein Science and Metabolomics and Proteomics Core Facility, Helmholtz Zentrum Munich - German Research Center for Environmental Health, Neuherberg, Germany
[31]Scientific Service Group Biomolecular Mass Spectrometry, Max Planck Institute for Heart and Lung Research, Bad Nauheim, Germany
[32]German Centre for Cardiovascular Research (DZHK), Partner Site Rhine-Main, Max Planck Institute of Heart and Lung Research, Bad Nauheim, Germany
[33]Klinik für Herz-Kreislauferkrankungen, Deutsches Herzzentrum München, Technical University of Munich, Munich, Germany
[34]Institute of Epidemiology and Medical Biometry, University of Ulm, Ulm, Germany
[35]Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway
[36]Nord-Trøndelag Health Study HUNT Research Centre, Faculty of Medicine, Norwegian University of Science and Technology, Levanger, Norway
[37]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland
[38]School of Pharmaceutical Sciences, University of Geneva, Geneva, Switzerland
[39]Department of Endocrinology and Diabetology, Medical Faculty and University Hospital Düsseldorf, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany
[40]Manchester Heart Centre and Manchester Academic Health Science Centre, Manchester University NHS Foundation Trust, Manchester, United Kingdom
[41]Department SHIP/Clinical-Epidemiological Research, Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany
[42]German Center for Cardiovascular Research, partner site Greifswald, Greifswald, Germany

Appendix B

Publication: DNAm-based signatures of accelerated aging and mortality in blood are associated with low renal function

Matías-García *et al. Clin Epigenet* (2021) 13:121
https://doi.org/10.1186/s13148-021-01082-w

Clinical Epigenetics

RESEARCH Open Access

DNAm-based signatures of accelerated aging and mortality in blood are associated with low renal function

Pamela R. Matías-García[1,2,3,4*], Cavin K. Ward-Caviness[5], Laura M. Raffield[6], Xu Gao[7], Yan Zhang[8,9], Rory Wilson[2,3], Xīn Gào[8], Jana Nano[3,10], Andrew Bostom[11], Elena Colicino[12], Adolfo Correa[13], Brent Coull[14], Charles Eaton[11,15], Lifang Hou[16], Allan C. Just[12], Sonja Kunze[2,3], Leslie Lange[17], Ethan Lange[17], Xihong Lin[18], Simin Liu[19], Jamaji C. Nwanaji-Enwerem[20], Alex Reiner[21], Jincheng Shen[22], Ben Schöttker[8,23], Pantel Vokonas[18], Yinan Zheng[16], Bessie Young[24,25], Joel Schwartz[20], Steve Horvath[26], Ake Lu[26], Eric A. Whitsel[27,28], Wolfgang Koenig[4,29,30], Jerzy Adamski[31,32,33], Juliane Winkelmann[34,35,36,37], Hermann Brenner[8,23], Andrea A. Baccarelli[7], Christian Gieger[2,3], Annette Peters[3,4,10], Nora Franceschini[27†] and Melanie Waldenberger[2,3,4*†]

Abstract

Background: The difference between an individual's chronological and DNA methylation predicted age (DNAmAge), termed DNAmAge acceleration (DNAmAA), can capture life-long environmental exposures and age-related physiological changes reflected in methylation status. Several studies have linked DNAmAA to morbidity and mortality, yet its relationship with kidney function has not been assessed. We evaluated the associations between seven DNAm aging and lifespan predictors (as well as GrimAge components) and five kidney traits (estimated glomerular filtration rate [eGFR], urine albumin-to-creatinine ratio [uACR], serum urate, microalbuminuria and chronic kidney disease [CKD]) in up to 9688 European, African American and Hispanic/Latino individuals from seven population-based studies.

Results: We identified 23 significant associations in our large trans-ethnic meta-analysis ($p < 1.43E{-}03$ and consistent direction of effect across studies). Age acceleration measured by the Extrinsic and PhenoAge estimators, as well as Zhang's 10-CpG epigenetic mortality risk score (MRS), were associated with all parameters of poor kidney health (lower eGFR, prevalent CKD, higher uACR, microalbuminuria and higher serum urate). Six of these associations were independently observed in European and African American populations. MRS in particular was consistently associated with eGFR ($\beta = -0.12$, 95% CI = [−0.16, −0.08] change in log-transformed eGFR per unit increase in MRS, $p = 4.39E{-}08$), prevalent CKD (odds ratio (OR) = 1.78 [1.47, 2.16], p = 2.71E-09) and higher serum urate levels ($\beta = 0.12$

*Correspondence: pamela.matias@helmholtz-muenchen.de; waldenberger@helmholtz-muenchen.de
†Nora Franceschini and Melanie Waldenberger shared senior authorship
[1] TUM School of Medicine, Technical University of Munich, Munich, Germany
[2] Research Unit Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich/Neuherberg, Germany
Full list of author information is available at the end of the article

[0.07, 0.16], $p = 2.08E{-}06$). The "first-generation" clocks (Hannum, Horvath) and GrimAge showed different patterns of association with the kidney traits. Three of the DNAm-estimated components of GrimAge, namely adrenomedullin, plasminogen-activation inhibition 1 and pack years, were positively associated with higher uACR, serum urate and microalbuminuria.

Conclusion: DNAmAge acceleration and DNAm mortality predictors estimated in whole blood were associated with multiple kidney traits, including eGFR and CKD, in this multi-ethnic study. Epigenetic biomarkers which reflect the systemic effects of age-related mechanisms such as immunosenescence, inflammaging and oxidative stress may have important mechanistic or prognostic roles in kidney disease. Our study highlights new findings linking kidney disease to biological aging, and opportunities warranting future investigation into DNA methylation biomarkers for prognostic or risk stratification in kidney disease.

Keywords: Aging, Kidney function, Epigenetic age acceleration, DNAm age, Glomerular filtration rate, UACR, Serum urate

Background

The kidneys are responsible for maintenance of homeostasis and blood filtration, and their function is most commonly clinically assessed by measuring serum creatinine levels to estimate glomerular filtration rate (eGFR) [1, 2]. Chronic kidney disease (CKD), defined by low eGFR (< 60 ml/min/m^2) and/or presence of protein in urine, is an increasingly prevalent non-communicable disease with a considerable burden worldwide [3–5]. Increased urinary albumin-to-creatinine ratio (uACR) is a marker of kidney injury, measured to identify early kidney damage which can precede eGFR decline (for example, diabetic nephropathy) [6]. Albuminuria is a predictor of CKD progression and mortality [6, 7], whereas high serum levels of urate, a molecule of purine nucleotide metabolism excreted by the kidney, is a risk factor for incident cardiovascular and kidney disease, and also a biomarker of low eGFR [8].

DNA methylation (DNAm), defined as the covalent addition of a methyl group to a DNA nucleotide (usually the cytosine of a cytosine-guanine dinucleotide [CpG]), is the most extensively studied epigenetic mechanism, and its role in numerous conditions and diseases has been demonstrated [9]. Age-predicting algorithms based on the percentages of DNAm observed at sets of CpGs, such as the ones proposed by Hannum [10] and Horvath [11], have been used to predict an individual's age (DNAmAge) and assess biological aging by calculating the difference between an individual's predicted and chronological age—a concept known as DNAmAge acceleration (DNAmAA) [12, 13]. Other measures have been derived to assess specific aspects of aging mechanisms, such as intrinsic epigenetic age acceleration (IEAA), which assesses aging independent of blood immune system changes [14], or extrinsic epigenetic age acceleration (EEAA) [15, 16], which specifically estimates aging as related to the immune system and reflected in changes in blood immune cell-type proportions. A "second generation" of DNAm-based aging signatures incorporated physiological markers to better capture changes in traditional biological aging biomarkers [12]. PhenoAge was developed as a marker meant to mirror physiological dysregulation as reflected in changes in age and 9 additional age-related features, such as C-reactive protein and serum glucose [17]. GrimAge is a mortality predictor based on mortality-related DNAm-estimated traits [18]. Another DNAm-based lifespan predictor, the 10-CpG epigenetic mortality risk score (MRS), stands out for its simplicity and its recent validation [19, 20].

Multiple studies have shown, although with varying findings, a positive relationship between DNAmAge measured in blood and aging-related diseases and mortality [12, 13, 21]. DNAmAge is associated with all-cause mortality [22, 23], frailty [24], cognitive function and physical fitness [25], body mass index (BMI) [26] and obesity [27], lifetime stress [28] and a number of other age-related conditions [12]. The available evidence points to DNAmAge as a potential global biomarker of biological aging and health, though potential for publication bias must be considered [21]. Although some of the "second-generation" DNAm-based aging measures include proteins or markers known to be associated with kidney function [17, 18], whether these and other DNAm-based predictors are correlated with different parameters of kidney aging and low function has not been investigated [13, 29, 30].

We evaluated the association between five kidney traits (eGFR, prevalent CKD, uACR, microalbuminuria and serum urate) and seven DNAm-based age and/or lifespan predictors (HannumAA, HorvathAA, EEAA, IEAA, PhenoAA, GrimAA and MRS) in up to seven population-based studies in a large trans-ethnic meta-analysis. We also evaluated kidney trait associations with secondary DNAm-based predictors: categorical epigenetic mortality risk score (MRS) variables, and eight DNAm-estimated traits underlying the GrimAge mortality predictor.

We additionally performed ethnic-specific meta-analyses to identify robust associations across cohorts of different ethnicities.

Results

Study characteristics

We conducted a trans-ethnic meta-analysis of seven DNAm-based age/lifespan predictors and five kidney traits using data from seven population-based cohorts (Fig. 1, population characteristics in Table 1). Serum creatinine-based traits (eGFR and prevalent CKD) were available for all European ancestry ($k=5$), African American ($k=2$) and one additional Hispanic/Latino study ($N=9688$). Sample sizes for the other traits were smaller: serum urate was available in four studies with European ancestry and one study with African American participants ($N=5903$), while uACR and microalbuminuria were available in three of the European ancestry studies and one African American study ($N=4110$). Additional details on the cohorts are provided in Additional file 1: Table S1.

DNAm-based age and lifespan predictors were used as independent variables and kidney traits as dependent variables in covariate-adjusted regression models. Study-level results showed associations were slightly attenuated after the inclusion of additional covariates in model 2 (namely BMI, log-transformed triglycerides, HDL, hypertension, smoking status and diabetes) in comparison with model 1 (basic model adjusting for chronological age and sex; Additional file 1: Table S2). Sensitivity analyses showed that "crude" bivariate correlations were largely attenuated after adjustment for chronological age; the introduction of additional variables did not significantly alter the observed effects, despite a slight increase in the coefficient after the introduction of smoking in models with MRS and GrimAge (Additional file 3: Note

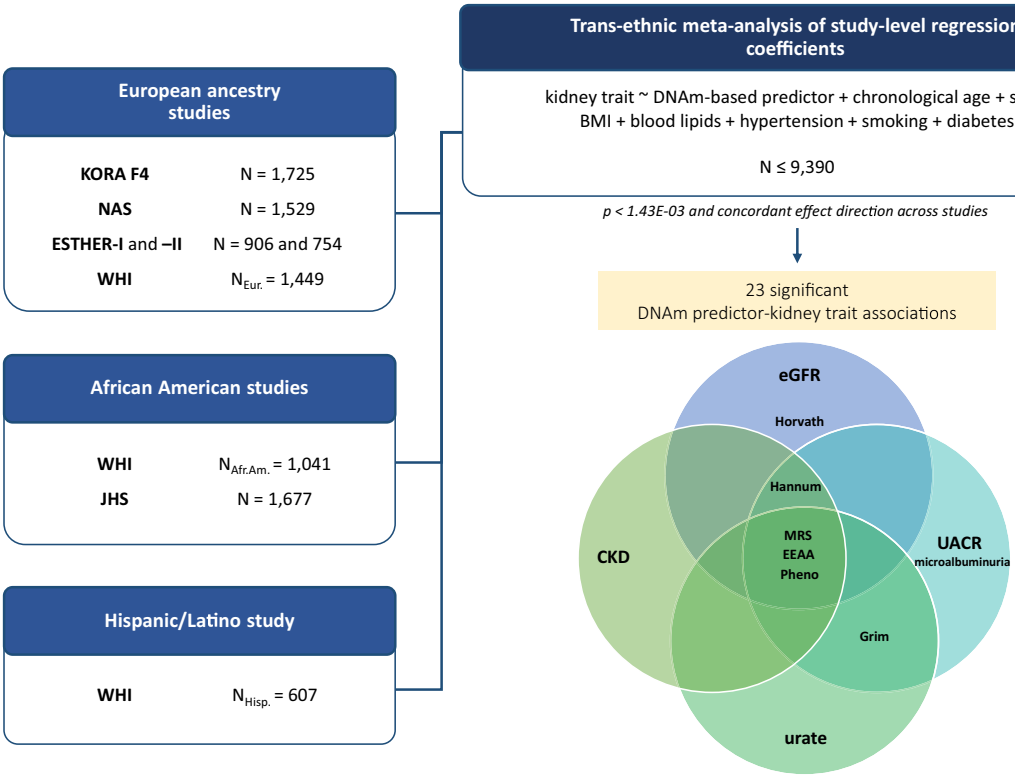

Fig. 1 Study design. Regression analyses were conducted in each study by modeling DNAm-based predictors as independent variables and kidney traits as dependent variables, adjusting for confounders. Results from the fully adjusted model (with chronological age, sex, BMI, log-transformed triglycerides, HDL, hypertension, smoking status, diabetes as covariates and baseline eGFR for serum urate analyses) were meta-analyzed using inverse-variance weighted fixed-effects and random-effects models. We based our main interpretations on the fixed-effects results; if heterogeneity was large ($I^2>0.50$ and Cochran's Q $p_{het}<0.05$), we based our interpretations on the random-effects results. The Venn diagram shows the set of 23 statistically significant associations between DNAm-based predictors and kidney traits identified in the trans-ethnic meta-analysis ($p<1.43E-03$ and consistent direction of effect across studies)

Table 1 Population characteristics

Traits	KORA	ESTHER-I	ESTHER-II	NAS	WHI			JHS
Ethnic background	**Eur**	**Eur**	**Eur**	**Eur**	**Afr. Am**	**His**	**Eur**	**Afr. Am**
N	1725	906	754	1529	1041	607	1449	1677
Age	60.98 (8.88)	62.00 (6.53)	62.76 (6.75)	74.61 (7.05)	61.83 (6.67)	60.93 (6.64)	66.32 (6.73)	56.23 (12.31)
Male	843 (48.9)	435 (48.0)	316 (41.9)	1529 (100)	0 (0)	0 (0)	0 (0)	649 (37.17)
BMI	28.11 (4.78)	27.75 (4.25)	27.47 (4.76)	28.00 (4.13)	31.61 (6.43)	29.21 (5.22)	28.84 (5.83)	32.02 (7.37)
Smoking status								
Never smoker	720 (41.7)	443 (48.9)	354 (47.0)	478 (31.3)	496 (48.06)	376 (62.46)	766 (53.34)	1490 (85.88)
Ever smoker	1003 (58.2)	463 (51.1)	400 (53.0)	1051 (68.7)	536 (51.94)	226 (37.54)	670 (46.66)	245 (14.12)
Serum creatinine	0.91 (0.27)	0.69 (0.31)	0.85 (0.31)	1.11 (0.45)	0.82 (0.21)	0.73 (0.23)	0.74 (0.14)	0.96 (0.59)
eGFR	86.77 (16.02)	99.77 (21.15)	86.63 (18.94)	69.32 (16.14)	92.31 (19.65)	88.64 (15.39)	83.69 (13.33)	93.57 (22.37)
CKD	99 (5.7)	54 (5.9)	72 (9.5)	407 (26.6)	59 (5.67)	31 (5.11)	82 (5.66)	115 (6.59)
uACR **	6.15 (3.85, 11.97)	9.14 (5.49, 17.98)	8.92 (5.34, 16.53)	NA	NA	NA	NA	5.95 (3.95,13.23)
Microalbuminuria	150 (8.7)	125 (13.8)	104 (13.8)	NA	NA	NA	NA	106 (13.75)
Serum urate	5.37 (1.46)	4.22 (1.49)	4.85 (1.47)	6.13 (1.51)	NA	NA	NA	5.64 (1.70)
Diabetes	158 (9.2)	141 (15.56)	150 (19.89)	235 (15.4)	166 (15.95)	71 (11.70)	96 (6.64)	433 (24.81)
Hypertension	788 (45.7)	516 (55.95)	446 (59.15)	1130 (73.9)	561 (56.21)	207 (35.94)	471 (35.33)	1027 (58.82)
HDL cholesterol	56.47 (14.64)	51.70 (15.79)	53.21 (15.45)	48.81 (12.85)	54.76 (13.78)	51.05 (13.02)	51.38 (11.92)	51.35 (14.73)
Triglycerides **	110 (77, 158)	89.70 (58.30, 140.50)	111.85 (76.50, 116.30)	114 (83, 158)	104 (75, 141)	140 (105, 188)	133 (95, 184)	92 (64, 129)
C-reactive protein **	1.27 (0.63, 2.655)	1.62 (0.83, 3.46)	2.40 (1.06, 4.72)	1.47 (0.75, 3.04)	NA	NA	NA	2.71 (1.18, 5.96)

Population characteristics of all participating studies. The means and standards deviation (SD) are shown for continuous traits, and *N* (%) for categorical traits. **Skewed variables, for which median and (1st, 3rd quartile) are shown. The sample sizes presented here for each of the studies correspond to the number of observations with information on DNAm-predictors, serum-based creatinine kidney traits, chronological age and sex. Age was measured in years at time of participation in study; BMI in kg/m^2; serum creatinine in mg/dL; eGFR, serum-creatinine-based estimated glomerular filtration rate in mL/min/1.73 m^2; CKD: prevalent chronic kidney disease, defined as eGFR < 60 ml/min/1.73 m^2; uACR in mg/g; microalbuminuria was defined as uACR ≥ 30 mg/g; serum urate in mg/dl; prevalent diabetes was defined based on use of glucose lowering drugs or fasting plasma glucose ≥ 126 mg/dl; hypertension defined using the Joint National Committee (JNC) VII definition (blood pressure > 140/90 mm Hg or use of anti-hypertensive medications); HDL cholesterol and triglycerides in mg/dl; C-reactive protein in mg/L; NA denotes the trait was not available. Abbreviations used in ethnic background row: Eur., European ancestry; Afr.Am.., African American; His., Hispanic/Latino. An extended version of this table is shown in Additional file 1: Table S1

S1). Estimates from the fully adjusted model were meta-analyzed using inverse-variance weighted fixed-effects and random-effects models. We based our main interpretations on the fixed-effects results; if heterogeneity was large ($I^2 > 0.50$ and Cochran's Q $p_{het} < 0.05$), we based our interpretations on the random-effects results (Methods).

Meta-analysis of associations between DNAm-based predictors and kidney traits

We identified 23 significant DNAm-based predictor-kidney trait associations ($p < 1.43E{-}03$ and concordant direction of effect across studies; Table 2). Three interesting groups worth further discussion are: (1) PhenoAA, MRS and EEAA were associated with all parameters of poor kidney health (lower eGFR, prevalent CKD, higher uACR or microalbuminuria and higher serum urate); (2) the "first-generation" epigenetic aging markers, where HannumAA was associated with all kidney traits but serum urate, and HorvathAA was only associated with lower eGFR; and (3) an analogous measure to age acceleration in GrimAge was associated with uACR, microalbuminuria and serum urate (Fig. 1). Six associations between DNAm-based predictors and kidney traits were replicated across ethnic groups, and ethnic-specific replication was observed for 16 associations in total.

PhenoAA, EEAA and MRS universally associated with poor kidney health

From these three DNAm-based predictors, associations with MRS had the smallest p-values: one MRS unit increase had a -0.12 (95% CI = [-0.16, -0.08]) change in one standard deviation (SD) of log-transformed eGFR ($p = 4.39E{-}08$) and was associated with 78% [47–116%] increased odds of prevalent CKD ($p = 2.71E{-}09$) (Table 2). Although high heterogeneity was identified in associations between MRS and CKD, uACR and

Table 2 Trans-ethnic meta-analyses of associations between kidney traits and DNAm-based age and lifespan predictors

Clock	eGFR					CKD				
	β	95% CI	*p*	I^2	p_{het}	OR	95% CI	*p*	I^2	p_{het}
HorvathAA	−0.006	−0.01, −0.003	**5.15E−04**	38.696	0.121	1.019	1.003, 1.034	0.016	13.663	0.323
HannumAA	−0.007	−0.011, −0.004	**1.05E−04**	0	0.816	1.033	1.016, 1.05	**8.55E−05**	44.898	0.08
GrimAA	−0.006	−0.01, −0.002	1.94E−03	63.474	0.008	1.027	1.01, 1.044	1.27E−03	77.714	5.2E−05
PhenoAA	−0.005	−0.008, −0.002	**2.62E−04**	0	0.564	1.031	1.018, 1.044	**3.19E−06**	33.885	0.158
EEAA	−0.008	−0.012, −0.005	**2.09E−06**	43.328	0.09	1.038	1.022, 1.055	**3.41E−06**	49.636	0.053
MRS	−0.117	−0.158, −0.075	**4.39E−08**	27.67	0.208	1.784	1.474, 2.159	**2.71E−09**	68.295	0.002[a]
IEAA	−0.004	−0.008, 0	0.051	16.08	0.303	1.007	0.99, 1.025	0.427	2.365	0.411
Clock	**uACR**					**Microalbuminuria**				
	β	95% CI	*p*	I^2	p_{het}	OR	95% CI	*p*	I^2	p_{het}
HorvathAA	0.002	−0.004, 0.008	0.606	0	0.632	1.014	0.992, 1.036	0.223	0	0.703
HannumAA	0.014	0.009, 0.02	**2.04E−06**	0	0.967	1.054	1.032, 1.076	**1.08E−06**	0	0.898
GrimAA	0.029	0.021, 0.037	**1.07E−12**	7.22	0.357	1.106	1.074, 1.138	**7.58E−12**	0	0.59
PhenoAA	0.01	0.005, 0.015	**2.71E−05**	0	0.787	1.035	1.017, 1.053	**8.96E−05**	0	0.612
EEAA	0.013	0.008, 0.017	**4.62E−07**	0	0.989	1.048	1.03, 1.066	**1.42E−07**	0	0.86
MRS	0.252	0.179, 0.324	**1.01E−11**	76.128	0.006[a]	2.238	1.734, 2.889	**6.14E−10**	67.72	0.026[a]
IEAA	0.002	−0.004, 0.008	0.529	0	0.847	1.015	0.991, 1.04	0.214	0	0.769

Clock	Urate				
	β	95% CI	*p*	I^2	p_{het}
HorvathAA	0.003	−0.001, 0.007	0.12	0	0.453
HannumAA	0.005	0.001, 0.009	0.011	0	0.919
GrimAA	0.009	0.004, 0.013	**1.17E−04**	56.31	0.057
PhenoAA	0.009	0.006, 0.012	**4.71E−08**	0	0.432
EEAA	0.007	0.004, 0.011	**4.37E−05**	41.60	0.144
MRS	0.115	0.067, 0.162	**2.08E−06**	12.16	0.336
IEAA	−0.001	−0.005, 0.003	0.675	0	0.797

Results from trans-ethnic meta-analyses of associations between kidney traits and DNAm-based age and lifespan predictors in up to seven population-based studies. Study-level associations were adjusted for chronological age, sex, BMI, blood lipids, hypertension, smoking and diabetes. Beta coefficients are given as changes in one standard deviation (SD) of the continuous kidney trait. Fully adjusted associations of serum-creatinine-based traits (eGFR, CKD) are based on $N \leq 9390$ observations, whereas the sample size for urinary albumin-based traits (uACR, microalbuminuria) is $N \leq 4406$ and for urate $N \leq 5769$. I^2 is the heterogeneity statistic, and (*Q*) p_{het} corresponds to Cochran's *Q* heterogeneity statistic

Shown in bold are statistically significant associations ($p < 1.43E\text{-}03$ and consistent direction of effect across studies) with either no evidence of heterogeneity in the fixed-effects model or supporting findings from the random-effects model

[a] p_{het} of MRS with CKD, uACR and microalbuminuria < 0.05, therefore reported association based on significant random-effects models: MRS-uACR: $\beta = 0.248$ [0.1, 0.397], $p = 1.061E{-}03$; MRS-CKD: OR = 1.915 [1.316, 2.786], $p = 6.89E{-}04$; and MRS-microalbuminuria: OR = 2.197 [1.403, 3.439], $p = 5.82E{-}04$ (Additional file 1: Table S3)

microalbuminuria, the results from random-effects models (Additional file 1: Table S3) were consistent with the estimates produced by the fixed-effect model: for example, one MRS unit increase was associated with a 0.25 [0.10, 0.40] change in one SD of log-transformed UACR ($p = 1.06E{-}03$). Figure 2 shows the study-level regression coefficients of the association between the continuous kidney traits and DNAm-based predictors standardized to one SD deviation in both terms to allow for their comparison. While the strength of association with serum urate of these three DNAm-based predictors was similar, MRS and EEAA had the largest effects on eGFR (Fig. 2, standardized effects in Additional file 1: Table S4). Similar observations were done for the binary kidney traits (Additional file 2: Fig. S1).

Figure 3 shows results from ethnic-specific meta-analyses significant in both European ancestry and African American meta-analyses ($p < 1.43E{-}03$, Additional file 1: Table S3). MRS stands out for its replication across all subgroups, including the small Hispanic/Latino cohort (Fig. 3a, study-level results in Additional file 1: Table S2). The eGFR-EEAA and CKD-PhenoAA effects were also replicated at the Bonferroni-corrected significance level across ethnic-specific meta-analyses (Fig. 3b), whereas

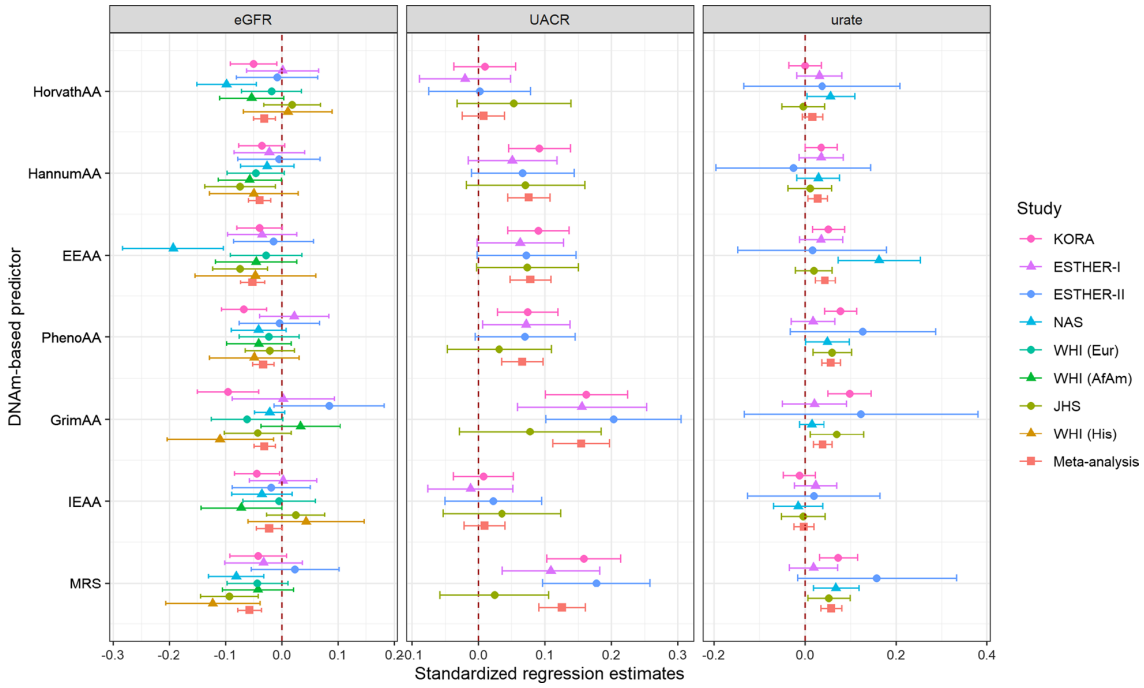

Fig. 2 Standardized effect estimates from DNAm-based predictors and kidney traits. Scatter plot showing the effect estimates from the DNAmAge and lifespan predictors across continuous kidney traits for individual studies and trans-ethnic fixed-effects meta-analysis. Effect estimates have been standardized to one SD in both variables to allow for comparison of effect sizes. The legend shows the combination of shape and color coding assigned for the studies

their "complementary" associations (CKD-EEAA and eGFR-PhenoAA) were nominally significant (Additional file 2: Fig. S2A). Further effects observed in both European ancestry meta-analysis and the African American cohort were the associations of higher serum urate with MRS and PhenoAA (Fig. 3c). On the other hand, the associations between EEAA, PhenoAA and MRS with higher uACR (and prevalent microalbuminuria) identified in the trans-ethnic meta-analysis were mostly driven by the effects from the European ancestry cohorts (Additional file 2: Figs. S3 and S4), as well as that of serum urate and EEAA (Additional file 2: Fig. S5A). Most notably, the association between uACR and MRS was replicated at the Bonferroni-corrected level in two of the European ancestry cohorts (KORA and ESTHER-II), and nominally significant in the third one (Additional file 1: Table S2).

Associations with "first-generation" DNAm clocks and GrimAge

HorvathAA, age acceleration measured by the "first-generation" DNAm-based predictor HorvathAge, was exclusively associated with low eGFR: a one year difference between DNAm-estimated age and chronological age was associated with a -0.006 $[-0.01, -0.003]$ change in log-transformed eGFR ($p = 5.15E{-}04$, Table 2). HannumAA, another "first-generation" DNAm-based age predictor, was also associated with eGFR; the strength of the association with both "first-generation" DNAm-based predictors was similar (Fig. 2). HannumAA was additionally associated with CKD, uACR and microalbuminuria (Additional file 2: Figs. S2–S4B). While the effects observed for HannumAA with eGFR and CKD were robustly replicated by cohorts with African American participants (Additional file 2: Fig. S2B), the association between HorvathAA and eGFR was mostly driven by studies with European ancestry (Additional file 2: Fig. S2C).

A year of GrimAge acceleration was associated with an increase of 0.03 SD of log-transformed uACR and a 10.6% increase in the odds of having microalbuminuria, both early markers of renal damage (Additional file 2: Figs S3–S4C). Likewise, a one-year difference in GrimAge was associated with higher serum urate levels, a risk factor for cardiorenal disease (Additional file 2: Fig. S5B). The GrimAA effects on uACR were the largest across DNAm-based predictors and

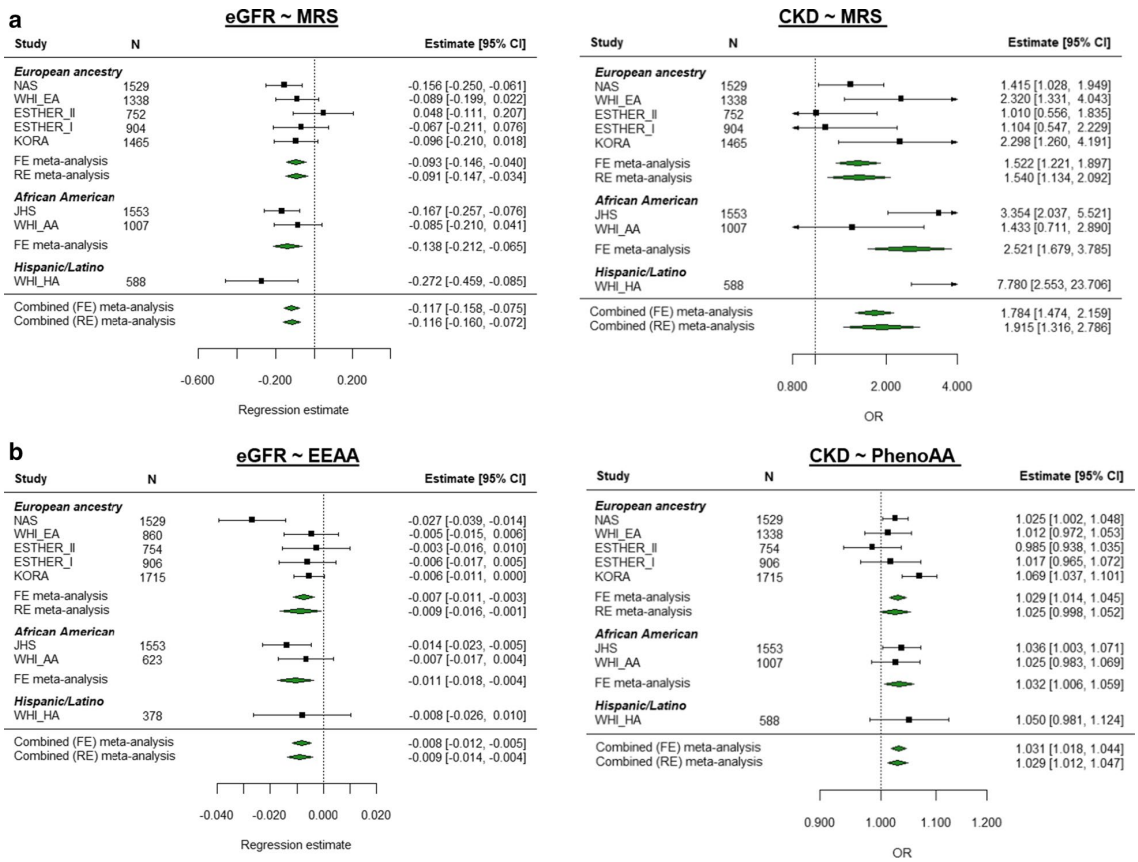

Fig. 3 Multivariate regression meta-analysis of association between kidney traits and MRS, EEAA and PhenoAA. Multivariate regression models were used to assess the relationship between kidney traits and age acceleration as measured by seven DNAm-based predictors of age and/or lifespan, including Zhang's 10-CpG mortality risk score (MRS), PhenoAge and extrinsic epigenetic age acceleration (EEAA). Results from the fully adjusted model included chronological age, sex, BMI, log-transformed triglycerides, HDL, hypertension, smoking status and diabetes as covariates, and baseline eGFR for serum urate analyses. Inverse-variance weighted fixed-effects models meta-analysis was conducted, where if heterogeneity was observed ([*Q*] $p_{het} < 0.05$), a random-effects model was further interpreted. Individual panels show forest plots with the study-level and meta-analytic results for each association between kidney trait and DNAm-based predictor from the fully adjusted model. The rows correspond to the different studies and the sample size for each analysis (*N*). For eGFR and urate, the regression estimates represent the change in one standard deviation of the kidney trait per unit change in the MRS or per one year of age acceleration for PhenoAA and EEA. For CKD, the estimate column corresponds to the odds ratio (OR). The *x*-axis shows the estimates obtained from either the regression model (for single studies, data point shape is a black square) or the meta-analytic estimate (data point shape is a green diamond for serum creatinine-based traits and a red diamond for urate) with their 95% CI. Estimate: regression coefficient for the continuous traits and OR for CKD, 95% CI: 95% confidence interval of the estimate

kidney traits (Fig. 2), and a similar effect was observed for microalbuminuria (Additional file 2: Fig. S1).

Associations with categorical MRS and GrimAge components

In the categorical MRS, where risk groups were defined based on the number of CpGs with methylation levels beyond pre-defined high-risk thresholds, a 0.431 (95% CI = [0.29, 0.57]) SD increase in log-uACR was observed for individuals with > 5 "aberrantly" methylated CpGs (high risk MRS) compared to those with 0–1 aberrantly methylated loci (low risk MRS) ($p = 4.01E\text{-}09$, Additional file 2: Fig. S6A). Similar associations were observed for microalbuminuria (OR = 2.075 [1.582, 2.722], $p = 1.36E\text{-}07$; Additional file 1: Table S3). The moderate MRS category (defined as those with 2–5 "aberrantly" methylated loci) was also associated with serum urate (Additional file 2: Fig. S6B). These effects were mostly driven by European ancestry studies, where effects had a similar strength of association and replicated within this group

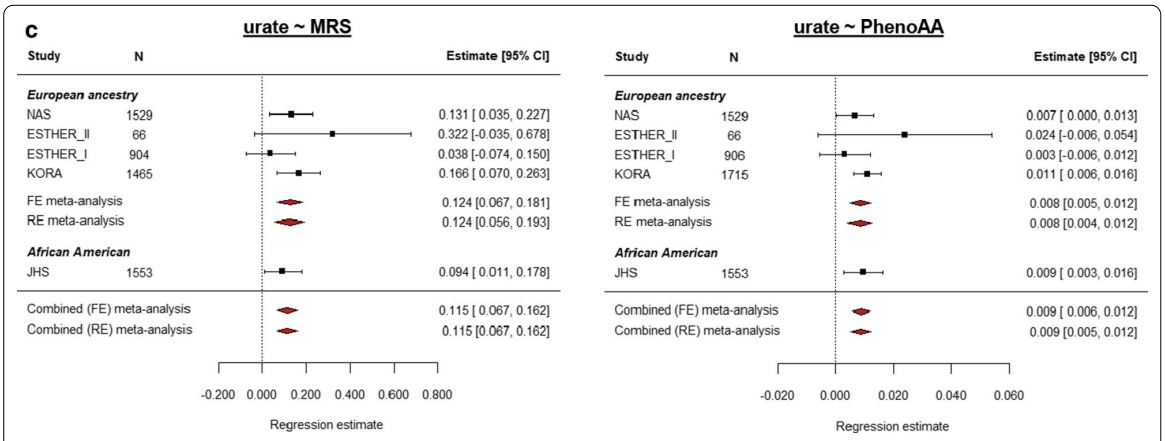

Fig. 3 continued

(Additional file 2: Fig. S7). None of these effects replicated in the African American cohort, perhaps due to the smaller sample size and reduced statistical power.

In regard to the secondary analyses with the eight DNAm-based components that constitute GrimAge, we found associations for three of them. DNAm-estimated adrenomedullin (DNAmADM), plasminogen-activation inhibition 1 (DNAmPAI1) and pack years (DNAmPACKYRS) were positively associated with higher uACR, higher serum urate levels and microalbuminuria (Additional file 1: Table S5). Of note, the effects identified for serum urate replicated in both the European-ancestry analyses and in the African American cohort (Additional file 1: Table S2), whereas those in uACR and microalbuminuria were mostly replicated within the European studies.

Discussion

We identified 23 associations between kidney traits and DNAm-based predictors of aging and/or mortality in a large trans-ethnic meta-analysis from up to seven multiethnic population-based cohorts. PhenoAA, MRS and EEAA were associated with all parameters of poor kidney health (lower eGFR, prevalent CKD, higher uACR or microalbuminuria and higher serum urate). Distinct patterns of association were observed with age acceleration in the "first-generation" clocks (Hannum and Horvath) and an analogous measure in GrimAge: HorvathAA was only associated with lower eGFR, while HannumAA was associated with all kidney traits but serum urate. Finally, GrimAA was associated with uACR, microalbuminuria and serum urate.

Sensitivity analyses showed that the inclusion of chronological age in a "crude" model lead to large attenuation of the correlation between DNAm-based predictors and kidney traits, as expected when adjusting for variables confounding the exposure-outcome association. The posterior introduction of additional variables did not further alter the observed effects. A slight increase in the coefficient after the introduction of smoking was observed for MRS and GrimAge, which may be explained in general by the strong correlation between smoking and DNAm/DNAmAge acceleration in blood [31–34], and specifically in relation to the these two markers as they either directly incorporate cigarette smoking into its formulation [19] or capture the effects of cigarette smoking [20].

From the three "universally" associated DNAm-based predictors, MRS was robustly associated with multiple kidney phenotypes, some of which replicated across ethnic-specific analyses (i.e., eGFR, prevalent CKD and urate). Together with EEAA and GrimAA, MRS showed the largest effects in comparison with the other DNAm-based predictors. Moreover, MRS moderate and high risk categories [20] were also associated with eGFR and markers of kidney injury in the secondary analyses. Recent evidence suggests the MRS is a DNAm-based biomarker which reflects mortality risks by capturing the effects of oxidative stress and systemic inflammation, as well as inflammation-driven changes in immune cell counts—all mechanisms shared by numerous chronic diseases [35–38]. Particularly in kidney aging, mitochondrial dysfunction, uremia-induced epigenetic changes and the production of reactive oxygen species in the glomeruli that lead to barrier function impairment and albuminuria are all mechanisms promoting oxidative stress [8, 39–41]. These factors may thus explain the associations between kidney traits and MRS here identified. The grounding of the MRS

in inflammation and oxidative stress mechanisms may also explain the predictive power of this predictor in regard to outcomes from cancer to cardiovascular disease mortality [20, 21].

EEAA and PhenoAA, also associated with all kidney traits, are both extrinsic aging measures (i.e., tracking changes in blood cell composition) that also capture (intrinsic) aging-related physiological dysregulation [12, 18]. EEAA, identified as a measure of immune system aging, is considered better at predicting age-related decline of tissue performance [12, 14, 18]. Increased allostatic load, activation of stress and pro-aging pathways, impairment of protective pathways as well as exogenous lifestyle and environmental factors are all factors driving premature aging [17, 39]. Immunosenescence [42, 43], systemic low-grade inflammation ('inflammaging') [44] and oxidative stress [37, 39] contribute to an increased allostatic load and are mechanisms present in kidney aging [17, 45]. In particular, the immune system plays an important role in the incidence, development and resolution of renal disease [43]. A signature of 447 genes involved in renal aging further confirmed the relevance of cellular pathways common to immune function and renal physiological decline [46]. Likewise, CKD patients show signs of premature immunological aging (such as poor naïve T-cell frequency and reduced thymic output), a status induced by ureamia (high concentrations of serum urate in blood) that is associated with poor clinical outcomes [47]. At the cellular level, telomere attrition [39, 48] and the cytokine secretory profile of senescent cells promote inflammation and lead to fibrotic damage [37, 39, 43, 45], further linking inflammaging to renal dysfunction [9]. Our findings, in line with the positive association between PhenoAge and albumin excretion rate identified in 499 subjects with type 1 diabetes [49], may be thus explained by the relationship between these DNAm-based predictors and the aging-related changes to the immune system, low-grade chronic inflammation and oxidative stress that impact renal disease [37, 47]. Moreover, EEAA, PhenoAge and the MRS have shown a stronger predictive association with time to death than HannumAA, HorvathAA and IEAA, suggesting they better reflect mortality risks associated with biological aging [16, 20]—and based on our findings, also better reflect immune system changes associated with kidney aging. The correlation between these DNAm-based predictors and poor kidney health may also explain their strong associations with mortality, as the ensuing contribution of renal disease to physiological dysregulation may increase mortality risk [6, 7, 39, 50].

Distinct patterns of association were identified with two of the "first-generation" clocks, HorvathAA and HannumAA. HorvathAA, thought to track cell-intrinsic aging (e.g., epigenetic stability mechanisms, cell growth and survival as well as organismal development) [11, 12, 16, 18], was exclusively associated with eGFR in our study. The statistical power derived from the larger sample size in our study likely explains this positive association, unlike prior studies reporting null findings with eGFR [49, 51]. Considering tissue from individuals with renal disease was included in the development of this pan-tissue marker [11], this epigenetic marker reflects alterations in ubiquitous cell-intrinsic pathways that our findings suggest may be relevant for renal (glomerular) function. Genome-wide association studies of the Hannum and Horvath DNAm-based predictors have shown that, even though they capture different aspects of aging, both markers are influenced by genes associated with metabolic and immune system pathways [52, 53]. Consequently, these DNAm-based predictors may also somewhat be reflective of the immune molecular mechanisms previously described. An additional factor of potential relevance for the eGFR-HorvathAA association may be aberrant glucocorticoid signaling, given its potentially pathogenic role in renal function [54] and the enrichment of glucocorticoid response elements in this DNAm marker [29].

HannumAA, associated with lower renal function and markers of early renal damage (uACR and microalbuminuria) in our study, is also strongly correlated with blood cell counts [16, 23] and is sensitive to environmental influences [15, 53]. Moreover, it has also been associated with higher levels of inflammatory biomarkers, creatinine and certain lipid classes in individuals of European ancestry [14, 15, 55]. Like other extrinsic measures, Hannum seems to be a better marker for later-life diseases and mortality than Horvath [22] and has even been proposed as a prognostic marker of pathological metabolic processes [56]. HannumAA in cancerous kidney tissue vs normal samples has also been reported [10], thus offering further evidence on the relevance of our findings in blood to renal disease.

The association between one-year difference in the GrimAge predictor and higher uACR, serum urate and microalbuminuria—but not eGFR or prevalent CKD—suggests this DNAm-based predictor might be more sensitive to systemic inflammation and early renal damage. Albuminuria changes can occur before eGFR decline in early kidney disease [6]: early structural glomerular lesions in patients with normal eGFR are better correlated with changes in uACR than with GFR decline, where the latter might not be present yet [57]. Moreover, albuminuria is a predictor of CKD progression and mortality independently from eGFR changes, which suggests they represent two independent mechanisms underlying renal disease progression [6, 7]. Our findings are in line

with prior reports of GrimAge association with albumin excretion in T1D patients and non-diabetic subjects [19, 49], where lower power may explain the null findings reported in [51]. Asymptomatic hyperuricaemia, or high serum urate levels, is involved in pro-inflammatory mechanisms and is associated with a high risk of cardiovascular and renal disease [8]. Serum urate, both in its crystal and soluble forms, activates innate immunity and triggers DNAm epigenetic mechanisms (e.g., promoting cytokine secretion, pro-inflammatory pathways including oxidative stress) leading to persisting inflammation and an increased allostatic load [8, 39]. These effects may, in turn, explain our findings in relation to GrimAge. A high degree of heterogeneity in the eGFR- and CKD-GrimAA analyses was observed, similarly to reports in prior studies [58, 59].

Overall, the observed associations did not show a clear pattern across kidney traits, thus supporting the proposed notion that the existing DNAm-based predictors might reflect different aspects of biological aging. This is in line with their differential association with risk factors, intermediate phenotypes and diseases [12, 58, 60, 61], and their inclusion of non-overlapping CpGs sets [19, 20, 60, 62]. The CpG overlap between DNAm-based predictors was assessed in detail by Liu et.al, where the lack of CpG overlap may be explained by the redundancy of the methylome: CpGs selected in the construction of different DNAm predictors may represent different aging hallmarks or pathways, despite the potential biological similarities of their genomic regions [62]. Our findings are also consistent with the many associations observed with EEAA and other blood immune system correlated DNAm-based predictors [12], and with the lack of associations with IEAA observed in prior studies [52, 61].

DNAm-estimated adrenomedullin (DNAmADM), plasminogen activator inhibitor-1 (DNAmPA1) and smoking pack years (DNAmPackYears) were positively associated with decreased renal function. Consistent with our findings, patients with chronic cardiorenal diseases have higher blood levels of ADM [63, 64] and PAI-1 [65–67], where the first is a potential biomarker of CKD progression [64, 68] and the latter a risk factor for cardiorenal disease [66, 69]. Several factors involved in kidney disease pathogenesis (e.g., oxidative stress, inflammation) induce PAI-1 expression (66, 70, 71), which in turn has been linked to fibrosis, glomeruli damage and other pathogenic mechanisms in renal disease [65, 67], as well as to thrombosis and an increased hypercoagulable state—a shared phenotype of inflammaging [37] and renal disease [72]. Moreover, the effects of smoking in DNAm [31, 32] and their association to DNAmAge acceleration in blood are well known [33, 34]. PhenoAge and MRS capture effects of cigarette smoking [18, 20, 33, 58, 60], whereas GrimAge specifically incorporates cigarette smoking into its formulation [19]. Cigarette smoking has been associated with renal function decline and increased inflammation [73], where several mechanisms explaining the negative effects of smoking on renal function (e.g., oxidative stress, endothelial dysfunction, immune function modulation) contribute to renal disease progression [74]. Overall, our findings are in line with the roles described in the literature for the studied DNAm-estimated markers, and further support the notion that associations between epigenetic aging and health outcomes may be mediated by age-related pro-inflammatory mechanisms [75]. Moreover, they suggest DNAm-based estimates might prove to be valuable proxies in settings where such variables are not available (e.g., limitations in the clinical use of ADM [68] and self-reported smoking [34]).

Strengths of this work are the large sample size and inclusion of multiple independent studies involving multi-ethnic populations. We comprehensively addressed biological aging and lifespan as predicted by DNAm and assessed multiple kidney traits reflecting different aspects of renal health. Our results in regard to PhenoAge and GrimAge represent confirmatory findings to some extent, given that renal function variables were included in the derivation of these algorithms [18, 19]. Of note, although the MRS was derived using data of two of the cohorts included in this study, it has been independently validated [20, 21] and the replication of its associations across multiple cohorts suggest our findings are not a product of data overfitting. All in all, our study meets the considerations proposed by a recent literature review and meta-analysis on the topic [22].

Limitations of this study include the estimation of DNAm markers in blood samples rather than renal tissue, although there is currently no epigenetic age predictor derived in kidney tissue. Age-related methylation changes can be tissue-specific [76] and show inter-individual variation [46], yet associations between eGFR and DNAm in blood have been demonstrated to be relevant to kidney traits [46, 77]. Moreover, this remains the only viable approach for research conducted in population-based studies, where taking renal biopsies from participants is not done due to practical and ethical considerations. Despite the bias inherent to the calculation of eGFR using equations that systematically produce higher values for individuals identified as black [2, 78, 79], associations with trans-ethnic replication in our study featured lower eGFR (CKD). Nevertheless, future kidney research would benefit from the development and use of methods relying on filtration markers independent from muscle mass and/or moving beyond race as a variable [78]. The lack of trans-ethnic replication

of all associations may be explained by multiple factors, most notably the smaller sample sizes from non-European studies, or ethnic biases in DNAm-based predictors (as those reported for PhenoAge in [80]). Future studies using larger, homogeneous sample sizes from diverse ethnic groups are needed to address the generalizability of our findings, and to interrogate the contributions of environmental and social determinants of health disparities in epigenetic aging. Our models assumed a linear relationship in the age range studied here, and residual confounding after adjustment for the covariates included in our regression analyses is a possibility. Another potential limitation is that the cross-sectional nature of the study does not allow to draw conclusions on temporal relationships between DNAm and renal phenotypes, although DNAm patterns reflect lifetime environmental exposures and genetic factors. Our findings do not provide a mechanistic or causal explanation for renal aging and blood epigenetic aging markers, but should be considered hypothesis-generating research. Preliminary results from the largest genome-wide association (GWAS) study of DNAm-based aging and lifespan predictors found no evidence of causal effects on renal outcomes (uACR, eGFR, albuminuria and serum urate, among 150 studied traits) [81]. Nevertheless, DNAm-based aging/lifespan signatures could still be a valuable biomarker of kidney disease prognosis, risk stratification or kidney-related outcomes. Future research should aim to expand our understanding of epigenetic aging in chronic diseases and on the clinical utility of the DNAm-based predictors.

Conclusion

In this study of multi-ethnic population-based cohorts, kidney traits were robustly associated with DNAm-based aging and lifespan predictors measured in whole blood, as well as with some secondary DNAm-estimated markers. Our findings are consistent with a body of literature on the role immunosenescence, inflammaging and oxidative stress play in renal function and damage, as well as offer evidence on the relevance of cell-intrinsic aging mechanisms. DNAm age and lifespan predictors seem to capture the contribution of multiple CpGs to pathological changes common to systemic inflammation and renal disease, highlighting the systemic nature of age-related physiological functional decline. Future research in longitudinal studies is required to evaluate the translational value of our findings as either prognostic biomarkers for disease progression and mortality, or as means to enhance risk stratification; functional studies to explore the complex physiological interplay between epigenetic mechanisms and biological aging are also warranted.

Methods

Study design

The association between kidney traits as dependent variables and DNAm aging/lifespan predictors as independent variables was modeled using linear regression following a meta-analytic approach (Fig. 1). Study-level results from up to seven studies were included in the meta-analyses: four studies with participants of European ancestry, one study of African American participants and three substudies from the WHI with European American, African American and Hispanic/Latino participants. The studies were KORA (Kooperative Gesundheitsforschung in der Region Augsburg), NAS (Normative Aging Study), ESTHER (Epidemiologische Studie zu Chancen der Verhütung, Früherkennung und optimierten THerapie chronischer ERkrankungen in der älteren Bevölkerung), WHI (Women's Health Initiative) and the Jackson Heart Study (JSH). The ESTHER, NAS and WHI studies contributed multiple sets of data that were analyzed separately: data sets from ESTHER corresponded to two surveys with non-overlapping sets of participants, whereas data from NAS were longitudinal and collected over consecutive examinations and analyzed taking into account these repeated measures. Further details on the data collection and methods used in each study are available in Additional file 3.

Outcome definition

Serum creatinine values obtained with a Jaffé assay before 2009 were calibrated by multiplying by 0.95 [82] and used to calculate estimated glomerular filtration rate (eGFR) as per the CKD-EPI equation [2] in its implementation in the R package *nephro* [83]. Prevalent chronic kidney disease (CKD) was defined as eGFR < 60 ml/min/1.73 m^2 [84]. eGFR and urinary albumin-creatinine ratio (UACR) were log transformed prior to statistical analysis. Microalbuminuria was defined as uACR ≥ 30 mg/g. Serum urate was also studied.

DNAmAge assessment

Methylation was measured using the Illumina Infinium HumanMethylation450K or EPIC array in whole blood and used to estimate measures of DNAmAge and mortality (additional details on each predictor are described in Additional file 3: Note S2). Five DNAmAge and lifespan predictors were calculated using the online DNAm Age calculator (https://dnamage.genetics.ucla.edu/) [11]: Hannum's estimate (HannumAge), ExtrinsicAge (EEAA) [10], Horvath's estimate (HorvathAge) [11], PhenoAge [18] and GrimAge [19]. IEAA, a marker capturing cell-intrinsic aging properties that are independent of blood cell types, was derived by regressing HorvathAge on cell counts [10]. Quality control was conducted as in previous

meta-analyses of epigenetic measures [16], with exclusion of individuals with mismatching predicted and reported sex data. Age acceleration (AA) measures were calculated in each study as the difference between the predicted DNAmAge and chronological age, with chronological age included in all models as an adjustment for known chronological age effects across the lifespan. Defining age acceleration as the difference, rather than the residual of chronological age regressed on epigenetic age, has advantages as it is an individual measure as opposed to a population measure, and is not defined to have mean 0 in each population as is the case for the residual measure. A sixth measure, the 10-CpG-based epigenetic mortality risk score (MRS) in its continuous form was calculated as the sum of the methylation β values multiplied by the regression coefficients of each of the ten CpGs for all-cause mortality, as described in [20].

Further measures of epigenetic aging were included in the secondary analysis: the risk-level MRS variable was built based on the cumulative number of "aberrantly" methylated CpG sites, defined by the cut-offs derived from the 4th quartile of the CpG positively correlated with mortality (cg08362785) and the 1st quartile of the other nine loci defined in [20]. Participants were then assigned to one of three risk levels based on the total number of "aberrantly" methylated CpGs: low risk, MRS = 0–1; moderate risk, MRS = 2–5; and high risk, MRS > 5. Finally, to better understand GrimAge, we also included in the analysis its eight underlying traits (smoking pack-years, adrenomedullin, beta-2 microglobulin, cystatin C, growth differentiation factor 15, leptin, plasminogen activation inhibitor 1, tissue inhibitor metalloproteinase 1) [19] if a renal phenotype was associated with GrimAge.

Statistical analysis

Linear and logistic regression models were run with kidney traits as outcomes and measures of DNAm-based age acceleration and lifespan as predictors, including covariates to adjust for potential confounding by biological and technical factors. Chronological age and sex were included in a basic model, whereas additional adjustment for BMI (kg/m^2), log transformed triglycerides, HDL, hypertension, smoking status (current/ever, never) and diabetes was done in the fully adjusted model. Linear regression models for serum urate additionally adjusted for baseline eGFR. Details on the study-specific definition or exclusion of variables, as well as additional information on all of the cohorts, are given in Additional file 3: Note S3. Leukocyte count (either measured or estimated by the Houseman approach [85]) was additionally included in the regression models for Horvath's estimate as to obtain the intrinsic age acceleration measure (IEAA). In the secondary analyses, the aforementioned covariates from the basic and the fully adjusted models were used, with the exception of no smoking adjustment for DNAm-predicted pack years. All measures of association between epigenetic markers and continuous renal traits (eGFR, uACR, urate) were standardized to the standard deviation of the given renal trait as to obtain estimates comparable across renal traits.

All outcomes were available in at least one cohort of European ancestry and African American studies (Additional file 1: Table S1), although eGFR and CKD were the only outcomes reported by all participating studies. Each cohort provided regression estimates and standard errors, which were pooled using inverse-variance fixed-effects and random-effects models. Between-study heterogeneity was assessed using Cochran's Q and I^2 statistics. High heterogeneity was defined as $I^2 > 0.50$ and (Q) $p_{het} < 0.05$. If high heterogeneity was detected in the fixed-effects model, the random-effects model was interpreted. All statistical analyses were conducted using R version 3.5.3 [86], where meta-analyses were conducted using the *metafor* package v2.0 [87]. Multiple testing was addressed by correcting the significance level for the total number of statistical tests (i.e., Bonferroni correction, 0.05/7 epigenetic markers * 5 renal traits). Associations were considered significant if $p < 1.43$ E−03 in the trans-ethnic meta-analysis and if they had consistent direction of effect across studies. Ethnic-specific replication was defined as associations with consistent direction of effect reaching nominal statistical significance ($p < 0.05$) in two or more studies.

Abbreviations

AA: Age acceleration defined as the difference between chronological and DNAmAge; HorvathAA: Age acceleration calculated with Horvath's DNAmAge predictor; HannumAA: Age acceleration calculated with Hannum's DNAmAge predictor; CKD: Chronic kidney disease; CpG: Cytosine-phosphate-guanine dinucleotide; DNAm: DNA methylation; DNAmAge: Estimated "biological" age using DNA methylation information, also known in the literature as epigenetic age; EEAA: "Extrinsic epigenetic age acceleration," calculated with Hannum's DNAmAge predictor; eGFR: Estimated glomerular filtration rate; ESTHER: Epidemiologische Studie zu Chancen der Verhütung, Früherkennung und optimierten THerapie chronischer ERkrankungen in der älteren Bevölkerung; GrimAA: Measure analogous to other age acceleration markers calculated with DNAm GrimAge predictor; JHS: Jackson Heart Study; KORA: Kooperative Gesundheitsforschung in der Region Augsburg; IEAA: Intrinsic epigenetic age acceleration; MRS: Epigenetic mortality risk score; NAS: US Department of Veterans Affairs' Normative Ageing Study; PhenoAA: Age acceleration calculated with DNAm PhenoAge predictor; uACR: Urinary albumin-to-creatinine ratio; WHI: Women's Health Initiative.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13148-021-01082-w.

Additional file 1: Tables S1–S5 showing the results from all conducted analyses across renal parameters and DNAm-based predictors, providing both estimates from individual studies and meta-analyses.

Additional file 2: Figures S1–S7 showing forest plots for all significant associations not presented in the main figures, as well as secondary traits (categorical MRS), comparison of MRS and DNAmAA effect sizes across renal traits.

Additional file 3: Notes including additional information on sensitivity analyses, the derivation of the DNAm-based predictors, cohort-specific details (background and study design, as well as data collection and preprocessing), and all legends to supplemental tables and figures.

Acknowledgements
We are grateful to all study participants of ESTHER, JHS, KORA, NAS and WHI for their invaluable contributions to these studies, as well as all members of field staff conducting the studies.

Authors' contributions
PRMG, CKWC, LMR, RW, XuG, JA, JW, AP, CG, NF and MW contributed to the design of the study. PRMG, LMR, XuG, JS and YaZ conducted the study-level data processing and/or analyses. PRMG carried out meta-analyses and was the major contributor in writing the manuscript. CKWC, LMR, XuG, YaZ, RW, XiG, JN, EC, BC, LH, ACJ, SK, XL, JCNE, JiS, BS, PV, YiZ, JoS, EAW, WK, JA, JW, HB, AAB, CG, NF, MW oversaw additional input and revisions to manuscript drafts. CKWC, LMR, XuG, YaZ, RW, XiG, JA, JW, CG, NF and MW provided statistical guidance throughout the study, as well as actively took part in the interpretation of data. AB, EC, AC, BC, CE, LH, ACJ, SK, LL, EL, SL, AR, JiS, BY, JoS, SH, AL, EAW, SH made contributions in the acquisition of data in their respective studies. All authors read and approved the final manuscript.

Funding
The KORA study was initiated and financed by the Helmholtz Zentrum München–German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Furthermore, KORA research has been supported within the Munich Center of Health Sciences (MC-Health), Ludwig-Maximilians-Universität, as part of LMUinnovativ. The ESTHER study was supported by the Baden-Württemberg State Ministry of Science, Research and Arts (Stuttgart, Germany), the Federal Ministry of Education and Research (Berlin, Germany), and the Federal Ministry of Family Affairs, Senior Citizens, Women and Youth (Berlin, Germany). The sponsors had no role in the study design, in the collection, analysis, and interpretation of data and preparation, review, or approval of the manuscript. The Normative Aging Study is supported by the National Institute of Environmental Health Sciences (grants P30ES009089, R01ES021733, R01ES025225, and R01ES027747). The VA Normative Aging Study is supported by the Cooperative Studies Program/Epidemiology Research and Information Center of the U.S. Department of Veterans Affairs and is a component of the Massachusetts Veterans Epidemiology Research and Information Center, Boston, Massachusetts. The WHI program is funded by the National Heart, Lung and Blood Institute, National Institutes of Health, U.S. Department of Health and Human Services through contracts HHSN268201600018C, HHSN268201600001C, HHSN268201600002C, HHSN268201600003C and HHSN268201600004C. The authors thank the WHI investigators and staff for their dedication, and the study participants for making the program possible. A full listing of WHI investigators can be found at: http://www.whi.org/researchers/Documents%20%20Write%20a%20Paper/WHI%20Investigator%20Long%20List.pdf. This study was supported by the National Institutes of Health R01- MD012765, R01-DK117445 and R21-HL140385 to NF; by NIH/NIEHS R01-ES020836 to LH, AB and EAW; and by NIH/NHLBI contract 60442456 BAA23 to SH. The Jackson Heart Study (JHS) is supported and conducted in collaboration with Jackson State University (HHSN268201800013I), Tougaloo College (HHSN268201800014I), the Mississippi State Department of Health (HHSN268201800015I) and the University of Mississippi Medical Center (HHSN268201800010I, HHSN268201800011I and HHSN268201800012I) contracts from the National Heart, Lung and Blood Institute (NHLBI) and the National Institute on Minority Health and Health Disparities (NIMHD). The authors also wish to thank the staffs and participants of the JHS. The views expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Heart, Lung and Blood Institute; the National Institutes of Health; or the US Department of Health and Human Services. The funders had no role in the design and conduct of the study, in the collection, analysis and interpretation of the data, and in the preparation, review or approval of the manuscript. The project described was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant KL2TR002490 (LMR). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. LMR was also funded by T32 HL129982 and R21-HL140385.

Availability of data and materials
The dataset(s) supporting the conclusions of this article is(are) included within the article (and its additional file(s)). The informed consent given by the study participants does not cover posting of participant level phenotype data in public databases. Pre-existing data access policies for each of the five studies state research data requests can be submitted to each steering committee. Study-specific details regarding such requests are described in the next points: KORA data are available upon request from KORA Project Application Self-Service Tool (https://epi.helmholtz-muenchen.de/); data requests can be submitted online and are subject to approval by the KORA Board. ESTHER data are not allowed to be publicly available due to restrictions of informed consent. However, the use of the data for collaboration projects has been and will remain the approach for data sharing. Data used in this analysis were produced and used in accordance with the policies of the Jackson Heart Study under contracts from the National, Heart, Lung and Blood Institute and are not the domain of the authors but that of the Jackson Heart Study. These data are available to other researchers for purposes of reproducing the results or replicating the procedures by submitting a manuscript proposal to the Jackson Heart Study at jhspub@umc.edu, as described at https://www.jacksonheartstudy.org/Research/Publications#submitmanuscript. Data updates for the Jackson Heart Study are also deposited regularly in the National Institutes of Health data repositories, dbGaP (https://www.ncbi.nlm.nih.gov/gap/) and BioLincc (https://biolincc.nhlbi.nih.gov/home/). NAS data are available on request due to privacy/ethical restrictions. Data that support the findings of this study are available from AAB upon reasonable request. The DNA methylation datasets for WHI are publicly available under dbGAP access number phs000200.v10.p3 or upon request to www.whi.org. WHI datasets are also available through BioLINCC https://biolincc.nhlbi.nih.gov/studies/whi_ctos/. Scripts used in data processing and statistical analyses have been made publicly accessible at https://ascgitlab.helmholtz-muenchen.de/pamela.matias/dnam-aging-in-kora-f4 and https://ascgitlab.helmholtz-muenchen.de/pamela.matias/kidney-dnamage.

Declarations

Ethics approval and consent to participate
All participants provided written consent, and study-specific details are described next: ESTHER: The study was approved by the ethics committees of the University of Heidelberg and of the Medical Association of Saarland. All participants provided written informed consent. JHS: All participants included in this analysis provided written, informed consent for use of genetic data, and all study protocols conform to the 1975 Declaration of Helsinki guidelines. The study was approved by the Institutional Review Boards of the participating institutions (University of Mississippi Medical Center, Jackson State University and Tougaloo College). KORA: The KORA cohort ethical approval was granted by the ethics committee of the Bavarian Medical Association (REC reference numbers: F4: #06068) and all were carried out in accordance with the principles of the Declaration of Helsinki. All research participants have signed informed consent prior to taking part in any research activities. The KORA data protection procedures were approved by the responsible data protection officer of the Helmholtz Zentrum München. NAS: The NAS was approved by the Department of Veterans Affairs Boston Healthcare System and written informed consent was obtained from each subject before participation. WHI: All study participants have provided written consent to participate in genetic studies.

Consent for publication
Not applicable.

Competing interests
The authors declare the following competing interests: WK reports personal fees from AstraZeneca, Novartis, Pfizer, The Medicines Company, DalCor, Kowa, Amgen, Corvidia, Daiichi-Sankyo, Berlin-Chemie, Sanofi, Genentech, and Bristol-Myers Squibb, grants and non-financial support from Singulex, Abbott, Roche Diagnostics, Beckmann, all outside the submitted work.

Author details
[1]TUM School of Medicine, Technical University of Munich, Munich, Germany. [2]Research Unit Molecular Epidemiology, Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich/Neuherberg, Germany. [3]Institute of Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich/Neuherberg, Germany. [4]German Center for Cardiovascular Research (DZHK), Partner Site Munich Heart Alliance, Munich, Germany. [5]Center for Public Health and Environmental Assessment, US Environmental Protection Agency, Chapel Hill, NC, USA. [6]Department of Genetics, University of North Carolina, Chapel Hill, NC, USA. [7]Laboratory of Precision Environmental Health, Mailman School of Public Health, Columbia University, New York, NY, USA. [8]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany. [9]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. [10]German Center for Diabetes Research (DZD), Neuherberg, Germany. [11]Center For Primary Care and Prevention, Memorial Hospital of Rhode Island, Pawtucket, RI, USA. [12]Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA. [13]Departments of Medicine and Pediatrics, University of Mississippi Medical Center, Jackson, MS, USA. [14]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [15]Department of Family Medicine, Warren Alpert Medical School, Brown University, Providence, RI, USA. [16]Department of Preventive Medicine, Feinberg School of Medicine, NorthwesternUniversity, Chicago, IL, USA. [17]Department of Medicine, University of Colorado Denver, Anschutz Medical Campus, Aurora, CO, USA. [18]Veterans Affairs Normative Aging Study, Veterans Affairs Boston Healthcare System, Department of Medicine, Boston University School of Medicine, Boston, MA, USA. [19]Department of Epidemiology, School of Public Health, Brown University, Providence, RI, USA. [20]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA. [21]Department of Epidemiology, University of Washington, Seattle, WA, USA. [22]Department of Population Health Sciences, School of Medicine, University of Utah, Salt Lake City, UT, USA. [23]Network Aging Research, University of Heidelberg, Heidelberg, Germany. [24]Nephrology, Hospital and Specialty Medicine and Center for Innovation for Veteran-Centered and Value Driven Care, Veterans Affairs Puget Sound Health Care System, Seattle, WA, USA. [25]Division of Nephrology, Kidney Research Institute, University of Washington, Seattle, WA, USA. [26]Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA. [27]Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA. [28]Department of Medicine, School of Medicine, University of North Carolina, Chapel Hill, NC, USA. [29]Deutsches Herzzentrum München, Technische Universität München, Munich, Germany. [30]Institute of Epidemiology and Medical Biometry, University of Ulm, Ulm, Germany. [31]Research Unit Molecular Endocrinology and Metabolism, Genome Analysis Center, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich/Neuherberg, Germany. [32]Chair for Experimental Genetics, Technical University of Munich, Freising-Weihenstephan, Germany. [33]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore. [34]Institute of Neurogenomics, Helmholtz Zentrum München, German Research Center for Environmental Health, Munich/Neuherberg, Germany. [35]Chair Neurogenetics, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany. [36]Institute of Human Genetics, Klinikum rechts der Isar, Technical University of Munich, Munich, Germany. [37]Munich Cluster for Systems Neurology (SyNergy), Munich, Germany.

Received: 29 November 2020 Accepted: 18 April 2021
Published online: 02 June 2021

References

1. Levey AS, Inker LA, Coresh J. GFR estimation: from physiology to public health. Am J Kidney Dis Off J Natl Kidney Found. 2014;63(5):820–34.
2. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Internal Med. 2009;150(9):604–12.
3. Eckardt KU, Coresh J, Devuyst O, Johnson RJ, Kottgen A, Levey AS, et al. Evolving importance of kidney disease: from subspecialty to global health burden. Lancet (London, England). 2013;382(9887):158–69.
4. Hill NR, Fatoba ST, Oke JL, Hirst JA, O'Callaghan CA, Lasserson DS, et al. Global prevalence of chronic kidney disease: a systematic review and meta-analysis. PLoS ONE. 2016;11(7):e0158765.
5. Levey AS, Atkins R, Coresh J, Cohen EP, Collins AJ, Eckardt KU, et al. Chronic kidney disease as a global public health problem: approaches and initiatives: a position statement from Kidney Disease Improving Global Outcomes. Kidney Int. 2007;72(3):247–59.
6. 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 (London, England). 2010;375(9731):2073–81.
7. Astor BC, Matsushita K, Gansevoort RT, van der Velde M, Woodward M, Levey AS, et al. Lower estimated glomerular filtration rate and higher albuminuria are associated with mortality and end-stage renal disease. A collaborative meta-analysis of kidney disease population cohorts. Kidney Int. 2011;79(12):1331–40.
8. Joosten LAB, Crişan TO, Bjornstad P, Johnson RJ. Asymptomatic hyperuricaemia: a silent activator of the innate immune system. Nat Rev Rheumatol. 2020;16(2):75–86.
9. Salameh Y, Bejaoui Y, El Hajj N. DNA methylation biomarkers in aging and age-related diseases. Front Genetics. 2020;11:171.
10. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. Mol Cell. 2013;49(2):359–67.
11. Horvath S. DNA methylation age of human tissues and cell types. Genome Biol. 2013;14(10):R115.
12. Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genetics. 2018;19(6):371–84.
13. Dhingra R, Nwanaji-Enwerem JC, Samet M, Ward-Caviness CK. DNA methylation age—environmental influences, health impacts, and its role in environmental epidemiology. Curr Environ Health Rep. 2018;5(3):317–27.
14. Horvath S, Gurven M, Levine ME, Trumble BC, Kaplan H, Allayee H, et al. An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. Genome Biol. 2016;17(1):171.
15. Quach A, Levine ME, Tanaka T, Lu AT, Chen BH, Ferrucci L, et al. Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. Aging. 2017;9(2):419–46.
16. Chen BH, Marioni RE, Colicino E, Peters MJ, Ward-Caviness CK, Tsai PC, et al. DNA methylation-based measures of biological age: meta-analysis predicting time to death. Aging. 2016;8(9):1844–65.
17. Levine ME, Lu AT, Quach A, Chen BH, Assimes TL, Bandinelli S, et al. An epigenetic biomarker of aging for lifespan and healthspan. Aging. 2018;10(4):573–91.
18. Lu AT, Quach A, Wilson JG, Reiner AP, Aviv A, Raj K, et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. Aging. 2019;11(2):303–27.
19. Zhang Y, Wilson R, Heiss J, Breitling LP, Saum K-U, Schöttker B, et al. DNA methylation signatures in peripheral blood strongly predict all-cause mortality. Nat Commun. 2017;8:14617.
20. Gao X, Colicino E, Shen J, Just AC, Nwanaji-Enwerem JC, Wang C, et al. Comparative validation of an epigenetic mortality risk score with three aging biomarkers for predicting mortality risks among older adult males. Int J Epidemiol. 2019;48(6):1958–71.
21. Fransquet PD, Wrigglesworth J, Woods RL, Ernst ME, Ryan J. The epigenetic clock as a predictor of disease and mortality risk: a systematic review and meta-analysis. Clin Epigenetics. 2019;11(1):62.
22. Marioni RE, Shah S, McRae AF, Chen BH, Colicino E, Harris SE, et al. DNA methylation age of blood predicts all-cause mortality in later life. Genome Biol. 2015;16:25.

23. Perna L, Zhang Y, Mons U, Holleczek B, Saum KU, Brenner H. Epigenetic age acceleration predicts cancer, cardiovascular, and all-cause mortality in a German case cohort. Clin Epigenetics. 2016;8:64.
24. Breitling LP, Saum KU, Perna L, Schottker B, Holleczek B, Brenner H. Frailty is associated with the epigenetic clock but not with telomere length in a German cohort. Clin Epigenetics. 2016;8:21.
25. Marioni RE, Shah S, McRae AF, Ritchie SJ, Muniz-Terrera G, Harris SE, et al. The epigenetic clock is correlated with physical and cognitive fitness in the Lothian Birth Cohort 1936. Int J Epidemiol. 2015;44(4):1388–96.
26. Ryan J, Wrigglesworth J, Loong J, Fransquet PD, Woods RL. A systematic review and meta-analysis of environmental, lifestyle, and health factors associated with DNA methylation age. J Gerontol Ser A Biol Sci Med Sci. 2020;75(3):481–94.
27. Horvath S, Erhart W, Brosch M, Ammerpohl O, von Schönfels W, Ahrens M, et al. Obesity accelerates epigenetic aging of human liver. Proc Natl Acad Sci USA. 2014;111(43):15538–43.
28. Zannas AS, Arloth J, Carrillo-Roa T, Iurato S, Röh S, Ressler KJ, et al. Lifetime stress accelerates epigenetic aging in an urban, African American cohort: relevance of glucocorticoid signaling. Genome Biol. 2015;16:266.
29. Shiels PG, McGuinness D, Eriksson M, Kooman JP, Stenvinkel P. The role of epigenetics in renal ageing. Nat Rev Nephrol. 2017;13(8):471–82.
30. Jung M, Pfeifer GP. Aging and DNA methylation. BMC Biol. 2015;13:7.
31. Joehanes R, Just AC, Marioni RE, Pilling LC, Reynolds LM, Mandaviya PR, et al. Epigenetic signatures of cigarette smoking. Circ Cardiovas Genetics. 2016;9(5):436–47.
32. Wilson R, Wahl S, Pfeiffer L, Ward-Caviness CK, Kunze S, Kretschmer A, et al. The dynamics of smoking-related disturbed methylation: a two time-point study of methylation change in smokers, non-smokers and former smokers. BMC Genomics. 2017;18(1):805.
33. Yang Y, Gao X, Just AC, Colicino E, Wang C, Coull BA, et al. Smoking-related DNA methylation is associated with DNA methylation phenotypic age acceleration: the veterans affairs normative aging study. Int J Environ Res Public Health. 2019;16(13):2356.
34. Gao X, Zhang Y, Breitling LP, Brenner H. Relationship of tobacco smoking and smoking-related DNA methylation with epigenetic age acceleration. Oncotarget. 2016;7(30):46878–89.
35. Gao X, Gào X, Zhang Y, Holleczek B, Schöttker B, Brenner H. Oxidative stress and epigenetic mortality risk score: associations with all-cause mortality among elderly people. Eur J Epidemiol. 2019;34(5):451–62.
36. Ward-Caviness CK, Pu S, Martin CL, Galea S, Uddin M, Wildman DE, et al. Epigenetic predictors of all-cause mortality are associated with objective measures of neighborhood disadvantage in an urban population. Clin Epigenetics. 2020;12(1):44.
37. Franceschi C, Campisi J. Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. J Gerontol Ser A Biol Sci Med Sci. 2014;69(Suppl 1):S4-9.
38. Wang C, Just A, Heiss J, Coull BA, Hou L, Zheng Y, et al. Biomarkers of aging and lung function in the normative aging study. Aging. 2020;12(12):11942–66.
39. Kooman JP, Kotanko P, Schols AM, Shiels PG, Stenvinkel P. Chronic kidney disease and premature ageing. Nat Rev Nephrol. 2014;10(12):732–42.
40. Young GH, Wu VC. KLOTHO methylation is linked to uremic toxins and chronic kidney disease. Kidney Int. 2012;81(7):611–2.
41. McCarthy ET, Sharma R, Sharma M, Li JZ, Ge XL, Dileepan KN, et al. TNF-alpha increases albumin permeability of isolated rat glomeruli through the generation of superoxide. J Am Soc Nephrol JASN. 1998;9(3):433–8.
42. Mueller AL, McNamara MS, Sinclair DA. Why does COVID-19 disproportionately affect older people? Aging. 2020;12(10):9959–81.
43. Tecklenborg J, Clayton D, Siebert S, Coley SM. The role of the immune system in kidney disease. Clin Exp Immunol. 2018;192(2):142–50.
44. Franceschi C, Bonafè M, Valensin S, Olivieri F, De Luca M, Ottaviani E, et al. Inflamm-aging: an evolutionary perspective on immunosenescence. Ann N Y Acad Sci. 2000;908:244–54.
45. Rowland J, Akbarov A, Maan A, Eales J, Dormer J, Tomaszewski M. Tick-tock chimes the kidney clock—from biology of renal ageing to clinical applications. Kidney Blood Press Res. 2018;43(1):55–67.
46. Rodwell GE, Sonu R, Zahn JM, Lund J, Wilhelmy J, Wang L, et al. A transcriptional profile of aging in the human kidney. PLoS Biol. 2004;2(12):e427.
47. Crépin T, Legendre M, Carron C, Vachey C, Courivaud C, Rebibou JM, et al. Uraemia-induced immune senescence and clinical outcomes in chronic kidney disease patients. Nephrol Dial Transplant Off Publ Eur Dial Transplant Assoc Eur Renal Assoc. 2020;35(4):624–32.
48. Carrero JJ, Stenvinkel P, Fellström B, Qureshi AR, Lamb K, Heimbürger O, et al. Telomere attrition is associated with inflammation, low fetuin-A levels and high mortality in prevalent haemodialysis patients. J Internal Med. 2008;263(3):302–12.
49. Roshandel D, Chen Z, Canty AJ, Bull SB, Natarajan R, Paterson AD, et al. DNA methylation age calculators reveal association with diabetic neuropathy in type 1 diabetes. Clin Epigenetics. 2020;12(1):52.
50. O'Hare AM, Choi AI, Bertenthal D, Bacchetti P, Garg AX, Kaufman JS, et al. Age affects outcomes in chronic kidney disease. J Am Soc Nephrol JASN. 2007;18(10):2758–65.
51. Smith JA, Raisky J, Ratliff SM, Liu J, Kardia SLR, Turner ST, et al. Intrinsic and extrinsic epigenetic age acceleration are associated with hypertensive target organ damage in older African Americans. BMC Med Genomics. 2019;12(1):141.
52. Lu AT, Xue L, Salfati EL, Chen BH, Ferrucci L, Levy D, et al. GWAS of epigenetic aging rates in blood reveals a critical role for TERT. Nat Commun. 2018;9(1):387.
53. Gibson J, Russ TC, Clarke TK, Howard DM, Hillary RF, Evans KL, et al. A meta-analysis of genome-wide association studies of epigenetic age acceleration. PLoS Genet. 2019;15(11):e1008104.
54. Hunter RW, Ivy JR, Bailey MA. Glucocorticoids and renal Na(+) transport: implications for hypertension and salt sensitivity. J Physiol. 2014;592(8):1731–44.
55. Irvin MR, Aslibekyan S, Do A, Zhi D, Hidalgo B, Claas SA, et al. Metabolic and inflammatory biomarkers are associated with epigenetic aging acceleration estimates in the GOLDN study. Clin Epigenetics. 2018;10:56.
56. Morrison FG, Logue MW, Guetta R, Maniates H, Stone A, Schichman SA, et al. Investigation of bidirectional longitudinal associations between advanced epigenetic age and peripheral biomarkers of inflammation and metabolic syndrome. Aging. 2019;11(11):3487–504.
57. Looker HC, Mauer M, Saulnier PJ, Harder JL, Nair V, Boustany-Kari CM, et al. Changes in albuminuria but not GFR are associated with early changes in kidney structure in type 2 diabetes. J Am Soc Nephrol JASN. 2019;30(6):1049–59.
58. Fiorito G, McCrory C, Robinson O, Carmeli C, Rosales CO, Zhang Y, et al. Socioeconomic position, lifestyle habits and biomarkers of epigenetic aging: a multi-cohort analysis. Aging. 2019;11(7):2045–70.
59. Rezwan FI, Imboden M, Amaral AFS, Wielscher M, Jeong A, Triebner K, et al. Association of adult lung function with accelerated biological aging. Aging. 2020;12(1):518–42.
60. Zhao W, Ammous F, Ratliff S, Liu J, Yu M, Mosley TH, et al. Education and lifestyle factors are associated with dna methylation clocks in older African Americans. Int J Environ Res Public Health. 2019;16(17):3141.
61. Liu Z, Chen BH, Assimes TL, Ferrucci L, Horvath S, Levine ME. The role of epigenetic aging in education and racial/ethnic mortality disparities among older U.S. Women. Psychoneuroendocrinology. 2019;104:18–24.
62. Liu Z, Leung D, Thrush K, Zhao W, Ratliff S, Tanaka T, et al. Underlying features of epigenetic aging clocks in vivo and in vitro. Aging Cell. 2020;19(10):e13229.
63. Jougasaki M, Burnett JC Jr. Adrenomedullin as a renal regulator peptide. Nephrol Dial Transplant. 2000;15(3):293–5.
64. Kronenberg F. Emerging risk factors and markers of chronic kidney disease progression. Nat Rev Nephrol. 2009;5(12):677–89.
65. Ma L-J, Fogo AB. PAI-1 and kidney fibrosis. Front Biosci (Landmark Ed). 2009;14:2028–41.
66. Vaughan DE. PAI-1 and atherothrombosis. J Thromb Haemost JTH. 2005;3(8):1879–83.
67. Eddy AA, Fogo AB. Plasminogen activator inhibitor-1 in chronic kidney disease: evidence and mechanisms of action. J Am Soc Nephrol JASN. 2006;17(11):2999–3012.
68. Dieplinger B, Mueller T, Kollerits B, Struck J, Ritz E, von Eckardstein A, et al. Pro-A-type natriuretic peptide and pro-adrenomedullin predict progression of chronic kidney disease: the MMKD study. Kidney Int. 2009;75(4):408–14.
69. Yamamoto K, Takeshita K, Kojima T, Takamatsu J, Saito H. Aging and plasminogen activator inhibitor-1 (PAI-1) regulation: implication in the pathogenesis of thrombotic disorders in the elderly. Cardiovasc Res. 2005;66(2):276–85.

70. Eddy AA. Plasminogen activator inhibitor-1 and the kidney. Am J Physiol Renal Physiol. 2002;283(2):F209–20.
71. D'Elia JA, Bayliss G, Gleason RE, Weinrauch LA. Cardiovascular-renal complications and the possible role of plasminogen activator inhibitor: a review. Clin Kidney J. 2016;9(5):705–12.
72. Lutz J, Menke J, Sollinger D, Schinzel H, Thürmel K. Haemostasis in chronic kidney disease. Nephrol Dial Transplant. 2013;29(1):29–40.
73. Hall ME, Wang W, Okhomina V, Agarwal M, Hall JE, Dreisbach AW, et al. Cigarette smoking and chronic kidney disease in African Americans in the Jackson heart study. J Am Heart Assoc. 2016;5(6):e003280.
74. Orth SR, Hallan SI. Smoking: a risk factor for progression of chronic kidney disease and for cardiovascular morbidity and mortality in renal patients–absence of evidence or evidence of absence? Clin J Am Soc Nephrol CJASN. 2008;3(1):226–36.
75. Ward-Caviness CK, Huffman JE, Evertt K, Germain M, Van Dongen J, Hill WD, et al. DNA methylation age is associated with an altered hemostatic profile in a multi-ethnic meta-analysis. Blood. 2018;132:1842–50.
76. Slieker RC, Relton CL, Gaunt TR, Slagboom PE, Heijmans BT. Age-related DNA methylation changes are tissue-specific with ELOVL2 promoter methylation as exception. Epigenetics Chromatin. 2018;11(1):25.
77. Chu AY, Tin A, Schlosser P, Ko Y-A, Qiu C, Yao C, et al. Epigenome-wide association studies identify DNA methylation associated with kidney function. Nat Commun. 2017;8(1):1286.
78. Eneanya ND, Yang W, Reese PP. Reconsidering the consequences of using race to estimate kidney function. JAMA. 2019;322(2):113–4.
79. Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. N Engl J Med. 2020;383(9):874–82.
80. Philibert R, Beach SRH, Lei MK, Gibbons FX, Gerrard M, Simons RL, et al. Array-based epigenetic aging indices may be racially biased. Genes. 2020;11(6):685.
81. McCartney DL, Min JL, Richmond RC, Lu AT, Sobczyk MK, Davies G, et al. Genome-wide association studies identify 137 loci for DNA methylation biomarkers of ageing. bioRxiv. 2020; 2020.06.29.133702.
82. Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, et al. Decline in estimated glomerular filtration rate and subsequent risk of end-stage renal disease and mortalityrisk of end-stage renal disease and mortalityrisk of end-stage renal disease and mortality. JAMA. 2014;311(24):2518–31.
83. Pattaro C, Riegler P, Stifter G, Modenese M, Minelli C, Pramstaller PP. Estimating the glomerular filtration rate in the general population using different equations: effects on classification and association. Nephron Clin Pract. 2013;123(1–2):102–11.
84. Jha V, Garcia-Garcia G, Iseki K, Li Z, Naicker S, Plattner B, et al. Chronic kidney disease: global dimension and perspectives. Lancet (London, England). 2013;382(9888):260–72.
85. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinform. 2012;13:86.
86. R Core Team. R: a language and environment for statistical computing. 2015.
87. Viechtbauer W. Conducting meta-analyses in R with the metafor package. J Stat Softw. 2010;36(3):48.

Publisher's Note

Appendix C

Permission to reuse the quote from (McKusick & Ruddle, 1987) for its inclusion in the opening page of this dissertation was obtained on May 15th, 2021. Details and conditions on the license are included in this appendix.

The bibliography included in this document lists literature referenced in Chapter 1 to Chapter 4. Please note that some of references were included in the bibliography of the first-author publications (Matías-García et al., 2021a) and (Matías-García et al., 2021b).

Figures reproduced and/or adapted from other publications (Davey Smith & Hemani, 2014; Franks & Pomares-Millan, 2020; Miranda-Gonçalves et al., 2018) – namely Figure 1.1, Figure 1.2 and Figure 2.3 – , are contained in articles published under a Creative Commons Attribution International License (CC BY) and thus no permission is necessary to reuse article contents given a proper citation is in place. A copy of this license is available at http://creativecommons.org/licenses/by/4.0/. Emojis used in this dissertation are part of Twitter's open source emoji set, Twemoji, and are also licensed under the CC-BY 4.0. More information can be found at https://twemoji.twitter.com/.

Permissions to reuse content from (Field et al., 2018) for its inclusion in Figure 1.3 and Figure 1.5, from (Yu et al., 2020) for Figure 1.4 and from (J. G. Smith & Gerszten, 2017) for Figure 1.6 were obtained on May 17th, 2021. Permission to reuse content reproduced from (Inker et al., 2014) for its use in Figure 1.7 was obtained on May 18th, 2021. The licenses' details and conditions are appended to this dissertation.

Figure 2.1 and Figure 2.2 were created with BioRender.com.

Bibliography

Anderson, G. L., Manson, J., Wallace, R., Lund, B., Hall, D., Davis, S., Shumaker, S., Wang, C. Y., Stein, E., & Prentice, R. L. (2003). Implementation of the Women's Health Initiative study design. *Ann Epidemiol*, *13*(9 Suppl), S5–17. https://doi.org/10.1016/s1047-2797(03)00043-7

Astor, B. C., Matsushita, K., Gansevoort, R. T., van der Velde, M., Woodward, M., Levey, A. S., Jong, P. E., Coresh, J., Astor, B. C., Matsushita, K., Gansevoort, R. T., van der Velde, M., Woodward, M., Levey, A. S., de Jong, P. E., Coresh, J., El-Nahas, M., Eckardt, K. U., Kasiske, B. L., Wright, J., Appel, L., Greene, T., Levin, A., Djurdjev, O., Wheeler, D. C., Landray, M. J., Townend, J. N., Emberson, J., Clark, L. E., Macleod, A., Marks, A., Ali, T., Fluck, N., Prescott, G., Smith, D. H., Weinstein, J. R., Johnson, E. S., Thorp, M. L., Wetzels, J. F., Blankestijn, P. J., van Zuilen, A. D., Menon, V., Sarnak, M., Beck, G., Kronenberg, F., Kollerits, B., Froissart, M., Stengel, B., Metzger, M., Remuzzi, G., Ruggenenti, P., Perna, A., Heerspink, H. J., Brenner, B., de Zeeuw, D., Rossing, P., Parving, H. H., Auguste, P., Veldhuis, K., Wang, Y., Camarata, L., Thomas, B., & Manley, T. (2011). Lower estimated glomerular filtration rate and higher albuminuria are associated with mortality and end-stage renal disease. a collaborative meta-analysis of kidney disease population cohorts. *Kidney Int*, *79*(12), 1331–40. https://doi.org/10.1038/ki.2010.550

Bell, C. G., Lowe, R., Adams, P. D., Baccarelli, A. A., Beck, S., Bell, J. T., Christensen, B. C., Gladyshev, V. N., Heijmans, B. T., Horvath, S., Ideker, T., Issa, J.-P. J., Kelsey, K. T., Marioni, R. E., Reik, W., Relton, C. L., Schalkwyk, L. C., Teschendorff, A. E., Wagner, W., Zhang, K., & Rakyan, V. K. (2019). DNA methylation aging clocks: Challenges and recommendations. *Genome Biology*, *20*(1), 249. https://doi.org/10.1186/s13059-019-1824-y

Bibikova, M., Barnes, B., Tsan, C., Ho, V., Klotzle, B., Le, J. M., Delano, D., Zhang, L., Schroth, G. P., Gunderson, K. L., Fan, J. B., & Shen, R. (2011). High density DNA methylation array with single CpG site resolution. *Genomics*, *98*(4), 288–95. https://doi.org/10.1016/j.ygeno.2011.07.007

Breeze, C. E., Batorsky, A., Lee, M. K., Szeto, M. D., Xu, X., McCartney, D. L., Jiang, R., Patki, A., Kramer, H. J., Eales, J. M., Raffield, L., Lange, L., Lange, E., Durda, P., Liu, Y., Tracy, R. P., Van Den Berg, D., Evans, K. L., Kraus, W. E., Shah, S., Tiwari, H. K., Hou, L., Whitsel, E. A., Jiang, X., Charchar, F. J., Baccarelli, A. A., Rich, S. S., Morris, A. P., Irvin, M. R., Arnett, D. K., Hauser, E. R., Rotter, J. I., Correa, A., Hayward, C., Horvath, S., Marioni, R. E., Tomaszewski, M., Beck, S., Berndt, S. I., London, S. J., Mychaleckyj, J. C., & Franceschini, N. (2021). Epigenome-wide association study of kidney function identifies trans-ethnic and ethnic-specific loci. *Genome Med*, *13*(1), 74. https://doi.org/10.1186/s13073-021-00877-z

Carlsson, A. C., Ingelsson, E., Sundstrom, J., Carrero, J. J., Gustafsson, S., Feldreich, T., Stenemo, M., Larsson, A., Lind, L., & Arnlov, J. (2017). Use of proteomics to investigate kidney function decline over 5 years. *Clin J Am Soc Nephrol*, *12*(8), 1226–1235. https://doi.org/10.2215/cjn.08780816

Carpenter, M. A., Crow, R., Steffes, M., Rock, W., Heilbraun, J., Evans, G., Skelton, T., Jensen, R., & Sarpong, D. (2004). Laboratory, reading center, and coordinating center data management methods in the Jackson Heart Study. *Am J Med Sci*, *328*(3), 131–44. https://doi.org/10.1097/00000441-200409000-00001

Chen, B. H., Marioni, R. E., Colicino, E., Peters, M. J., Ward-Caviness, C. K., Tsai, P. C., Roetker, N. S., Just, A. C., Demerath, E. W., Guan, W., Bressler, J., Fornage, M., Studenski, S., Vandiver, A. R., Moore, A. Z., Tanaka, T., Kiel, D. P., Liang, L., Vokonas, P., Schwartz, J., Lunetta, K. L., Murabito, J. M., Bandinelli, S., Hernandez, D. G., Melzer, D., Nalls, M., Pilling, L. C., Price, T. R., Singleton, A. B., Gieger, C., Holle, R., Kretschmer, A., Kronenberg, F., Kunze, S., Linseisen, J., Meisinger, C., Rathmann, W., Waldenberger, M., Visscher, P. M., Shah, S., Wray, N. R., McRae, A. F., Franco, O. H.,

Hofman, A., Uitterlinden, A. G., Absher, D., Assimes, T., Levine, M. E., Lu, A. T., Tsao, P. S., Hou, L., Manson, J. E., Carty, C. L., LaCroix, A. Z., Reiner, A. P., Spector, T. D., Feinberg, A. P., Levy, D., Baccarelli, A., van Meurs, J., Bell, J. T., Peters, A., Deary, I. J., Pankow, J. S., Ferrucci, L., & Horvath, S. (2016). DNA methylation-based measures of biological age: Meta-analysis predicting time to death. *Aging (Albany NY)*, *8*(9), 1844–1865. https://doi.org/10.18632/aging.101020

Christensson, A., Ash, J. A., DeLisle, R. K., Gaspar, F. W., Ostroff, R., Grubb, A., Lindstrom, V., Bruun, L., & Williams, S. A. (2017). The impact of the glomerular filtration rate on the human plasma proteome. *Proteomics Clin Appl*. https://doi.org/10.1002/prca.201700067

Chu, A. Y., Tin, A., Schlosser, P., Ko, Y. A., Qiu, C., Yao, C., Joehanes, R., Grams, M. E., Liang, L., Gluck, C. A., Liu, C., Coresh, J., Hwang, S. J., Levy, D., Boerwinkle, E., Pankow, J. S., Yang, Q., Fornage, M., Fox, C. S., Susztak, K., & Köttgen, A. (2017). Epigenome-wide association studies identify DNA methylation associated with kidney function. *Nat Commun*, *8*(1), 1286. https://doi.org/10.1038/s41467-017-01297-7

Clark, C. J., & Sage, E. H. (2008). A prototypic matricellular protein in the tumor microenvironment–where there's SPARC, there's fire. *J Cell Biochem*, *104*(3), 721–32. https://doi.org/10.1002/jcb.21688

Davey Smith, G., & Hemani, G. (2014). Mendelian randomization: Genetic anchors for causal inference in epidemiological studies. *Hum Mol Genet*, *23*(R1), R89–98. https://doi.org/10.1093/hmg/ddu328

Dhingra, R., Nwanaji-Enwerem, J. C., Samet, M., & Ward-Caviness, C. K. (2018). DNA methylation age—environmental influences, health impacts, and its role in environmental epidemiology. *Current Environmental Health Reports*, *5*(3), 317–327. https://doi.org/10.1007/s40572-018-0203-2

Dieplinger, B., Mueller, T., Kollerits, B., Struck, J., Ritz, E., von Eckardstein, A., Haltmayer, M., & Kronenberg, F. (2009). Pro-a-type natriuretic peptide and pro-adrenomedullin predict progression of chronic kidney disease: The

MMKD study. *Kidney Int*, *75*(4), 408–14. https://doi.org/10.1038/ki.2008.560

Eckardt, K. U., Coresh, J., Devuyst, O., Johnson, R. J., Kottgen, A., Levey, A. S., & Levin, A. (2013). Evolving importance of kidney disease: From subspecialty to global health burden. *Lancet*, *382*(9887), 158–69. https://doi.org/10.1016/s0140-6736(13)60439-0

Ellington, A. A., Kullo, I. J., Bailey, K. R., & Klee, G. G. (2010). Antibody-based protein multiplex platforms: Technical and operational challenges. *Clin Chem*, *56*(2), 186–93. https://doi.org/10.1373/clinchem.2009.127514

Emilsson, V., Ilkov, M., Lamb, J. R., Finkel, N., Gudmundsson, E. F., Pitts, R., Hoover, H., Gudmundsdottir, V., Horman, S. R., Aspelund, T., Shu, L., Trifonov, V., Sigurdsson, S., Manolescu, A., Zhu, J., Olafsson, O., Jakobsdottir, J., Lesley, S. A., To, J., Zhang, J., Harris, T. B., Launer, L. J., Zhang, B., Eiriksdottir, G., Yang, X., Orth, A. P., Jennings, L. L., & Gudnason, V. (2018). Co-regulatory networks of human serum proteins link genetics to disease. *Science*, *361*(6404), 769–773. https://doi.org/10.1126/science.aaq1327

Feng, D., Ngov, C., Henley, N., Boufaied, N., & Gerarduzzi, C. (2019). Characterization of matricellular protein expression signatures in mechanistically diverse mouse models of kidney injury. *Scientific Reports*, *9*(1), 16736. https://doi.org/10.1038/s41598-019-52961-5

Field, A. E., Robertson, N. A., Wang, T., Havas, A., Ideker, T., & Adams, P. D. (2018). DNA methylation clocks in aging: Categories, causes, and consequences. *Mol Cell*, *71*(6), 882–895. https://doi.org/10.1016/j.molcel.2018.08.008

Folkersen, L., Gustafsson, S., Wang, Q., Hansen, D. H., Hedman, Å. K., Schork, A., Page, K., Zhernakova, D. V., Wu, Y., Peters, J., Eriksson, N., Bergen, S. E., Boutin, T. S., Bretherick, A. D., Enroth, S., Kalnapenkis, A., Gådin, J. R., Suur, B. E., Chen, Y., Matic, L., Gale, J. D., Lee, J., Zhang, W., Quazi, A., Ala-Korpela, M., Choi, S. H., Claringbould, A., Danesh, J., Davey Smith, G., de Masi, F., Elmståhl, S., Engström, G., Fauman, E., Fernandez, C., Franke, L., Franks, P. W., Giedraitis, V., Haley, C., Hamsten, A., Inga-

son, A., Johansson, Å., Joshi, P. K., Lind, L., Lindgren, C. M., Lubitz, S., Palmer, T., Macdonald-Dunlop, E., Magnusson, M., Melander, O., Michaelsson, K., Morris, A. P., Mägi, R., Nagle, M. W., Nilsson, P. M., Nilsson, J., Orho-Melander, M., Polasek, O., Prins, B., Pålsson, E., Qi, T., Sjögren, M., Sundström, J., Surendran, P., Võsa, U., Werge, T., Wernersson, R., Westra, H.-J., Yang, J., Zhernakova, A., Ärnlöv, J., Fu, J., Smith, J. G., Esko, T., Hayward, C., Gyllensten, U., Landen, M., Siegbahn, A., Wilson, J. F., Wallentin, L., Butterworth, A. S., Holmes, M. V., Ingelsson, E., & Mälarstig, A. (2020). Genomic and drug target evaluation of 90 cardiovascular proteins in 30,931 individuals. *Nature Metabolism*, *2*(10), 1135–1148. https://doi.org/10.1038/s42255-020-00287-2

Foulkes, A. (2009). *Applied statistical genetics with R: For population-based association studies*. Springer New York. https://books.google.de/books?id=gwW8jArHnwwC

Franceschi, C., & Campisi, J. (2014). Chronic inflammation (inflammaging) and its potential contribution to age-associated diseases. *J Gerontol A Biol Sci Med Sci*, *69 Suppl 1*, S4–9. https://doi.org/10.1093/gerona/glu057

Francki, A., & Sage, E. H. (2001). SPARC and the kidney glomerulus: Matricellular proteins exhibit diverse functions under normal and pathological conditions. *Trends Cardiovasc Med*, *11*(1), 32–7. https://doi.org/10.1016/s1050-1738(01)00081-0

Franks, P. W., & Pomares-Millan, H. (2020). Next-generation epidemiology: The role of high-resolution molecular phenotyping in diabetes research. *Diabetologia*, *63*(12), 2521–2532. https://doi.org/10.1007/s00125-020-05246-w

Fransquet, P. D., Wrigglesworth, J., Woods, R. L., Ernst, M. E., & Ryan, J. (2019). The epigenetic clock as a predictor of disease and mortality risk: A systematic review and meta-analysis. *Clinical Epigenetics*, *11*(1), 62. https://doi.org/10.1186/s13148-019-0656-7

Fulwyler, M. J., & McHugh, T. M. (1990). Flow microsphere immunoassay for the quantitative and simultaneous detection of multiple soluble analytes. *Methods Cell Biol*, *33*, 613–29.

Gailly, P., Jouret, F., Martin, D., Debaix, H., Parreira, K. S., Nishita, T., Blanchard, A., Antignac, C., Willnow, T. E., Courtoy, P. J., Scheinman, S. J., Christensen, E. I., & Devuyst, O. (2008). A novel renal carbonic anhydrase type III plays a role in proximal tubule dysfunction. *Kidney Int*, *74*(1), 52–61. https://doi.org/10.1038/sj.ki.5002794

Gao, X., Colicino, E., Shen, J., Just, A. C., Nwanaji-Enwerem, J. C., Wang, C., Coull, B., Lin, X., Vokonas, P., Zheng, Y., Hou, L., Schwartz, J., & Baccarelli, A. A. (2019). Comparative validation of an epigenetic mortality risk score with three aging biomarkers for predicting mortality risks among older adult males. *Int J Epidemiol*, *48*(6), 1958–1971. https://doi.org/10.1093/ije/dyz082

Gao, X., Gào, X., Zhang, Y., Holleczek, B., Schöttker, B., & Brenner, H. (2019). Oxidative stress and epigenetic mortality risk score: Associations with all-cause mortality among elderly people. *European Journal of Epidemiology*, *34*(5), 451–462. https://doi.org/10.1007/s10654-019-00493-7

Garcia-Closas, M., Vermeulen, R., Cox, D., Lan, Q., Caporaso, N., & Rothman, N. (2011). Population-based study designs in molecular epidemiology. In N. Rothman, P. Hainaut, P. Schulte, M. Smith, P. Boffetta, F. Perera, W. H. Organization, & I. A. f. R. o. Cancer (Eds.), *Molecular epidemiology: Principles and practices* (pp. 251–259). International Agency for Research on Cancer, World Health Organization. https://books.google.de/books?id=PaEzzgEACAAJ

Geyer, P. E., Kulak, N. A., Pichler, G., Holdt, L. M., Teupser, D., & Mann, M. (2016). Plasma proteome profiling to assess human health and disease. *Cell Syst*, *2*(3), 185–95. https://doi.org/10.1016/j.cels.2016.02.015

Gibson, J., Russ, T. C., Clarke, T. K., Howard, D. M., Hillary, R. F., Evans, K. L., Walker, R. M., Bermingham, M. L., Morris, S. W., Campbell, A., Hayward, C., Murray, A. D., Porteous, D. J., Horvath, S., Lu, A. T., McIntosh, A. M., Whalley, H. C., & Marioni, R. E. (2019). A meta-analysis of genome-wide association studies of epigenetic age acceleration. *PLoS Genet*, *15*(11), e1008104. https://doi.org/10.1371/journal.pgen.1008104

Glassock, R. J., & Winearls, C. (2008). Screening for CKD with eGFR: Doubts and dangers. *Clin J Am Soc Nephrol*, *3*(5), 1563–8. https://doi.org/10.2215/cjn.00960208

Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the global burden of disease study 2013. (2015). *Lancet*, *385*(9963), 117–71. https://doi.org/10.1016/s0140-6736(14)61682-2

Gold, L., Ayers, D., Bertino, J., Bock, C., Bock, A., Brody, E. N., Carter, J., Dalby, A. B., Eaton, B. E., Fitzwater, T., Flather, D., Forbes, A., Foreman, T., Fowler, C., Gawande, B., Goss, M., Gunn, M., Gupta, S., Halladay, D., Heil, J., Heilig, J., Hicke, B., Husar, G., Janjic, N., Jarvis, T., Jennings, S., Katilius, E., Keeney, T. R., Kim, N., Koch, T. H., Kraemer, S., Kroiss, L., Le, N., Levine, D., Lindsey, W., Lollo, B., Mayfield, W., Mehan, M., Mehler, R., Nelson, S. K., Nelson, M., Nieuwlandt, D., Nikrad, M., Ochsner, U., Ostroff, R. M., Otis, M., Parker, T., Pietrasiewicz, S., Resnicow, D. I., Rohloff, J., Sanders, G., Sattin, S., Schneider, D., Singer, B., Stanton, M., Sterkel, A., Stewart, A., Stratford, S., Vaught, J. D., Vrkljan, M., Walker, J. J., Watrobka, M., Waugh, S., Weiss, A., Wilcox, S. K., Wolfson, A., Wolk, S. K., Zhang, C., & Zichi, D. (2010). Aptamer-based multiplexed proteomic technology for biomarker discovery. *PLoS One*, *5*(12), e15004. https://doi.org/10.1371/journal.pone.0015004

Good, D. M., Zurbig, P., Argiles, A., Bauer, H. W., Behrens, G., Coon, J. J., Dakna, M., Decramer, S., Delles, C., Dominiczak, A. F., Ehrich, J. H., Eitner, F., Fliser, D., Frommberger, M., Ganser, A., Girolami, M. A., Golovko, I., Gwinner, W., Haubitz, M., Herget-Rosenthal, S., Jankowski, J., Jahn, H., Jerums, G., Julian, B. A., Kellmann, M., Kliem, V., Kolch, W., Krolewski, A. S., Luppi, M., Massy, Z., Melter, M., Neususs, C., Novak, J., Peter, K., Rossing, K., Rupprecht, H., Schanstra, J. P., Schiffer, E., Stolzenburg, J. U., Tarnow, L., Theodorescu, D., Thongboonkerd, V., Vanholder, R., Weissinger, E. M., Mischak, H., & Schmitt-Kopplin, P. (2010). Naturally occurring human urinary peptides for use in diagnosis of chronic kidney

disease. *Mol Cell Proteomics*, *9*(11), 2424–37. https://doi.org/10.1074/mcp.M110.001917

Gorski, M., Jung, B., Li, Y., Matias-Garcia, P. R., Wuttke, M., Coassin, S., Thio, C. H. L., Kleber, M. E., Winkler, T. W., Wanner, V., Chai, J. F., Chu, A. Y., Cocca, M., Feitosa, M. F., Ghasemi, S., Hoppmann, A., Horn, K., Li, M., Nutile, T., Scholz, M., Sieber, K. B., Teumer, A., Tin, A., Wang, J., Tayo, B. O., Ahluwalia, T. S., Almgren, P., Bakker, S. J. L., Banas, B., Bansal, N., Biggs, M. L., Boerwinkle, E., Bottinger, E. P., Brenner, H., Carroll, R. J., Chalmers, J., Chee, M. L., Chee, M. L., Cheng, C. Y., Coresh, J., de Borst, M. H., Degenhardt, F., Eckardt, K. U., Endlich, K., Franke, A., Freitag-Wolf, S., Gampawar, P., Gansevoort, R. T., Ghanbari, M., Gieger, C., Hamet, P., Ho, K., Hofer, E., Holleczek, B., Xian Foo, V. H., Hutri-Kähönen, N., Hwang, S. J., Ikram, M. A., Josyula, N. S., Kähönen, M., Khor, C. C., Koenig, W., Kramer, H., Krämer, B. K., Kühnel, B., Lange, L. A., Lehtimäki, T., Lieb, W., Loos, R. J. F., Lukas, M. A., Lyytikäinen, L. P., Meisinger, C., Meitinger, T., Melander, O., Milaneschi, Y., Mishra, P. P., Mononen, N., Mychaleckyj, J. C., Nadkarni, G. N., Nauck, M., Nikus, K., Ning, B., Nolte, I. M., O'Donoghue, M. L., Orho-Melander, M., Pendergrass, S. A., Penninx, B., Preuss, M. H., Psaty, B. M., Raffield, L. M., Raitakari, O. T., Rettig, R., Rheinberger, M., Rice, K. M., Rosenkranz, A. R., Rossing, P., Rotter, J. I., Sabanayagam, C., Schmidt, H., Schmidt, R., et al. (2021). Meta-analysis uncovers genome-wide significant variants for rapid kidney function decline. *Kidney Int*, *99*(4), 926–939. https://doi.org/10.1016/j.kint.2020.09.030

Greenland, S. (2005). Regression methods for epidemiologic analysis. In W. Ahrens & I. Pigeot (Eds.), *Handbook of epidemiology*. Springer. https://books.google.de/books?id=u0aIngEACAAJ

Guttmacher, A. E., & Collins, F. S. (2002). Genomic medicine — a primer. *New England Journal of Medicine*, *347*(19), 1512–1520. https://doi.org/10.1056/NEJMra012240

Hall, M. E., Wang, W., Okhomina, V., Agarwal, M., Hall, J. E., Dreisbach, A. W., Juncos, L. A., Winniford, M. D., Payne, T. J., Robertson, R. M., Bhatnagar,

A., & Young, B. A. (2016). Cigarette smoking and chronic kidney disease in African Americans in the Jackson Heart Study. *J Am Heart Assoc*, *5*(6). https://doi.org/10.1161/jaha.116.003280

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., Klotzle, B., Bibikova, M., Fan, J. B., Gao, Y., Deconde, R., Chen, M., Rajapakse, I., Friend, S., Ideker, T., & Zhang, K. (2013). Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell*, *49*(2), 359–367. https://doi.org/10.1016/j.molcel.2012.10.016

Haring, R., & Wallaschofski, H. (2012). Diving through the "-omics": The case for deep phenotyping and systems epidemiology. *Omics*, *16*(5), 231–4. https://doi.org/10.1089/omi.2011.0108

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2021). *Doing meta-analysis with r: A hands-on guide* (1st). Chapman & Hall/CRC Press. https://www.routledge.com/Doing-Meta-Analysis-with-R-A-Hands-On-Guide/Harrer-Cuijpers-Furukawa-Ebert/p/book/9780367610074

Hasin, Y., Seldin, M., & Lusis, A. (2017). Multi-omics approaches to disease. *Genome Biology*, *18*(1), 83. https://doi.org/10.1186/s13059-017-1215-1

Hattori, N., & Ushijima, T. (2011). Chapter 8 - analysis of gene-specific DNA methylation. In T. Tollefsbol (Ed.), *Handbook of epigenetics* (pp. 125–134). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-375709-8.00008-3

Haycock, P. C., Burgess, S., Wade, K. H., Bowden, J., Relton, C., & Davey Smith, G. (2016). Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. *The American Journal of Clinical Nutrition*, *103*(4), 965–978. https://doi.org/10.3945/ajcn.115.118216

Heiberger, R., & Holland, B. (2015). *Statistical analysis and data display: An intermediate course with examples in r*. Springer New York. https://books.google.de/books?id=t0NECwAAQBAJ

Hernan, M. A. (2018). The c-word: Scientific euphemisms do not improve causal inference from observational data. *Am J Public Health*, *108*(5), 616–619. https://doi.org/10.2105/ajph.2018.304337

Hill, N. R., Fatoba, S. T., Oke, J. L., Hirst, J. A., O'Callaghan, C. A., Lasserson, D. S., & Hobbs, F. D. (2016). Global prevalence of chronic kidney disease - a systematic review and meta-analysis. *PLoS One*, *11*(7), e0158765. https://doi.org/10.1371/journal.pone.0158765

Holle, R., Happich, M., Lowel, H., & Wichmann, H. E. (2005). KORA–a research platform for population based health research. *Gesundheitswesen*, *67 Suppl 1*, S19–25. https://doi.org/10.1055/s-2005-858235

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biol*, *14*(10), R115. https://doi.org/10.1186/gb-2013-14-10-r115

Horvath, S., & Raj, K. (2018). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet*, *19*(6), 371–384. https://doi.org/10.1038/s41576-018-0004-3

Houseman, E. A., Accomando, W. P., Koestler, D. C., Christensen, B. C., Marsit, C. J., Nelson, H. H., Wiencke, J. K., & Kelsey, K. T. (2012). DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics*, *13*, 86. https://doi.org/10.1186/1471-2105-13-86

Howard, B. V., Van Horn, L., Hsia, J., Manson, J. E., Stefanick, M. L., Wassertheil-Smoller, S., Kuller, L. H., LaCroix, A. Z., Langer, R. D., Lasser, N. L., Lewis, C. E., Limacher, M. C., Margolis, K. L., Mysiw, W. J., Ockene, J. K., Parker, L. M., Perri, M. G., Phillips, L., Prentice, R. L., Robbins, J., Rossouw, J. E., Sarto, G. E., Schatz, I. J., Snetselaar, L. G., Stevens, V. J., Tinker, L. F., Trevisan, M., Vitolins, M. Z., Anderson, G. L., Assaf, A. R., Bassford, T., Beresford, S. A., Black, H. R., Brunner, R. L., Brzyski, R. G., Caan, B., Chlebowski, R. T., Gass, M., Granek, I., Greenland, P., Hays, J., Heber, D., Heiss, G., Hendrix, S. L., Hubbell, F. A., Johnson, K. C., & Kotchen, J. M. (2006). Low-fat dietary pattern and risk of cardiovascular disease: The Women's Health Initiative Randomized Controlled Dietary Modification Trial. *Jama*, *295*(6), 655–66. https://doi.org/10.1001/jama.295.6.655

Inker, L. A., Astor, B. C., Fox, C. H., Isakova, T., Lash, J. P., Peralta, C. A., Kurella Tamura, M., & Feldman, H. I. (2014). KDOQI us commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD. *Am J Kidney Dis*, *63*(5), 713–35. https://doi.org/10.1053/j.ajkd.2014.01.416

Jackson, R. D., LaCroix, A. Z., Gass, M., Wallace, R. B., Robbins, J., Lewis, C. E., Bassford, T., Beresford, S. A., Black, H. R., Blanchette, P., Bonds, D. E., Brunner, R. L., Brzyski, R. G., Caan, B., Cauley, J. A., Chlebowski, R. T., Cummings, S. R., Granek, I., Hays, J., Heiss, G., Hendrix, S. L., Howard, B. V., Hsia, J., Hubbell, F. A., Johnson, K. C., Judd, H., Kotchen, J. M., Kuller, L. H., Langer, R. D., Lasser, N. L., Limacher, M. C., Ludlam, S., Manson, J. E., Margolis, K. L., McGowan, J., Ockene, J. K., O'Sullivan, M. J., Phillips, L., Prentice, R. L., Sarto, G. E., Stefanick, M. L., Van Horn, L., Wactawski-Wende, J., Whitlock, E., Anderson, G. L., Assaf, A. R., & Barad, D. (2006). Calcium plus vitamin D supplementation and the risk of fractures. *N Engl J Med*, *354*(7), 669–83. https://doi.org/10.1056/NEJMoa055218

Jha, V., Garcia-Garcia, G., Iseki, K., Li, Z., Naicker, S., Plattner, B., Saran, R., Wang, A. Y., & Yang, C. W. (2013). Chronic kidney disease: Global dimension and perspectives. *Lancet*, *382*(9888), 260–72. https://doi.org/10.1016/s0140-6736(13)60687-x

Jiang, X., Eales, J. M., Scannali, D., Nazgiewicz, A., Prestes, P., Maier, M., Denniff, M., Xu, X., Saluja, S., Cano-Gamez, E., Wystrychowski, W., Szulinska, M., Antczak, A., Byars, S., Skrypnik, D., Glyda, M., Król, R., Zywiec, J., Zukowska-Szczechowska, E., Burrell, L. M., Woolf, A. S., Greenstein, A., Bogdanski, P., Keavney, B., Morris, A. P., Heagerty, A., Williams, B., Harrap, S. B., Trynka, G., Samani, N. J., Guzik, T. J., Charchar, F. J., & Tomaszewski, M. (2020). Hypertension and renin-angiotensin system blockers are not associated with expression of angiotensin-converting enzyme 2 (ACE2) in the kidney. *European Heart Journal*. https://doi.org/10.1093/eurheartj/ehaa794

Joosten, L. A. B., Crişan, T. O., Bjornstad, P., & Johnson, R. J. (2020). Asymptomatic hyperuricaemia: A silent activator of the innate immune system. *Nature Reviews Rheumatology*, *16*(2), 75–86. https://doi.org/10.1038/s41584-019-0334-3

Jovanovic, D., Krstivojevic, P., Obradovic, I., Durdevic, V., & Dukanovic, L. (2003). Serum cystatin c and beta2-microglobulin as markers of glomerular filtration rate. *Ren Fail*, *25*(1), 123–33.

Ju, W., Greene, C. S., Eichinger, F., Nair, V., Hodgin, J. B., Bitzer, M., Lee, Y. S., Zhu, Q., Kehata, M., Li, M., Jiang, S., Rastaldi, M. P., Cohen, C. D., Troyanskaya, O. G., & Kretzler, M. (2013). Defining cell-type specificity at the transcriptional level in human disease. *Genome Res*, *23*(11), 1862–73. https://doi.org/10.1101/gr.155697.113

Jung, M., & Pfeifer, G. P. (2015). Aging and DNA methylation. *BMC Biol*, *13*, 7. https://doi.org/10.1186/s12915-015-0118-4

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

Khoury, M., Millikan, R., & Gwinn, M. (2008). Genetic and molecular epidemiology. In K. Rothman, S. Greenland, & T. Lash (Eds.), *Modern epidemiology* (pp. 564–579). Wolters Kluwer Health/Lippincott Williams & Wilkins. https://books.google.de/books?id=Z3vjT9ALxHUC

Kooman, J. P., Kotanko, P., Schols, A. M., Shiels, P. G., & Stenvinkel, P. (2014). Chronic kidney disease and premature ageing. *Nat Rev Nephrol*, *10*(12), 732–42. https://doi.org/10.1038/nrneph.2014.185

Krokstad, S., Langhammer, A., Hveem, K., Holmen, T., Midthjell, K., Stene, T., Bratberg, G., Heggland, J., & Holmen, J. (2012). Cohort profile: The HUNT study, norway. *International Journal of Epidemiology*, *42*(4), 968–977. https://doi.org/10.1093/ije/dys095

Kronenberg, F. (2009). Emerging risk factors and markers of chronic kidney disease progression. *Nat Rev Nephrol*, *5*(12), 677–89. https://doi.org/10.1038/nrneph.2009.173

Kulis, M., & Esteller, M. (2010). DNA methylation and cancer. *Adv Genet*, *70*, 27–56. https://doi.org/10.1016/b978-0-12-380866-0.60002-2

Lawlor, D. A., Harbord, R. M., Sterne, J. A., Timpson, N., & Davey Smith, G. (2008). Mendelian randomization: Using genes as instruments for making causal inferences in epidemiology. *Stat Med*, *27*(8), 1133–63. https://doi.org/10.1002/sim.3034

Levey, A. S., Atkins, R., Coresh, J., Cohen, E. P., Collins, A. J., Eckardt, K. U., Nahas, M. E., Jaber, B. L., Jadoul, M., Levin, A., Powe, N. R., Rossert, J., Wheeler, D. C., Lameire, N., & Eknoyan, G. (2007). Chronic kidney disease as a global public health problem: Approaches and initiatives - a position statement from Kidney Disease Improving Global Outcomes. *Kidney Int*, *72*(3), 247–59. https://doi.org/10.1038/sj.ki.5002343

Levey, A. S., Beto, J. A., Coronado, B. E., Eknoyan, G., Foley, R. N., Kasiske, B. L., Klag, M. J., Mailloux, L. U., Manske, C. L., Meyer, K. B., Parfrey, P. S., Pfeffer, M. A., Wenger, N. K., Wilson, P. W., & Wright, J., J. T. (1998). Controlling the epidemic of cardiovascular disease in chronic renal disease: What do we know? what do we need to learn? where do we go from here? National Kidney Foundation Task Force on Cardiovascular Disease. *Am J Kidney Dis*, *32*(5), 853–906. https://doi.org/10.1016/s0272-6386(98)70145-3

Levey, A. S., Stevens, L. A., Schmid, C. H., Zhang, Y. L., Castro, 3., A. F., Feldman, H. I., Kusek, J. W., Eggers, P., Van Lente, F., Greene, T., & Coresh, J. (2009). A new equation to estimate glomerular filtration rate. *Ann Intern Med*, *150*(9), 604–12.

Levin, A., Stevens, P. E., Bilous, R. W., Coresh, J., De Francisco, A. L. M., De Jong, P. E., Griffith, K. E., Hemmelgarn, B. R., Iseki, K., Lamb, E. J., Levey, A. S., Riella, M. C., Shlipak, M. G., Wang, H., White, C. T., & Winearls, C. G. (2013). Kidney disease: Improving global outcomes (KDIGO) CKD

work group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. *Kidney International Supplements*, *3*(1), 1–150. https://doi.org/10.1038/kisup.2012.73

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., Hou, L., Baccarelli, A. A., Stewart, J. D., Li, Y., Whitsel, E. A., Wilson, J. G., Reiner, A. P., Aviv, A., Lohman, K., Liu, Y., Ferrucci, L., & Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging (Albany NY)*, *10*(4), 573–591. https://doi.org/10.18632/aging.101414

Lim, U., & Song, M. A. (2018). DNA methylation as a biomarker of aging in epidemiologic studies. *Methods Mol Biol*, *1856*, 219–231. https://doi.org/10.1007/978-1-4939-8751-1_12

Lindenmeyer, M. T., Eichinger, F., Sen, K., Anders, H. J., Edenhofer, I., Mattinzoli, D., Kretzler, M., Rastaldi, M. P., & Cohen, C. D. (2010). Systematic analysis of a novel human renal glomerulus-enriched gene expression dataset. *PLoS One*, *5*(7), e11545. https://doi.org/10.1371/journal.pone.0011545

Lisa, C.-T. (2014). *Writing dissertation and grant proposals : Epidemiology, preventive medicine and biostatistics*. Chapman; Hall/CRC. http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=762282&site=ehost-live

Liu, Z., Leung, D., Thrush, K., Zhao, W., Ratliff, S., Tanaka, T., Schmitz, L. L., Smith, J. A., Ferrucci, L., & Levine, M. E. (2020). Underlying features of epigenetic aging clocks in vivo and in vitro. *Aging Cell*, *19*(10), e13229. https://doi.org/10.1111/acel.13229

Liyanage, T., Ninomiya, T., Jha, V., Neal, B., Patrice, H. M., Okpechi, I., Zhao, M.-h., Lv, J., Garg, A. X., Knight, J., Rodgers, A., Gallagher, M., Kotwal, S., Cass, A., & Perkovic, V. (2015). Worldwide access to treatment for end-stage kidney disease: A systematic review. *The Lancet*, *385*(9981), 1975–1982. https://doi.org/10.1016/S0140-6736(14)61601-9

Lu, A. T., Quach, A., Wilson, J. G., Reiner, A. P., Aviv, A., Raj, K., Hou, L., Baccarelli, A. A., Li, Y., Stewart, J. D., Whitsel, E. A., Assimes, T. L., Ferrucci, L., & Horvath, S. (2019). DNA methylation grimage strongly pre-

dicts lifespan and healthspan. *Aging (Albany NY)*, *11*(2), 303–327. https://doi.org/10.18632/aging.101684

Lu, A. T., Xue, L., Salfati, E. L., Chen, B. H., Ferrucci, L., Levy, D., Joehanes, R., Murabito, J. M., Kiel, D. P., Tsai, P. C., Yet, I., Bell, J. T., Mangino, M., Tanaka, T., McRae, A. F., Marioni, R. E., Visscher, P. M., Wray, N. R., Deary, I. J., Levine, M. E., Quach, A., Assimes, T., Tsao, P. S., Absher, D., Stewart, J. D., Li, Y., Reiner, A. P., Hou, L., Baccarelli, A. A., Whitsel, E. A., Aviv, A., Cardona, A., Day, F. R., Wareham, N. J., Perry, J. R. B., Ong, K. K., Raj, K., Lunetta, K. L., & Horvath, S. (2018). GWAS of epigenetic aging rates in blood reveals a critical role for tert. *Nat Commun*, *9*(1), 387. https://doi.org/10.1038/s41467-017-02697-5

Marimuthu, A., O'Meally, R. N., Chaerkady, R., Subbannayya, Y., Nanjappa, V., Kumar, P., Kelkar, D. S., Pinto, S. M., Sharma, R., Renuse, S., Goel, R., Christopher, R., Delanghe, B., Cole, R. N., Harsha, H. C., & Pandey, A. (2011). A comprehensive map of the human urinary proteome. *J Proteome Res*, *10*(6), 2734–43. https://doi.org/10.1021/pr2003038

Marioni, R. E., Shah, S., McRae, A. F., Chen, B. H., Colicino, E., Harris, S. E., Gibson, J., Henders, A. K., Redmond, P., Cox, S. R., Pattie, A., Corley, J., Murphy, L., Martin, N. G., Montgomery, G. W., Feinberg, A. P., Fallin, M. D., Multhaup, M. L., Jaffe, A. E., Joehanes, R., Schwartz, J., Just, A. C., Lunetta, K. L., Murabito, J. M., Starr, J. M., Horvath, S., Baccarelli, A. A., Levy, D., Visscher, P. M., Wray, N. R., & Deary, I. J. (2015). DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biol*, *16*, 25. https://doi.org/10.1186/s13059-015-0584-6

Martini, S., Nair, V., Keller, B. J., Eichinger, F., Hawkins, J. J., Randolph, A., Böger, C. A., Gadegbeku, C. A., Fox, C. S., Cohen, C. D., & Kretzler, M. (2014). Integrative biology identifies shared transcriptional networks in CKD. *J Am Soc Nephrol*, *25*(11), 2559–72. https://doi.org/10.1681/asn.2013080906

Matías-García, P. R., Ward-Caviness, C. K., Raffield, L. M., Gao, X., Zhang, Y., Wilson, R., Gào, X., Nano, J., Bostom, A., Colicino, E., Correa, A., Coull, B., Eaton, C., Hou, L., Just, A. C., Kunze, S., Lange, L., Lange, E., Lin,

X., Liu, S., Nwanaji-Enwerem, J. C., Reiner, A., Shen, J., Schöttker, B., Vokonas, P., Zheng, Y., Young, B., Schwartz, J., Horvath, S., Lu, A., Whitsel, E. A., Koenig, W., Adamski, J., Winkelmann, J., Brenner, H., Baccarelli, A. A., Gieger, C., Peters, A., Franceschini, N., & Waldenberger, M. (2021a). DNAm-based signatures of accelerated aging and mortality in blood are associated with low renal function. *Clin Epigenetics*, *13*(1), 121. https://doi.org/10.1186/s13148-021-01082-w

Matías-García, P. R., Wilson, R., Guo, Q., Zaghlool, S., Eales, J., Xu, X., Charchar, F., Dormer, J., Maalmi, H., Schlosser, P., Elhadad, M., Nano, J., Sharma, S., Peters, A., Fornoni, A., Mook-Kanamori, D., Winkelmann, J., Danesh, J., Di Angelantonio, E., Ouwehand, W., Watkins, N., Roberts, D., Petrera, A., Graumann, J., Koenig, W., Hveem, K., Jonasson, C., Köttgen, A., Butterworth, A., Prunotto, M., Hauck, S., Herder, C., Suhre, K., Gieger, C., Tomaszewski, M., Teumer, A., & Waldenberger, M. (2021b). Plasma proteomics of renal function: A trans-ethnic meta-analysis and Mendelian randomization study. *Journal of the American Society of Nephrology*. https://doi.org/10.1681/asn.2020071070

Matsushita, K., van der Velde, M., Astor, B. C., Woodward, M., Levey, A. S., de Jong, P. E., Coresh, J., & Gansevoort, R. T. (2010). Association of estimated glomerular filtration rate and albuminuria with all-cause and cardiovascular mortality in general population cohorts: A collaborative meta-analysis. *Lancet*, *375*(9731), 2073–81. https://doi.org/10.1016/s0140-6736(10)60674-5

McCarthy, E. T., Sharma, R., Sharma, M., Li, J. Z., Ge, X. L., Dileepan, K. N., & Savin, V. J. (1998). TNF-alpha increases albumin permeability of isolated rat glomeruli through the generation of superoxide. *J Am Soc Nephrol*, *9*(3), 433–8.

McGuire, A. L., Gabriel, S., Tishkoff, S. A., Wonkam, A., Chakravarti, A., Furlong, E. E. M., Treutlein, B., Meissner, A., Chang, H. Y., López-Bigas, N., Segal, E., & Kim, J.-S. (2020). The road ahead in genetics and genomics. *Nature Reviews Genetics*, *21*(10), 581–596. https://doi.org/10.1038/s41576-020-0272-6

McKusick, V. A., & Ruddle, F. H. (1987). A new discipline, a new name, a new journal. *Genomics*, *1*(1), 1–2. https://doi.org/https://doi.org/10.1016/0888-7543(87)90098-X

Mehta, A. J., Zanobetti, A., Bind, M. A., Kloog, I., Koutrakis, P., Sparrow, D., Vokonas, P. S., & Schwartz, J. D. (2016). Long-term exposure to ambient fine particulate matter and renal function in older men: The Veterans Administration Normative Aging Study. *Environ Health Perspect*, *124*(9), 1353–60. https://doi.org/10.1289/ehp.1510269

Miranda-Gonçalves, V., Lameirinhas, A., Henrique, R., & Jerónimo, C. (2018). Metabolism and epigenetic interplay in cancer: Regulation and putative therapeutic targets. *Frontiers in Genetics*, *9*(427). https://doi.org/10.3389/fgene.2018.00427

Mischak, H., Delles, C., Vlahou, A., & Vanholder, R. (2015). Proteomic biomarkers in kidney disease: Issues in development and implementation. *Nat Rev Nephrol*, *11*(4), 221–32. https://doi.org/10.1038/nrneph.2014.247

Mohammadi-Shemirani, P., Sjaarda, J., Gerstein, H. C., Treleaven, D. J., Walsh, M., Mann, J. F., McQueen, M. J., Hess, S., & Paré, G. (2019). A Mendelian randomization-based approach to identify early and sensitive diagnostic biomarkers of disease. *Clinical Chemistry*, *65*(3), 427–436. https://doi.org/10.1373/clinchem.2018.291104

Mook-Kanamori, D. O., Selim, M. M., Takiddin, A. H., Al-Homsi, H., Al-Mahmoud, K. A., Al-Obaidli, A., Zirie, M. A., Rowe, J., Yousri, N. A., Karoly, E. D., Kocher, T., Sekkal Gherbi, W., Chidiac, O. M., Mook-Kanamori, M. J., Abdul Kader, S., Al Muftah, W. A., McKeon, C., & Suhre, K. (2014). 1,5-anhydroglucitol in saliva is a noninvasive marker of short-term glycemic control. *J Clin Endocrinol Metab*, *99*(3), E479–83. https://doi.org/10.1210/jc.2013-3596

Moore, C., Bolton, T., Walker, M., Kaptoge, S., Allen, D., Daynes, M., Mehenny, S., Sambrook, J., Watkins, N. A., Miflin, G., Di Angelantonio, E., Ouwehand, W. H., Roberts, D. J., Danesh, J., & Thompson, S. G. (2016). Recruitment and representativeness of blood donors in the INTERVAL randomised trial

assessing varying inter-donation intervals. *Trials*, *17*(1), 458. https://doi.org/10.1186/s13063-016-1579-7

Morales-Nebreda, L., McLafferty, F. S., & Singer, B. D. (2018). DNA methylation as a transcriptional regulator of the immune system. *Transl Res*. https://doi.org/10.1016/j.trsl.2018.08.001

Morita, Y., Araki, H., Sugimoto, T., Takeuchi, K., Yamane, T., Maeda, T., Yamamoto, Y., Nishi, K., Asano, M., Shirahama-Noda, K., Nishimura, M., Uzu, T., Hara-Nishimura, I., Koya, D., Kashiwagi, A., & Ohkubo, I. (2007). Legumain/asparaginyl endopeptidase controls extracellular matrix remodeling through the degradation of fibronectin in mouse renal proximal tubular cells. *FEBS Lett*, *581*(7), 1417–24. https://doi.org/10.1016/j.febslet.2007.02.064

Mueller, A. L., McNamara, M. S., & Sinclair, D. A. (2020). Why does COVID-19 disproportionately affect older people? *Aging (Albany NY)*, *12*(10), 9959–9981. https://doi.org/10.18632/aging.103344

Muka, T., Koromani, F., Portilla, E., O'Connor, A., Bramer, W. M., Troup, J., Chowdhury, R., Dehghan, A., & Franco, O. H. (2016). The role of epigenetic modifications in cardiovascular disease: A systematic review. *Int J Cardiol*, *212*, 174–83. https://doi.org/10.1016/j.ijcard.2016.03.062

Nayor, M., Short, M. I., Rasheed, H., Lin, H., Jonasson, C., Yang, Q., Hveem, K., Felix, J. F., Morrison, A. C., Wild, P. S., Morley, M. P., Cappola, T. P., Benson, M. D., Ngo, D., Sinha, S., Keyes, M. J., Shen, D., Wang, T. J., Larson, M. G., Brumpton, B. M., Gerszten, R. E., Omland, T., & Vasan, R. S. (2020). Aptamer-based proteomic platform identifies novel protein predictors of incident heart failure and echocardiographic traits. *Circ Heart Fail*, *13*(5), e006749. https://doi.org/10.1161/circheartfailure.119.006749

Ngo, D., Wen, D., Gao, Y., Keyes, M. J., Drury, E. R., Katz, D. H., Benson, M. D., Sinha, S., Shen, D., Farrell, L. A., Peterson, B. D., Friedman, D. J., Elmariah, S., Young, B. A., Smith, J. G., Yang, Q., Vasan, R. S., Larson, M. G., Correa, A., Humphreys, B. D., Wang, T. J., Pollak, M. R., Wilson, J. G., Gerszten, R. E., & Rhee, E. P. (2020). Circulating testican-2

is a podocyte-derived marker of kidney health. *Proc Natl Acad Sci U S A*, *117*(40), 25026–25035. https://doi.org/10.1073/pnas.2009606117

Niewczas, M. A., Pavkov, M. E., Skupien, J., Smiles, A., Md Dom, Z. I., Wilson, J. M., Park, J., Nair, V., Schlafly, A., Saulnier, P. J., Satake, E., Simeone, C. A., Shah, H., Qiu, C., Looker, H. C., Fiorina, P., Ware, C. F., Sun, J. K., Doria, A., Kretzler, M., Susztak, K., Duffin, K. L., Nelson, R. G., & Krolewski, A. S. (2019). A signature of circulating inflammatory proteins and development of end-stage renal disease in diabetes. *Nat Med*, *25*(5), 805–813. https://doi.org/10.1038/s41591-019-0415-5

Nystrom, J., Fierlbeck, W., Granqvist, A., Kulak, S. C., & Ballermann, B. J. (2009). A human glomerular SAGE transcriptome database. *BMC Nephrol*, *10*, 13. https://doi.org/10.1186/1471-2369-10-13

Orth, S. R., & Hallan, S. I. (2008). Smoking: A risk factor for progression of chronic kidney disease and for cardiovascular morbidity and mortality in renal patients–absence of evidence or evidence of absence? *Clin J Am Soc Nephrol*, *3*(1), 226–36. https://doi.org/10.2215/cjn.03740907

Pierce, B. L., & Burgess, S. (2013). Efficient design for Mendelian randomization studies: Subsample and 2-sample instrumental variable estimators. *Am J Epidemiol*, *178*(7), 1177–84. https://doi.org/10.1093/aje/kwt084

Ping, P., Vondriska, T. M., Creighton, C. J., Gandhi, T. K., Yang, Z., Menon, R., Kwon, M. S., Cho, S. Y., Drwal, G., Kellmann, M., Peri, S., Suresh, S., Gronborg, M., Molina, H., Chaerkady, R., Rekha, B., Shet, A. S., Gerszten, R. E., Wu, H., Raftery, M., Wasinger, V., Schulz-Knappe, P., Hanash, S. M., Paik, Y. K., Hancock, W. S., States, D. J., Omenn, G. S., & Pandey, A. (2005). A functional annotation of subproteomes in human plasma. *Proteomics*, *5*(13), 3506–19. https://doi.org/10.1002/pmic.200500140

Prunotto, M., Farina, A., Lane, L., Pernin, A., Schifferli, J., Hochstrasser, D. F., Lescuyer, P., & Moll, S. (2013). Proteomic analysis of podocyte exosome-enriched fraction from normal human urine. *J Proteomics*, *82*, 193–229. https://doi.org/10.1016/j.jprot.2013.01.012

Quach, A., Levine, M. E., Tanaka, T., Lu, A. T., Chen, B. H., Ferrucci, L., Ritz, B., Bandinelli, S., Neuhouser, M. L., Beasley, J. M., Snetselaar, L., Wallace, R. B., Tsao, P. S., Absher, D., Assimes, T. L., Stewart, J. D., Li, Y., Hou, L., Baccarelli, A. A., Whitsel, E. A., & Horvath, S. (2017). Epigenetic clock analysis of diet, exercise, education, and lifestyle factors. *Aging (Albany NY)*, *9*(2), 419–446. https://doi.org/10.18632/aging.101168

R Core Team. (2019). *R: A language and environment for statistical computing*. https://www.R-project.org/

Rathmann, W., Strassburger, K., Heier, M., Holle, R., Thorand, B., Giani, G., & Meisinger, C. (2009). Incidence of type 2 diabetes in the elderly german population and the effect of clinical and lifestyle risk factors: KORA S4/F4 cohort study. *Diabet Med*, *26*(12), 1212–9. https://doi.org/10.1111/j.1464-5491.2009.02863.x

Rauch, T. A., & Pfeifer, G. P. (2011). Chapter 9 - methods for assessing genome-wide DNA methylation. In T. Tollefsbol (Ed.), *Handbook of epigenetics* (pp. 135–147). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-375709-8.00009-5

Raum, E., Rothenbacher, D., Low, M., Stegmaier, C., Ziegler, H., & Brenner, H. (2007). Changes of cardiovascular risk factors and their implications in subsequent birth cohorts of older adults in Germany: A life course approach. *Eur J Cardiovasc Prev Rehabil*, *14*(6), 809–14. https://doi.org/10.1097/HJR.0b013e3282eeb308

Rhee, C. M., & Kovesdy, C. P. (2015). Epidemiology: Spotlight on CKD deaths — increasing mortality worldwide. *Nat Rev Nephrol*, *11*(4), 199–200. https://doi.org/10.1038/nrneph.2015.25

Rohloff, J. C., Gelinas, A. D., Jarvis, T. C., Ochsner, U. A., Schneider, D. J., Gold, L., & Janjic, N. (2014). Nucleic acid ligands with protein-like side chains: Modified aptamers and their use as diagnostic and therapeutic agents. *Mol Ther Nucleic Acids*, *3*, e201. https://doi.org/10.1038/mtna.2014.49

Roshandel, D., Chen, Z., Canty, A. J., Bull, S. B., Natarajan, R., Paterson, A. D., Scherer, S., Miao, F., Zhang, L., Brown-Friday, J., Crandall, J., Engel, H.,

Engel, S., Martinez, H., Phillips, M., Reid, M., Shamoon, H., Sheindlin, J., Gubitosi-Klug, R., Wood, J., Mayer, L., Miller, D., Nayate, A., Novak, M., Pendegast, S., Singerman, L., Weiss, D., Zegarra, H., Brown, E., Crawford, P., Genuth, S., Palmert, M., Pugsley, P., Quin, J., Smith-Brewer, S., Dahms, W., McConnell, J., Beck, C., Farrell, K., Gaston, P., Trail, R., Monnier, V., Sell, D., Strauch, C., Gregory, N. S., Hanna, R., Chan, R., Kiss, S., Orlin, A., Rubin, M., Barron, S., Bosco, B., Brillon, D., Chang, S., Dwoskin, A., Heinemann, M., Jovanovic, L., Lackaye, M. E., Lee, T., Levy, B., Reppucci, V., Richardson, M., Campbell, R., Bhan, A., Jones, J. K., Kruger, D., Edwards, P. A., Remtema, H., Angus, E., Galprin, A., McLellan, M., Thomas, A., Carey, J. D., Whitehouse, F., Bergenstal, R., Dunnigan, S., Gunyou, K., Johnson, M., Carlson, A., Thomas, L., Birk, R., Callahan, P., Castle, G., Cuddihy, R., Franz, M., Freking, D., Gill, L., Gott, J., Hollander, P., Kendall, D., Laechelt, J., List, S., Mestrezat, W., Nelson, J., Olson, B., Rude, N., Spencer, M., Etzwiler, D., Morgan, K., Aiello, L. P., et al. (2020). DNA methylation age calculators reveal association with diabetic neuropathy in type 1 diabetes. *Clinical Epigenetics*, *12*(1), 52. https://doi.org/10.1186/s13148-020-00840-6

Rothman, K., Greenland, S., & Lash, T. (2008). Types of epidemiologic studies. In K. Rothman, S. Greenland, & T. Lash (Eds.), *Modern epidemiology* (pp. 87–99). Wolters Kluwer Health/Lippincott Williams & Wilkins. https://books.google.de/books?id=Z3vjT9ALxHUC

Rothman, N., World Health, O., Danmark, & International Agency for Research on, C. (2011). *Molecular epidemiology : Principles and practices*. International Agency for Research on Cancer, World Health Organization.

Rowland, J., Akbarov, A., Eales, J., Xu, X., Dormer, J. P., Guo, H., Denniff, M., Jiang, X., Ranjzad, P., Nazgiewicz, A., Prestes, P. R., Antczak, A., Szulinska, M., Wise, I. A., Zukowska-Szczechowska, E., Bogdanski, P., Woolf, A. S., Samani, N. J., Charchar, F. J., & Tomaszewski, M. (2019). Uncovering genetic mechanisms of kidney aging through transcriptomics, genomics, and epigenomics. *Kidney Int*, *95*(3), 624–635. https://doi.org/10.1016/j.kint.2018.10.029

Rowland, J., Akbarov, A., Maan, A., Eales, J., Dormer, J., & Tomaszewski, M. (2018). Tick-tock chimes the kidney clock – from biology of renal ageing to clinical applications. *Kidney and Blood Pressure Research*, *43*(1), 55–67. https://doi.org/10.1159/000486907

Sanchez-Nino, M. D., Sanz, A. B., Ramos, A. M., Fernandez-Fernandez, B., & Ortiz, A. (2017). Clinical proteomics in kidney disease as an exponential technology: Heading towards the disruptive phase. *Clin Kidney J*, *10*(2), 188–191. https://doi.org/10.1093/ckj/sfx023

Sarnak, M. J., Levey, A. S., Schoolwerth, A. C., Coresh, J., Culleton, B., Hamm, L. L., McCullough, P. A., Kasiske, B. L., Kelepouris, E., Klag, M. J., Parfrey, P., Pfeffer, M., Raij, L., Spinosa, D. J., & Wilson, P. W. (2003). Kidney disease as a risk factor for development of cardiovascular disease: A statement from the American Heart Association Councils on Kidney in Cardiovascular Disease, High Blood Pressure Research, Clinical Cardiology, and Epidemiology and Prevention. *Circulation*, *108*(17), 2154–69. https://doi.org/10.1161/01.cir.0000095676.90936.80

Schenk, S., Schoenhals, G. J., de Souza, G., & Mann, M. (2008). A high confidence, manually validated human blood plasma protein reference set. *BMC Med Genomics*, *1*, 41. https://doi.org/10.1186/1755-8794-1-41

Schormair, B., Plag, J., Kaffe, M., Gross, N., Czamara, D., Samtleben, W., Lichtner, P., Ströhle, A., Stefanidis, I., Vainas, A., Dardiotis, E., Sakkas, G. K., Gieger, C., Müller-Myhsok, B., Meitinger, T., Heemann, U., Hadjigeorgiou, G. M., Oexle, K., & Winkelmann, J. (2011). MEIS1 and BTBD9: Genetic association with restless leg syndrome in end stage renal disease. *Journal of medical genetics*, *48*(7), 462–466. https://doi.org/10.1136/jmg.2010.087858

Schulte, P., Rothman, N., Hainaut, P., Smith, M., Bofetta, P., & Perera, F. (2011). Molecular epidemiology: Linking molecular scale insights to population impacts. In N. Rothman, P. Hainaut, P. Schulte, M. Smith, P. Boffetta, F. Perera, W. H. Organization, & I. A. f. R. o. Cancer (Eds.), *Molecular epidemiology: Principles and practices* (pp. 1–7). International Agency for Research on Cancer, World Health Organization. https://books.google.de/books?id=PaEzzgEACAAJ

Schwab, K., Patterson, L. T., Aronow, B. J., Luckas, R., Liang, H.-C., & Potter, S. S. (2003). A catalogue of gene expression in the developing kidney. *Kidney International*, *64*(5), 1588–1604. https://doi.org/10.1046/j.1523-1755.2003.00276.x

Sekula, P., Del Greco M, F., Pattaro, C., & Köttgen, A. (2016). Mendelian randomization as an approach to assess causality using observational data. *Journal of the American Society of Nephrology : JASN*, *27*(11), 3253–3265. https://doi.org/10.1681/ASN.2016010098

Shiels, P. G., McGuinness, D., Eriksson, M., Kooman, J. P., & Stenvinkel, P. (2017). The role of epigenetics in renal ageing. *Nat Rev Nephrol*, *13*(8), 471–482. https://doi.org/10.1038/nrneph.2017.78

Smith, G. D., & Ebrahim, S. (2003). 'Mendelian randomization': Can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol*, *32*(1), 1–22. https://doi.org/10.1093/ije/dyg070

Smith, J. G., & Gerszten, R. E. (2017). Emerging affinity-based proteomic technologies for large-scale plasma profiling in cardiovascular disease. *Circulation*, *135*(17), 1651–1664. https://doi.org/doi:10.1161/CIRCULATIONAHA.116.025446

Smith, J. A., Raisky, J., Ratliff, S. M., Liu, J., Kardia, S. L. R., Turner, S. T., Mosley, T. H., & Zhao, W. (2019). Intrinsic and extrinsic epigenetic age acceleration are associated with hypertensive target organ damage in older African Americans. *BMC medical genomics*, *12*(1), 141–141. https://doi.org/10.1186/s12920-019-0585-5

Suhre, K., & Arnold, M. (2017). Connecting genetic risk to disease end points through the human blood plasma proteome. *8*, 14357. https://doi.org/10.1038/ncomms14357

Sun, B. B., Maranville, J. C., Peters, J. E., Stacey, D., Staley, J. R., Blackshaw, J., Burgess, S., Jiang, T., Paige, E., Surendran, P., Oliver-Williams, C., Kamat, M. A., Prins, B. P., Wilcox, S. K., Zimmerman, E. S., Chi, A., Bansal, N., Spain, S. L., Wood, A. M., Morrell, N. W., Bradley, J. R., Janjic, N., Roberts, D. J., Ouwehand, W. H., Todd, J. A., Soranzo, N.,

Suhre, K., Paul, D. S., Fox, C. S., Plenge, R. M., Danesh, J., Runz, H., & Butterworth, A. S. (2018). Genomic atlas of the human plasma proteome. *Nature*, *558*(7708), 73–79. https://doi.org/10.1038/s41586-018-0175-2

Suzuki, M. M., & Bird, A. (2008). DNA methylation landscapes: Provocative insights from epigenomics. *Nat Rev Genet*, *9*(6), 465–76. https://doi.org/10.1038/nrg2341

Tam, V., Patel, N., Turcotte, M., Bossé, Y., Paré, G., & Meyre, D. (2019). Benefits and limitations of genome-wide association studies. *Nat Rev Genet*, *20*(8), 467–484. https://doi.org/10.1038/s41576-019-0127-1

Taylor, J., H. A., Wilson, J. G., Jones, D. W., Sarpong, D. F., Srinivasan, A., Garrison, R. J., Nelson, C., & Wyatt, S. B. (2005). Toward resolution of cardiovascular health disparities in African Americans: Design and methods of the Jackson Heart Study. *Ethn Dis*, *15*(4 Suppl 6), S6-4–17.

Tecklenborg, J., Clayton, D., Siebert, S., & Coley, S. M. (2018). The role of the immune system in kidney disease. *Clin Exp Immunol*, *192*(2), 142–150. https://doi.org/10.1111/cei.13119

Teumer, A., Li, Y., Ghasemi, S., Prins, B. P., Wuttke, M., Hermle, T., Giri, A., Sieber, K. B., Qiu, C., Kirsten, H., Tin, A., Chu, A. Y., Bansal, N., Feitosa, M. F., Wang, L., Chai, J. F., Cocca, M., Fuchsberger, C., Gorski, M., Hoppmann, A., Horn, K., Li, M., Marten, J., Noce, D., Nutile, T., Sedaghat, S., Sveinbjornsson, G., Tayo, B. O., van der Most, P. J., Xu, Y., Yu, Z., Gerstner, L., Ärnlöv, J., Bakker, S. J. L., Baptista, D., Biggs, M. L., Boerwinkle, E., Brenner, H., Burkhardt, R., Carroll, R. J., Chee, M. L., Chee, M. L., Chen, M., Cheng, C. Y., Cook, J. P., Coresh, J., Corre, T., Danesh, J., de Borst, M. H., De Grandi, A., de Mutsert, R., de Vries, A. P. J., Degenhardt, F., Dittrich, K., Divers, J., Eckardt, K. U., Ehret, G., Endlich, K., Felix, J. F., Franco, O. H., Franke, A., Freedman, B. I., Freitag-Wolf, S., Gansevoort, R. T., Giedraitis, V., Gögele, M., Grundner-Culemann, F., Gudbjartsson, D. F., Gudnason, V., Hamet, P., Harris, T. B., Hicks, A. A., Holm, H., Foo, V. H. X., Hwang, S. J., Ikram, M. A., Ingelsson, E., Jaddoe, V. W. V., Jakobsdottir, J., Josyula, N. S., Jung, B., Kähönen, M., Khor, C. C., Kiess, W., Koenig, W., Körner, A., Kovacs, P., Kramer, H.,

Krämer, B. K., Kronenberg, F., Lange, L. A., Langefeld, C. D., Lee, J. J., Lehtimäki, T., Lieb, W., Lim, S. C., Lind, L., Lindgren, C. M., Liu, J., Loeffler, M., et al. (2019). Genome-wide association meta-analyses and fine-mapping elucidate pathways influencing albuminuria. *Nat Commun*, *10*(1), 4130. https://doi.org/10.1038/s41467-019-11576-0

The Women's Health Initiative Study Group. (1998). Design of the Women's Health Initiative clinical trial and observational study. *Control Clin Trials*, *19*(1), 61–109. https://doi.org/10.1016/s0197-2456(97)00078-0

Tin, A., Marten, J., Halperin Kuhns, V. L., Li, Y., Wuttke, M., Kirsten, H., Sieber, K. B., Qiu, C., Gorski, M., Yu, Z., Giri, A., Sveinbjornsson, G., Li, M., Chu, A. Y., Hoppmann, A., O'Connor, L. J., Prins, B., Nutile, T., Noce, D., Akiyama, M., Cocca, M., Ghasemi, S., van der Most, P. J., Horn, K., Xu, Y., Fuchsberger, C., Sedaghat, S., Afaq, S., Amin, N., Ärnlöv, J., Bakker, S. J. L., Bansal, N., Baptista, D., Bergmann, S., Biggs, M. L., Biino, G., Boerwinkle, E., Bottinger, E. P., Boutin, T. S., Brumat, M., Burkhardt, R., Campana, E., Campbell, A., Campbell, H., Carroll, R. J., Catamo, E., Chambers, J. C., Ciullo, M., Concas, M. P., Coresh, J., Corre, T., Cusi, D., Felicita, S. C., de Borst, M. H., De Grandi, A., de Mutsert, R., de Vries, A. P. J., Delgado, G., Demirkan, A., Devuyst, O., Dittrich, K., Eckardt, K. U., Ehret, G., Endlich, K., Evans, M. K., Gansevoort, R. T., Gasparini, P., Giedraitis, V., Gieger, C., Girotto, G., Gögele, M., Gordon, S. D., Gudbjartsson, D. F., Gudnason, V., Haller, T., Hamet, P., Harris, T. B., Hayward, C., Hicks, A. A., Hofer, E., Holm, H., Huang, W., Hutri-Kähönen, N., Hwang, S. J., Ikram, M. A., Lewis, R. M., Ingelsson, E., Jakobsdottir, J., Jonsdottir, I., Jonsson, H., Joshi, P. K., Josyula, N. S., Jung, B., Kähönen, M., Kamatani, Y., Kanai, M., Kerr, S. M., Kiess, W., Kleber, M. E., Koenig, W., et al. (2019). Target genes, variants, tissues and transcriptional pathways influencing human serum urate levels. *Nat Genet*, *51*(10), 1459–1474. https://doi.org/10.1038/s41588-019-0504-x

Tirumalai, R. S., Chan, K. C., Prieto, D. A., Issaq, H. J., Conrads, T. P., & Veenstra, T. D. (2003). Characterization of the low molecular weight human serum

proteome. *Mol Cell Proteomics*, *2*(10), 1096–103. https://doi.org/10.1074/mcp.M300031-MCP200

Tollefsbol, T. O. (2011). Chapter 1 - epigenetics: The new science of genetics. In T. Tollefsbol (Ed.), *Handbook of epigenetics* (pp. 1–6). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-375709-8.00001-0

Toto, R. D. (2004). Microalbuminuria: Definition, detection, and clinical significance. *J Clin Hypertens (Greenwich)*, *6*(11 Suppl 3), 2–7.

Trenkwalder, C., Allen, R., Hogl, B., Paulus, W., & Winkelmann, J. (2016). Restless legs syndrome associated with major diseases: A systematic review and new concept. *Neurology*, *86*(14), 1336–43. https://doi.org/10.1212/wnl.0000000000002542

Uhlén, M., Karlsson, M. J., Hober, A., Svensson, A.-S., Scheffel, J., Kotol, D., Zhong, W., Tebani, A., Strandberg, L., Edfors, F., Sjöstedt, E., Mulder, J., Mardinoglu, A., Berling, A., Ekblad, S., Dannemeyer, M., Kanje, S., Rockberg, J., Lundqvist, M., Malm, M., Volk, A.-L., Nilsson, P., Månberg, A., Dodig-Crnkovic, T., Pin, E., Zwahlen, M., Oksvold, P., von Feilitzen, K., Häussler, R. S., Hong, M.-G., Lindskog, C., Ponten, F., Katona, B., Vuu, J., Lindström, E., Nielsen, J., Robinson, J., Ayoglu, B., Mahdessian, D., Sullivan, D., Thul, P., Danielsson, F., Stadler, C., Lundberg, E., Bergström, G., Gummesson, A., Voldborg, B. G., Tegel, H., Hober, S., Forsström, B., Schwenk, J. M., Fagerberg, L., & Sivertsson, Å. (2019). The human secretome. *Science Signaling*, *12*(609), eaaz0274. https://doi.org/10.1126/scisignal.aaz0274

Van Vliet, A., Baelde, H. J., Vleming, L. J., de Heer, E., & Bruijn, J. A. (2001). Distribution of fibronectin isoforms in human renal disease. *J Pathol*, *193*(2), 256–62. https://doi.org/10.1002/1096-9896(2000)9999:9999⟨::aid-path783⟩3.0.co;2-p

Vaughan, D. E. (2005). PAI-1 and atherothrombosis. *J Thromb Haemost*, *3*(8), 1879–83. https://doi.org/10.1111/j.1538-7836.2005.01420.x

Viechtbauer, W. (2010). Conducting meta-analyses in r with the metafor package. *2010*, *36*(3), 48. https://doi.org/10.18637/jss.v036.i03

Vineis, P., Matullo, G., & Berwick, M. (2005). Molecular epidemiology. In W. Ahrens & I. Pigeot (Eds.), *Handbook of epidemiology* (pp. 1111–1138). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-26577-1_28

Visscher, P. M., Wray, N. R., Zhang, Q., Sklar, P., McCarthy, M. I., Brown, M. A., & Yang, J. (2017). 10 years of GWAS discovery: Biology, function, and translation. *Am J Hum Genet*, *101*(1), 5–22. https://doi.org/10.1016/j.ajhg.2017.06.005

Ward-Caviness, C. K., Pu, S., Martin, C. L., Galea, S., Uddin, M., Wildman, D. E., Koenen, K., & Aiello, A. E. (2020). Epigenetic predictors of all-cause mortality are associated with objective measures of neighborhood disadvantage in an urban population. *Clinical Epigenetics*, *12*(1), 44. https://doi.org/10.1186/s13148-020-00830-8

Williams, S. A., Kivimaki, M., Langenberg, C., Hingorani, A. D., Casas, J. P., Bouchard, C., Jonasson, C., Sarzynski, M. A., Shipley, M. J., Alexander, L., Ash, J., Bauer, T., Chadwick, J., Datta, G., DeLisle, R. K., Hagar, Y., Hinterberg, M., Ostroff, R., Weiss, S., Ganz, P., & Wareham, N. J. (2019). Plasma protein patterns as comprehensive indicators of health. *Nat Med*, *25*(12), 1851–1857. https://doi.org/10.1038/s41591-019-0665-2

Wilson, J. G., Rotimi, C. N., Ekunwe, L., Royal, C. D., Crump, M. E., Wyatt, S. B., Steffes, M. W., Adeyemo, A., Zhou, J., Taylor, J., H. A., & Jaquish, C. (2005). Study design for genetic analysis in the Jackson Heart Study. *Ethn Dis*, *15*(4 Suppl 6), S6-30–37.

Woroniecka, K. I., Park, A. S., Mohtat, D., Thomas, D. B., Pullman, J. M., & Susztak, K. (2011). Transcriptome analysis of human diabetic kidney disease. *Diabetes*, *60*(9), 2354–69. https://doi.org/10.2337/db10-1181

Wuttke, M., & Köttgen, A. (2016). Insights into kidney diseases from genome-wide association studies. *Nature Reviews Nephrology*, *12*, 549. https://doi.org/10.1038/nrneph.2016.107https://www.nature.com/articles/nrneph.2016.107#supplementary-information

Wuttke, M., Li, Y., Li, M., Sieber, K. B., Feitosa, M. F., Gorski, M., Tin, A., Wang, L., Chu, A. Y., Hoppmann, A., Kirsten, H., Giri, A., Chai, J.-F., Svein-

bjornsson, G., Tayo, B. O., Nutile, T., Fuchsberger, C., Marten, J., Cocca, M., Ghasemi, S., Xu, Y., Horn, K., Noce, D., van der Most, P. J., Sedaghat, S., Yu, Z., Akiyama, M., Afaq, S., Ahluwalia, T. S., Almgren, P., Amin, N., Ärnlöv, J., Bakker, S. J. L., Bansal, N., Baptista, D., Bergmann, S., Biggs, M. L., Biino, G., Boehnke, M., Boerwinkle, E., Boissel, M., Bottinger, E. P., Boutin, T. S., Brenner, H., Brumat, M., Burkhardt, R., Butterworth, A. S., Campana, E., Campbell, A., Campbell, H., Canouil, M., Carroll, R. J., Catamo, E., Chambers, J. C., Chee, M.-L., Chee, M.-L., Chen, X., Cheng, C.-Y., Cheng, Y., Christensen, K., Cifkova, R., Ciullo, M., Concas, M. P., Cook, J. P., Coresh, J., Corre, T., Sala, C. F., Cusi, D., Danesh, J., Daw, E. W., de Borst, M. H., De Grandi, A., de Mutsert, R., de Vries, A. P. J., Degenhardt, F., Delgado, G., Demirkan, A., Di Angelantonio, E., Dittrich, K., Divers, J., Dorajoo, R., Eckardt, K.-U., Ehret, G., Elliott, P., Endlich, K., Evans, M. K., Felix, J. F., Foo, V. H. X., Franco, O. H., Franke, A., Freedman, B. I., Freitag-Wolf, S., Friedlander, Y., Froguel, P., Gansevoort, R. T., Gao, H., Gasparini, P., Gaziano, J. M., Giedraitis, V., Gieger, C., et al. (2019). A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nature Genetics*, *51*(6), 957–972. https://doi.org/10.1038/s41588-019-0407-x

Xu, X., Eales, J. M., Akbarov, A., Guo, H., Becker, L., Talavera, D., Ashraf, F., Nawaz, J., Pramanik, S., Bowes, J., Jiang, X., Dormer, J., Denniff, M., Antczak, A., Szulinska, M., Wise, I., Prestes, P. R., Glyda, M., Bogdanski, P., Zukowska-Szczechowska, E., Berzuini, C., Woolf, A. S., Samani, N. J., Charchar, F. J., & Tomaszewski, M. (2018). Molecular insights into genome-wide association studies of chronic kidney disease-defining traits. *Nat Commun*, *9*(1), 4800. https://doi.org/10.1038/s41467-018-07260-4

Yamamoto, K., Takeshita, K., Kojima, T., Takamatsu, J., & Saito, H. (2005). Aging and plasminogen activator inhibitor-1 (PAI-1) regulation: Implication in the pathogenesis of thrombotic disorders in the elderly. *Cardiovasc Res*, *66*(2), 276–85. https://doi.org/10.1016/j.cardiores.2004.11.013

Young, G. H., & Wu, V. C. (2012). Klotho methylation is linked to uremic toxins and chronic kidney disease. *Kidney Int*, *81*(7), 611–2. https://doi.org/10.1038/ki.2011.461

Yu, M., Hazelton, W. D., Luebeck, G. E., & Grady, W. M. (2020). Epigenetic aging: More than just a clock when it comes to cancer. *Cancer Research*, *80*(3), 367–374. https://doi.org/10.1158/0008-5472.can-19-0924

Zanini, J. C., Pietzner, M., & Langenberg, C. (2020). Integrating genetics and the plasma proteome to predict the risk of type 2 diabetes. *Current Diabetes Reports*, *20*(11), 60. https://doi.org/10.1007/s11892-020-01340-w

Zaykin, D. V. (2011). Optimally weighted z-test is a powerful method for combining probabilities in meta-analysis. *J Evol Biol*, *24*(8), 1836–41. https://doi.org/10.1111/j.1420-9101.2011.02297.x

Zeilinger, S., Kuhnel, B., Klopp, N., Baurecht, H., Kleinschmidt, A., Gieger, C., Weidinger, S., Lattka, E., Adamski, J., Peters, A., Strauch, K., Waldenberger, M., & Illig, T. (2013). Tobacco smoking leads to extensive genome-wide changes in DNA methylation. *PLoS One*, *8*(5), e63812. https://doi.org/10.1371/journal.pone.0063812

Zhang, Y., Wilson, R., Heiss, J., Breitling, L. P., Saum, K.-U., Schöttker, B., Holleczek, B., Waldenberger, M., Peters, A., & Brenner, H. (2017). DNA methylation signatures in peripheral blood strongly predict all-cause mortality. *Nature Communications*, *8*, 14617. https://doi.org/10.1038/ncomms14617https://www.nature.com/articles/ncomms14617#supplementary-information

License details

The license details, terms and conditions provided by the publishers for the inclusion of the first-author publications and other content from external publications in this dissertation are available in the next pages.

SPRINGER NATURE

BioMed Central Limited
The Campus
4 Crinan Street
London N1 9XW
United Kingdom

T +44 20 7833 4000
F +44 20 7833 4640
www.springernature.com

Joseph S Hasan
Senior Journal Development Editor,
Medicine and Life Sciences – Journals
T + 1 (212) 451 8737
Joseph.Hasan@springernature.com

3 May 2021

Re: including "DNAm-based signatures of accelerated aging and mortality in blood are associated with low renal function" in a publication-based doctoral dissertation

To Whom It May Concern:

This is Joseph Hasan, a Senior Journal Development Editor at BMC. We publish the journal *Clinical Epigenetics*. Pamela R. Matías-García, as the first author of "DNAm-based signatures of accelerated aging and mortality in blood are associated with low renal function", is allowed to include this article on a publication-based doctoral dissertation to be submitted in June to the Technical University of Munich. The only requirement we have is that the original authors, citation details and publisher are identified.

Yours sincerely,

Joseph S Hasan

Joseph Hasan
Senior Journal Development Editor
Medicine and Life Sciences – Journals
BMC

Hi Pamela,

Thank you for sending these forms. At this time the only outstanding forms are those of Dr. Roberts. We still need his copyright form and disclosure form.

You have permission to use this paper for your dissertation as you've described below. We allow authors to deposit their articles in their internal institutional repositories while their paper is still under embargo.

I hope this more clearly answers your question.

Best,

Sydney Cough
Managing Editor, *JASN*
scough@asn-online.org

From: Pamela Matias <pamela.matias@helmholtz-muenchen.de>
Sent: Tuesday, May 4, 2021 5:05 AM
To: Sydney Cough <scough@asn-online.org>
Subject: Re: JASN-2020-07-1070.R2- Outstanding Forms Needed

Dear Sydney,

Hope this e-mail finds you well. Attached to this e-mail are the disclosure forms for two of the co-authors with outstanding forms, in case they were not able to upload them to the journal's website. May I ask if the forms from Prof. Dave Roberts are still outstanding?

I have a few questions regarding the consent for me to use this paper in my dissertation and compliance with the copyright held by the journal. By including this paper in my dissertation, a copy of it would be appended to it, which will be made available to the academic community with access to my university's library. Would this be a case of infringement to the copyright, or would this fall in the category of "fair use"?

Would it be possible to receive a letter from the journal with a clarification on the above mentioned? That would be really helpful.

Thank you again.

All the best,
Pamela

Von: "scough" <scough@asn-online.org>
An: "Matias Garcia, Pamela Raquel" <pamela.matias@helmholtz-muenchen.de>
Gesendet: Montag, 26. April 2021 18:53:37
Betreff: RE: JASN-2020-07-1070.R2- Outstanding Forms Needed

Hi Pamela,

Sure, no problem. The following forms are still outstanding.

Missing copyrights: David Roberts
Missing Disclosures: John Danesh, Nicholas Watkins, David Roberts

The paper is currently with the production team for copyediting and formatting. They'll reach out as soon as your proofs are ready.

You are welcome to include this paper in your dissertation. Please make sure to cite it accordingly. The paper will be online before that date, but will be published within an issue in August.

Please let me know if you have any other questions.

Best,

Sydney Cough
Managing Editor, *JASN*
scough@asn-online.org

From: Pamela Matias <pamela.matias@helmholtz-muenchen.de>
Sent: Monday, April 26, 2021 8:02 AM
To: Sydney Cough <scough@asn-online.org>
Subject: Re: JASN-2020-07-1070.R2- Outstanding Forms Needed

Dear Sydney,

Thank you for your continuous support with co-author queries. May I please know whose forms are still missing? I will again reach out to get these forms in place.

Once these forms are in place, what would be the next steps in the publication process? Is there any additional information missing from our side prior to receiving the proofs?

I would also like to enquire about the inclusion of this article in a publication-based dissertation. This dissertation is to be submitted by myself mid-June to the Technical University of Munich (TUM) for the awardment of a doctoral degree. May I please receive a letter of approval from the publisher for this purpose?

Thank you again.

Best regards,
Pamela Matías

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 15, 2021

This Agreement between Mrs. Pamela Matias ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5070281242579
License date	May 15, 2021
Licensed Content Publisher	Elsevier
Licensed Content Publication	Genomics
Licensed Content Title	A new discipline, a new name, a new journal
Licensed Content Author	Victor A. McKusick,Frank H. Ruddle
Licensed Content Date	Sep 1, 1987
Licensed Content Volume	1
Licensed Content Issue	1
Licensed Content Pages	2
Start Page	1
End Page	2

Type of Use	reuse in a thesis/dissertation
Portion	excerpt
Number of excerpts	3
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	A multiple omics approach to kidney function and related traits
Institution name	Technical University of Munich (TUM)
Expected presentation date	Jun 2021
Portions	Text from page 1: "In his essay entitled (...) One might add genetics, and now, genomics"
Requestor Location	Mrs. Pamela Matias Ingolstädter Landstr. 1 Ingolstädter Landstr. 1 Neuherberg, 85764 Germany Attn: Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Publisher Tax ID	GB 494 6272 12
Total	0.00 USD
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk here). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all

value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.10

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 17, 2021

This Agreement between Mrs. Pamela Matias ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5071380343322
License date	May 17, 2021
Licensed Content Publisher	Elsevier
Licensed Content Publication	Molecular Cell
Licensed Content Title	DNA Methylation Clocks in Aging: Categories, Causes, and Consequences
Licensed Content Author	Adam E. Field,Neil A. Robertson,Tina Wang,Aaron Havas,Trey Ideker,Peter D. Adams
Licensed Content Date	Sep 20, 2018
Licensed Content Volume	71
Licensed Content Issue	6
Licensed Content Pages	14
Start Page	882
End Page	895

Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	A multiple omics approach to kidney function and related traits
Institution name	Technical University of Munich (TUM)
Expected presentation date	Jun 2021
Portions	Figures 1 and 3
Requestor Location	Mrs. Pamela Matias Ingolstädter Landstr. 1 Ingolstädter Landstr. 1 Neuherberg, 85764 Germany Attn: Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Publisher Tax ID	GB 494 6272 12
Total	0.00 EUR
Terms and Conditions	

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk here). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all

value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.10

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

AMERICAN ASSOCIATION FOR CANCER RESEARCH LICENSE
TERMS AND CONDITIONS

May 17, 2021

This Agreement between Mrs. Pamela Matias ("You") and American Association for Cancer Research ("American Association for Cancer Research") consists of your license details and the terms and conditions provided by American Association for Cancer Research and Copyright Clearance Center.

License Number	5071300887983
License date	May 17, 2021
Licensed Content Publisher	American Association for Cancer Research
Licensed Content Publication	Cancer Research
Licensed Content Title	Epigenetic Aging: More Than Just a Clock When It Comes to Cancer
Licensed Content Author	Ming Yu,William D. Hazelton,Georg E. Luebeck,William M. Grady
Licensed Content Date	Feb 1, 2020
Licensed Content Volume	80
Licensed Content Issue	3
Type of Use	Thesis/Dissertation
Requestor type	academic/educational

Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Will you be translating?	no
Circulation	50
Territory of distribution	Worldwide
Title	A multiple omics approach to kidney function and related traits
Institution name	Technical University of Munich (TUM)
Expected presentation date	Jun 2021
Portions	Figure 1
Requestor Location	Mrs. Pamela Matias Ingolstädter Landstr. 1 Ingolstädter Landstr. 1 Neuherberg, 85764 Germany Attn: Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Total	0.00 EUR

Terms and Conditions

American Association for Cancer Research (AACR) Terms and Conditions

INTRODUCTION
The Publisher for this copyright material is the American Association for Cancer Research (AACR). By clicking "accept" in connection with completing this licensing transaction, you

agree to the following terms and conditions applying to this transaction. You also agree to the Billing and Payment terms and conditions established by Copyright Clearance Center (CCC) at the time you opened your Rightslink account.

LIMITED LICENSE
The AACR grants exclusively to you, the User, for onetime, non-exclusive use of this material for the purpose stated in your request and used only with a maximum distribution equal to the number you identified in the permission process. Any form of republication must be completed within one year although copies made before then may be distributed thereafter and any electronic posting is limited to a period of one year. Reproduction of this material is confined to the purpose and/or media for which permission is granted. Altering or modifying this material is not permitted. However, figures and illustrations may be minimally altered or modified to serve the new work.

GEOGRAPHIC SCOPE
Licenses may be exercised as noted in the permission process

RESERVATION OF RIGHTS
The AACR reserves all rights not specifically granted in the combination of 1) the license details provided by you and accepted in the course of this licensing transaction, 2) these terms and conditions , and 3) CCC's Billing and Payment terms and conditions.

DISCLAIMER
You may obtain permission via Rightslink to use material owned by AACR. When you are requesting permission to reuse a portion for an AACR publication, it is your responsibility to examine each portion of content as published to determine whether a credit to, or copyright notice of a third party owner is published next to the item. You must obtain permission from the third party to use any material which has been reprinted with permission from the said third party. If you have not obtained permission from the third party, AACR disclaims any responsibility for the use you make of items owned by them.

LICENSE CONTINGENT ON PAYMENT
While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you, either by the publisher or by the CCC, as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions, or any of the CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and the publisher reserves the right to take any and all action to protect its copyright in the materials.

COPYRIGHT NOTICE
You must include the following credit line in connection with your reproduction of the licensed material: "Reprinted (or adapted) from Publication Title, Copyright Year, Volume/Issue, Page Range, Author, Title of Article, with permission from AACR".

TRANSLATION
This permission is granted for non-exclusive world English rights only.

WARRANTIES
Publisher makes no representations or warranties with respect to the licensed material.

INDEMNIFICATION
You hereby indemnify and agree to hold harmless the publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

REVOCATION
The AACR reserves the right to revoke a license for any reason, including but not limited to advertising and promotional uses of AACR content, third party usage and incorrect figure source attribution.

NO TRANSFER OF LICENSE
This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

NO AMENDMENT EXCEPT IN WRITING
This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

OBJECTION TO CONTRARY TERMS
Publishers hereby objects to any terms contained in any purchase order, acknowledgement, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions together with CCC's Billing and Payment terms and conditions (which are incorporated herein) comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions, and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

THESIS/DISSERTATION TERMS
If your request is to reuse an article authored by you and published by the AACR in your dissertation/thesis, your thesis may be submitted to your institution in either in print or electronic form. Should your thesis be published commercially, please reapply.

ELECTRONIC RESERVE
If this license is made in connection with a course, and the Licensed Material or any portion thereof is to be posted to a website, the website is to be password protected and made available only to the students registered for the relevant course. The permission is granted for the duration of the course. All content posted to the website must maintain the copyright information notice.

JURISDICTION
This license transaction shall be governed by and construed in accordance with the laws of Pennsylvania. You hereby agree to submit to the jurisdiction of the federal and state courts located in Pennsylvania for purposes of resolving any disputes that may arise in connection with this licensing transaction.

Other Terms and Conditions:

v1.0

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

WOLTERS KLUWER HEALTH, INC. LICENSE
TERMS AND CONDITIONS

May 17, 2021

This Agreement between Mrs. Pamela Matias ("You") and Wolters Kluwer Health, Inc. ("Wolters Kluwer Health, Inc.") consists of your license details and the terms and conditions provided by Wolters Kluwer Health, Inc. and Copyright Clearance Center.

License Number	5071491370445
License date	May 17, 2021
Licensed Content Publisher	Wolters Kluwer Health, Inc.
Licensed Content Publication	Circulation
Licensed Content Title	Emerging Affinity-Based Proteomic Technologies for Large-Scale Plasma Profiling in Cardiovascular Disease
Licensed Content Author	J. Gustav Smith, Robert E. Gerszten
Licensed Content Date	Apr 25, 2017
Licensed Content Volume	135
Licensed Content Issue	17
Type of Use	Dissertation/Thesis
Requestor type	University/College
Sponsorship	No Sponsorship

Format	Print and electronic
Will this be posted online?	Yes, on a secure website
Portion	Figures/tables/illustrations
Number of figures/tables/illustrations	2
Author of this Wolters Kluwer article	No
Will you be translating?	No
Intend to modify/change the content	No
Title	A multiple omics approach to kidney function and related traits
Institution name	Technical University of Munich (TUM)
Expected presentation date	Jun 2021
Portions	Figure 2, Table 2
Requestor Location	Mrs. Pamela Matias Ingolstädter Landstr. 1 Ingolstädter Landstr. 1 Neuherberg, 85764 Germany Attn: Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Publisher Tax ID	EU826013006
Total	0.00 EUR

Terms and Conditions

Wolters Kluwer Health Inc. Terms and Conditions

1. **Duration of License:** Permission is granted for a one time use only. Rights herein do not apply to future reproductions, editions, revisions, or other derivative works. This permission shall be effective as of the date of execution by the parties for the maximum period of 12 months and should be renewed after the term expires.
    i. When content is to be republished in a book or journal the validity of this agreement should be the life of the book edition or journal issue.
    ii. When content is licensed for use on a website, internet, intranet, or any publicly accessible site (not including a journal or book), you agree to remove the material from such site after 12 months, or request to renew your permission license
2. **Credit Line:** A credit line must be prominently placed and include: For book content: the author(s), title of book, edition, copyright holder, year of publication; For journal content: the author(s), titles of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page; If a journal is published by a learned society the credit line must include the details of that society.
3. **Warranties:** The requestor warrants that the material shall not be used in any manner which may be considered derogatory to the title, content, authors of the material, or to Wolters Kluwer Health, Inc.
4. **Indemnity:** You hereby indemnify and hold harmless Wolters Kluwer Health, Inc. and its respective officers, directors, employees and agents, from and against any and all claims, costs, proceeding or demands arising out of your unauthorized use of the Licensed Material
5. **Geographical Scope:** Permission granted is non-exclusive and is valid throughout the world in the English language and the languages specified in the license.
6. **Copy of Content:** Wolters Kluwer Health, Inc. cannot supply the requestor with the original artwork, high-resolution images, electronic files or a clean copy of content.
7. **Validity:** Permission is valid if the borrowed material is original to a Wolters Kluwer Health, Inc. imprint (J.B Lippincott, Lippincott-Raven Publishers, Williams & Wilkins, Lea & Febiger, Harwal, Rapid Science, Little Brown & Company, Harper & Row Medical, American Journal of Nursing Co, and Urban & Schwarzenberg - English Language, Raven Press, Paul Hoeber, Springhouse, Ovid), and the Anatomical Chart Company
8. **Third Party Material:** This permission does not apply to content that is credited to publications other than Wolters Kluwer Health, Inc. or its Societies. For images credited to non-Wolters Kluwer Health, Inc. books or journals, you must obtain permission from the source referenced in the figure or table legend or credit line before making any use of the image(s), table(s) or other content.
9. **Adaptations:** Adaptations are protected by copyright. For images that have been adapted, permission must be sought from the rightsholder of the original material and the rightsholder of the adapted material.
10. **Modifications:** Wolters Kluwer Health, Inc. material is not permitted to be modified or adapted without written approval from Wolters Kluwer Health, Inc. with the exception of text size or color. The adaptation should be credited as follows: Adapted with permission from Wolters Kluwer Health, Inc.: [the author(s), title of book, edition, copyright holder, year of publication] or [the author(s), titles of article, title of journal, volume number, issue number, inclusive pages and website URL to the journal page].

11. **Full Text Articles:** Republication of full articles in English is prohibited.
12. **Branding and Marketing:** No drug name, trade name, drug logo, or trade logo can be included on the same page as material borrowed from *Diseases of the Colon & Rectum, Plastic Reconstructive Surgery, Obstetrics & Gynecology (The Green Journal), Critical Care Medicine, Pediatric Critical Care Medicine, the American Heart Association publications and the American Academy of Neurology publications.*
13. **Open Access:** Unless you are publishing content under the same Creative Commons license, the following statement must be added when reprinting material in Open Access journals: "The Creative Commons license does not apply to this content. Use of the material in any format is prohibited without written permission from the publisher, Wolters Kluwer Health, Inc. Please contact permissions@lww.com for further information."
14. **Translations:** The following disclaimer must appear on all translated copies: Wolters Kluwer Health, Inc. and its Societies take no responsibility for the accuracy of the translation from the published English original and are not liable for any errors which may occur.
15. **Published Ahead of Print (PAP):** Articles in the PAP stage of publication can be cited using the online publication date and the unique DOI number.
    i. Disclaimer: Articles appearing in the PAP section have been peer-reviewed and accepted for publication in the relevant journal and posted online before print publication. Articles appearing as PAP may contain statements, opinions, and information that have errors in facts, figures, or interpretation. Any final changes in manuscripts will be made at the time of print publication and will be reflected in the final electronic version of the issue. Accordingly, Wolters Kluwer Health, Inc., the editors, authors and their respective employees are not responsible or liable for the use of any such inaccurate or misleading data, opinion or information contained in the articles in this section.
16. **Termination of Contract:** Wolters Kluwer Health, Inc. must be notified within 90 days of the original license date if you opt not to use the requested material.
17. **Waived Permission Fee:** Permission fees that have been waived are not subject to future waivers, including similar requests or renewing a license.
18. **Contingent on payment:** You may exercise these rights licensed immediately upon issuance of the license, however until full payment is received either by the publisher or our authorized vendor, this license is not valid. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of Wolters Kluwer Health, Inc.'s other billing and payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement and publisher reserves the right to take any and all action to protect its copyright in the materials.
19. **STM Signatories Only:** Any permission granted for a particular edition will apply to subsequent editions and for editions in other languages, provided such editions are for the work as a whole in situ and do not involve the separate exploitation of the permitted illustrations or excerpts. Please view: STM Permissions Guidelines
20. **Warranties and Obligations:** LICENSOR further represents and warrants that, to the best of its knowledge and belief, LICENSEE's contemplated use of the Content as represented to LICENSOR does not infringe any valid rights to any third party.
21. **Breach:** If LICENSEE fails to comply with any provisions of this agreement, LICENSOR may serve written notice of breach of LICENSEE and, unless such breach is fully cured within fifteen (15) days from the receipt of notice by LICENSEE, LICENSOR may thereupon, at its option, serve notice of cancellation on LICENSEE, whereupon this Agreement shall immediately terminate.

22. **Assignment:** License conveyed hereunder by the LICENSOR shall not be assigned or granted in any manner conveyed to any third party by the LICENSEE without the consent in writing to the LICENSOR.
23. **Governing Law:** The laws of The State of New York shall govern interpretation of this Agreement and all rights and liabilities arising hereunder.
24. **Unlawful:** If any provision of this Agreement shall be found unlawful or otherwise legally unenforceable, all other conditions and provisions of this Agreement shall remain in full force and effect.

For Copyright Clearance Center / RightsLink Only:

1. **Service Description for Content Services:** Subject to these terms of use, any terms set forth on the particular order, and payment of the applicable fee, you may make the following uses of the ordered materials:
    i. **Content Rental:** You may access and view a single electronic copy of the materials ordered for the time period designated at the time the order is placed. Access to the materials will be provided through a dedicated content viewer or other portal, and access will be discontinued upon expiration of the designated time period. An order for Content Rental does not include any rights to print, download, save, create additional copies, to distribute or to reuse in any way the full text or parts of the materials.
    ii. **Content Purchase:** You may access and download a single electronic copy of the materials ordered. Copies will be provided by email or by such other means as publisher may make available from time to time. An order for Content Purchase does not include any rights to create additional copies or to distribute copies of the materials

Other Terms and Conditions:

v1.18

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.

ELSEVIER LICENSE
TERMS AND CONDITIONS

May 18, 2021

This Agreement between Mrs. Pamela Matias ("You") and Elsevier ("Elsevier") consists of your license details and the terms and conditions provided by Elsevier and Copyright Clearance Center.

License Number	5071950777968
License date	May 18, 2021
Licensed Content Publisher	Elsevier
Licensed Content Publication	American Journal of Kidney Diseases
Licensed Content Title	KDOQI US Commentary on the 2012 KDIGO Clinical Practice Guideline for the Evaluation and Management of CKD
Licensed Content Author	Lesley A. Inker,Brad C. Astor,Chester H. Fox,Tamara Isakova,James P. Lash,Carmen A. Peralta,Manjula Kurella Tamura,Harold I. Feldman
Licensed Content Date	May 1, 2014
Licensed Content Volume	63
Licensed Content Issue	5
Licensed Content Pages	23
Start Page	713

End Page	735
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1
Format	both print and electronic
Are you the author of this Elsevier article?	No
Will you be translating?	No
Title	A multiple omics approach to kidney function and related traits
Institution name	Technical University of Munich (TUM)
Expected presentation date	Jun 2021
Portions	Figure titled “Prognosis of CKD by GFR and Albuminuria Categories: KDIGO 2012”
Requestor Location	Mrs. Pamela Matias Ingolstädter Landstr. 1 Ingolstädter Landstr. 1 Neuherberg, 85764 Germany Attn: Helmholtz Zentrum München Deutsches Forschungszentrum für Gesundheit und Umwelt (GmbH)
Publisher Tax ID	GB 494 6272 12
Total	0.00 EUR

Terms and Conditions

INTRODUCTION

1. The publisher for this copyrighted material is Elsevier. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

GENERAL TERMS

2. Elsevier hereby grants you permission to reproduce the aforementioned material subject to the terms and conditions indicated.

3. Acknowledgement: If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source, permission must also be sought from that source. If such permission is not obtained then that material may not be included in your publication/copies. Suitable acknowledgement to the source must be made, either as a footnote or in a reference list at the end of your publication, as follows:

"Reprinted from Publication title, Vol /edition number, Author(s), Title of article / title of chapter, Pages No., Copyright (Year), with permission from Elsevier [OR APPLICABLE SOCIETY COPYRIGHT OWNER]." Also Lancet special credit - "Reprinted from The Lancet, Vol. number, Author(s), Title of article, Pages No., Copyright (Year), with permission from Elsevier."

4. Reproduction of this material is confined to the purpose and/or media for which permission is hereby given.

5. Altering/Modifying Material: Not Permitted. However figures and illustrations may be altered/adapted minimally to serve your work. Any other abbreviations, additions, deletions and/or any other alterations shall be made only with prior written authorization of Elsevier Ltd. (Please contact Elsevier's permissions helpdesk here). No modifications can be made to any Lancet figures/tables and they must be reproduced in full.

6. If the permission fee for the requested use of our material is waived in this instance, please be advised that your future requests for Elsevier materials may attract a fee.

7. Reservation of Rights: Publisher reserves all rights not specifically granted in the combination of (i) the license details provided by you and accepted in the course of this licensing transaction, (ii) these terms and conditions and (iii) CCC's Billing and Payment terms and conditions.

8. License Contingent Upon Payment: While you may exercise the rights licensed immediately upon issuance of the license at the end of the licensing process for the transaction, provided that you have disclosed complete and accurate details of your proposed use, no license is finally effective unless and until full payment is received from you (either by publisher or by CCC) as provided in CCC's Billing and Payment terms and conditions. If full payment is not received on a timely basis, then any license preliminarily granted shall be deemed automatically revoked and shall be void as if never granted. Further, in the event that you breach any of these terms and conditions or any of CCC's Billing and Payment terms and conditions, the license is automatically revoked and shall be void as if never granted. Use of materials as described in a revoked license, as well as any use of the materials beyond the scope of an unrevoked license, may constitute copyright infringement

and publisher reserves the right to take any and all action to protect its copyright in the materials.

9. Warranties: Publisher makes no representations or warranties with respect to the licensed material.

10. Indemnity: You hereby indemnify and agree to hold harmless publisher and CCC, and their respective officers, directors, employees and agents, from and against any and all claims arising out of your use of the licensed material other than as specifically authorized pursuant to this license.

11. No Transfer of License: This license is personal to you and may not be sublicensed, assigned, or transferred by you to any other person without publisher's written permission.

12. No Amendment Except in Writing: This license may not be amended except in a writing signed by both parties (or, in the case of publisher, by CCC on publisher's behalf).

13. Objection to Contrary Terms: Publisher hereby objects to any terms contained in any purchase order, acknowledgment, check endorsement or other writing prepared by you, which terms are inconsistent with these terms and conditions or CCC's Billing and Payment terms and conditions. These terms and conditions, together with CCC's Billing and Payment terms and conditions (which are incorporated herein), comprise the entire agreement between you and publisher (and CCC) concerning this licensing transaction. In the event of any conflict between your obligations established by these terms and conditions and those established by CCC's Billing and Payment terms and conditions, these terms and conditions shall control.

14. Revocation: Elsevier or Copyright Clearance Center may deny the permissions described in this License at their sole discretion, for any reason or no reason, with a full refund payable to you. Notice of such denial will be made using the contact information provided by you. Failure to receive such notice will not alter or invalidate the denial. In no event will Elsevier or Copyright Clearance Center be responsible or liable for any costs, expenses or damage incurred by you as a result of a denial of your permission request, other than a refund of the amount(s) paid by you to Elsevier and/or Copyright Clearance Center for denied permissions.

LIMITED LICENSE

The following terms and conditions apply only to specific license types:

15. **Translation**: This permission is granted for non-exclusive world **English** rights only unless your license was granted for translation rights. If you licensed translation rights you may only translate this content into the languages you requested. A professional translator must perform all translations and reproduce the content word for word preserving the integrity of the article.

16. **Posting licensed content on any Website**: The following terms and conditions apply as follows: Licensing material from an Elsevier journal: All content posted to the web site must maintain the copyright information line on the bottom of each image; A hyper-text must be included to the Homepage of the journal from which you are licensing at http://www.sciencedirect.com/science/journal/xxxxx or the Elsevier homepage for books at http://www.elsevier.com; Central Storage: This license does not include permission for a scanned version of the material to be stored in a central repository such as that provided by Heron/XanEdu.

Licensing material from an Elsevier book: A hyper-text link must be included to the Elsevier homepage at http://www.elsevier.com . All content posted to the web site must maintain the copyright information line on the bottom of each image.

Posting licensed content on Electronic reserve: In addition to the above the following clauses are applicable: The web site must be password-protected and made available only to bona fide students registered on a relevant course. This permission is granted for 1 year only. You may obtain a new license for future website posting.

17. **For journal authors:** the following clauses are applicable in addition to the above:

Preprints:

A preprint is an author's own write-up of research results and analysis, it has not been peer-reviewed, nor has it had any other value added to it by a publisher (such as formatting, copyright, technical enhancement etc.).

Authors can share their preprints anywhere at any time. Preprints should not be added to or enhanced in any way in order to appear more like, or to substitute for, the final versions of articles however authors can update their preprints on arXiv or RePEc with their Accepted Author Manuscript (see below).

If accepted for publication, we encourage authors to link from the preprint to their formal publication via its DOI. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help users to find, access, cite and use the best available version. Please note that Cell Press, The Lancet and some society-owned have different preprint policies. Information on these policies is available on the journal homepage.

Accepted Author Manuscripts: An accepted author manuscript is the manuscript of an article that has been accepted for publication and which typically includes author-incorporated changes suggested during submission, peer review and editor-author communications.

Authors can share their accepted author manuscript:

- immediately
    - via their non-commercial person homepage or blog
    - by updating a preprint in arXiv or RePEc with the accepted manuscript
    - via their research institute or institutional repository for internal institutional uses or as part of an invitation-only research collaboration work-group
    - directly by providing copies to their students or to research collaborators for their personal use
    - for private scholarly sharing as part of an invitation-only work group on commercial sites with which Elsevier has an agreement
- After the embargo period
    - via non-commercial hosting platforms such as their institutional repository
    - via commercial sites with which Elsevier has an agreement

In all cases accepted manuscripts should:

- link to the formal publication via its DOI
- bear a CC-BY-NC-ND license - this is easy to do
- if aggregated with other manuscripts, for example in a repository or other site, be shared in alignment with our hosting policy not be added to or enhanced in any way to

appear more like, or to substitute for, the published journal article.

Published journal article (JPA): A published journal article (PJA) is the definitive final record of published research that appears or will appear in the journal and embodies all value-adding publishing activities including peer review co-ordination, copy-editing, formatting, (if relevant) pagination and online enrichment.

Policies for sharing publishing journal articles differ for subscription and gold open access articles:

Subscription Articles: If you are an author, please share a link to your article rather than the full-text. Millions of researchers have access to the formal publications on ScienceDirect, and so links will help your users to find, access, cite, and use the best available version.

Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

If you are affiliated with a library that subscribes to ScienceDirect you have additional private sharing rights for others' research accessed under that agreement. This includes use for classroom teaching and internal training at the institution (including use in course packs and courseware programs), and inclusion of the article for grant funding purposes.

Gold Open Access Articles: May be shared according to the author-selected end-user license and should contain a CrossMark logo, the end user license, and a DOI link to the formal publication on ScienceDirect.

Please refer to Elsevier's posting policy for further information.

18. **For book authors** the following clauses are applicable in addition to the above: Authors are permitted to place a brief summary of their work online only. You are not allowed to download and post the published electronic version of your chapter, nor may you scan the printed edition to create an electronic version. **Posting to a repository:** Authors are permitted to post a summary of their chapter only in their institution's repository.

19. **Thesis/Dissertation**: If your license is for use in a thesis/dissertation your thesis may be submitted to your institution in either print or electronic form. Should your thesis be published commercially, please reapply for permission. These requirements include permission for the Library and Archives of Canada to supply single copies, on demand, of the complete thesis and include permission for Proquest/UMI to supply single copies, on demand, of the complete thesis. Should your thesis be published commercially, please reapply for permission. Theses and dissertations which contain embedded PJAs as part of the formal submission can be posted publicly by the awarding institution with DOI links back to the formal publications on ScienceDirect.

Elsevier Open Access Terms and Conditions

You can publish open access with Elsevier in hundreds of open access journals or in nearly 2000 established subscription journals that support open access publishing. Permitted third party re-use of these open access articles is defined by the author's choice of Creative Commons user license. See our open access license policy for more information.

Terms & Conditions applicable to all Open Access articles published with Elsevier:

Any reuse of the article must not represent the author as endorsing the adaptation of the article nor should the article be modified in such a way as to damage the author's honour or reputation. If any changes have been made, such changes must be clearly indicated.

The author(s) must be appropriately credited and we ask that you include the end user license and a DOI link to the formal publication on ScienceDirect.

If any part of the material to be used (for example, figures) has appeared in our publication with credit or acknowledgement to another source it is the responsibility of the user to ensure their reuse complies with the terms and conditions determined by the rights holder.

Additional Terms & Conditions applicable to each Creative Commons user license:

CC BY: The CC-BY license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article and to make commercial use of the Article (including reuse and/or resale of the Article by commercial entities), provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by/4.0.

CC BY NC SA: The CC BY-NC-SA license allows users to copy, to create extracts, abstracts and new works from the Article, to alter and revise the Article, provided this is not done for commercial purposes, and that the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, indicates if changes were made and the licensor is not represented as endorsing the use made of the work. Further, any new works must be made available on the same conditions. The full details of the license are available at http://creativecommons.org/licenses/by-nc-sa/4.0.

CC BY NC ND: The CC BY-NC-ND license allows users to copy and distribute the Article, provided this is not done for commercial purposes and further does not permit distribution of the Article if it is changed or edited in any way, and provided the user gives appropriate credit (with a link to the formal publication through the relevant DOI), provides a link to the license, and that the licensor is not represented as endorsing the use made of the work. The full details of the license are available at http://creativecommons.org/licenses/by-nc-nd/4.0. Any commercial reuse of Open Access articles published with a CC BY NC SA or CC BY NC ND license requires permission from Elsevier and will be subject to a fee.

Commercial reuse includes:

- Associating advertising with the full text of the Article
- Charging fees for document delivery or access
- Article aggregation
- Systematic distribution via e-mail lists or share buttons

Posting or linking by commercial companies for use by customers of those companies.

20. **Other Conditions**:

v1.10

Questions? customercare@copyright.com or +1-855-239-3415 (toll free in the US) or +1-978-646-2777.