

Review

Chronic Kidney Disease Cohort Studies: A Guide to Metabolome Analyses

Ulla T. Schultheiss ^{1,2} , Robin Kosch ³ , Fruzsina Kotsis ^{1,2} , Michael Altenbuchinger ⁴
and Helena U. Zacharias ^{5,6,*}

- ¹ Institute of Genetic Epidemiology, Faculty of Medicine and Medical Center, University of Freiburg, 79106 Freiburg, Germany; ulla.schultheiss@uniklinik-freiburg.de (U.T.S.); fruzsina.kotsis@uniklinik-freiburg.de (F.K.)
 - ² Department of Medicine IV—Nephrology and Primary Care, Faculty of Medicine and Medical Center, University of Freiburg, 79106 Freiburg, Germany
 - ³ Computational Biology, University of Hohenheim, 70599 Stuttgart, Germany; robin.kosch@uni-hohenheim.de
 - ⁴ Institute of Medical Bioinformatics, University Medical Center Göttingen, 37077 Göttingen, Germany; michael.altenbuchinger@med.uni-goettingen.de
 - ⁵ Department of Internal Medicine I, University Medical Center Schleswig-Holstein, Campus Kiel, 24105 Kiel, Germany
 - ⁶ Institute of Clinical Molecular Biology, Kiel University and University Medical Center Schleswig-Holstein, Campus Kiel, 24105 Kiel, Germany
- * Correspondence: h.zacharias@ikmb.uni-kiel.de; Tel.: +49-(0)431-500-22274

Abstract: Kidney diseases still pose one of the biggest challenges for global health, and their heterogeneity and often high comorbidity load seriously hinders the unraveling of their underlying pathomechanisms and the delivery of optimal patient care. Metabolomics, the quantitative study of small organic compounds, called metabolites, in a biological specimen, is gaining more and more importance in nephrology research. Conducting a metabolomics study in human kidney disease cohorts, however, requires thorough knowledge about the key workflow steps: study planning, sample collection, metabolomics data acquisition and preprocessing, statistical/bioinformatics data analysis, and results interpretation within a biomedical context. This review provides a guide for future metabolomics studies in human kidney disease cohorts. We will offer an overview of important a priori considerations for metabolomics cohort studies, available analytical as well as statistical/bioinformatics data analysis techniques, and subsequent interpretation of metabolic findings. We will further point out potential research questions for metabolomics studies in the context of kidney diseases and summarize the main results and data availability of important studies already conducted in this field.

Keywords: metabolomics study design; nephrology; chronic kidney disease; human cohort studies; epidemiology; kidney disease etiologies



Citation: Schultheiss, U.T.; Kosch, R.; Kotsis, F.; Altenbuchinger, M.; Zacharias, H.U. Chronic Kidney Disease Cohort Studies: A Guide to Metabolome Analyses. *Metabolites* **2021**, *11*, 460. <https://doi.org/10.3390/metabo11070460>

Academic Editor: Brian Drew

Received: 15 June 2021

Accepted: 12 July 2021

Published: 16 July 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Chronic kidney disease (CKD) has become one of the major global health burdens in the 21st century [1], with a typically chronic progressive disease course. Its extremely heterogeneous disease pattern and comorbidity load complicates the understanding of the underlying pathomechanisms and optimal patient treatment. Cohort studies form a suitable study design to investigate the associations between multiple exposures on the one hand and multiple outcomes on the other hand. They are particularly appropriate to study rare exposures or exposures for which randomization is not possible due to practical or ethical reasons. Even though randomized controlled trials (RCTs) are the gold standard for a research question of the effect of an exposure on an outcome, the majority of interventions investigated by RCTs in nephrology have so far been unable to demonstrate treatment benefits or have even caused harm [2]. This may well be due

to the aforementioned heterogeneity of CKD populations. Large observational studies are therefore needed to appropriately characterize CKD population cohorts and identify interventionally treatable subgroups.

Important findings on kidney disease pathophysiology have already been accomplished by omics science, i.e., genomics [3,4], epigenomics [5], transcriptomics [6], and proteomics [7]. One of the latest additions to the omics research field is metabolomics, the quantitative study of small organic compounds, called metabolites, present in a biological specimen [8]. Metabolites are the intermediate and/or final products of molecular interactions between different proteins, signaling cascades, and cellular environments, thus constituting the end of the omics cascade. Additionally, they can arise from exogenous sources, including food and drug intake, cosmetics, gut microbe–host co-metabolism, and others. The observation, analysis, and interpretation of the metabolites' entirety, i.e., the metabolome, can therefore provide us with a metabolic “snapshot” or “fingerprint” of the current state of an organism. The ability of metabolomics studies to provide deeper insights into fundamental disease pathomechanisms has already been demonstrated for numerous other chronic diseases, including diabetes [9,10], cardiovascular diseases [11], and cancer [12]. Metabolomics is increasingly recognized as a valuable tool in the field of nephrology [13]. The first important metabolomics studies investigated metabolites as uremic toxins [14]. The research field then shifted towards the identification of metabolites associated with the patient's glomerular filtration rate (GFR) [15,16], to metabolic fingerprints of adverse patient events [17], and has now extended towards the understanding of the underlying mechanisms in CKD progression. Especially the latter two goals make prospective CKD cohorts with measurements of the important patient parameters, e.g., GFR, and metabolites available at multiple time points a prerequisite. A multitude of different study questions can be tackled by applying metabolomics and for each study question a fitting study design is required. Conducting extensive metabolomics studies in large-scale cohorts of CKD patients might therefore enable the elucidation of important, possibly causal molecular disease traits and, consequently, improve CKD patient treatment.

This review provides a guide for future metabolomics studies in kidney disease cohorts using observational study designs. Conducting a thorough investigation of the metabolic changes related to impaired kidney function requires sophisticated study planning, metabolomics data acquisition and statistical/bioinformatics data analysis, as well as interpretation of the findings (Figure 1).

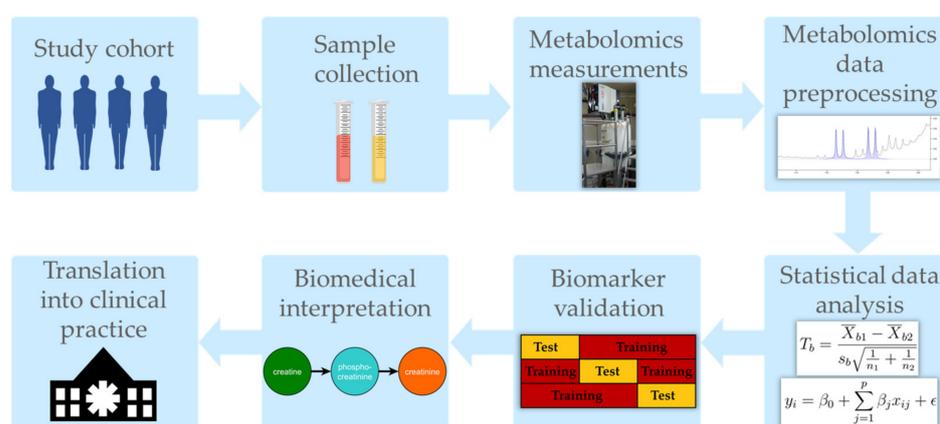


Figure 1. Schematic workflow of metabolomics studies in kidney disease cohorts.

Important a priori considerations for metabolomics cohort studies, the available analytical as well as statistical/bioinformatics data analysis techniques, and subsequent interpretation of metabolic findings will be given. We will further point out potential research questions for metabolomics CKD studies and summarize the main results of important metabolomics studies already conducted in this field. A comprehensive summary of the metabolic markers of CKD discussed throughout the text is given in Table 1.

Table 1. Important metabolite biomarkers of chronic kidney disease and its sequelae. Abbr.: AASK, African American Study of Kidney Disease and Hypertension study; ADPKD, autosomal dominant polycystic kidney disease; AKI, acute kidney injury; ARIC, Atherosclerosis Risk in Communities; ArMORR, Accelerated Mortality on Renal Replacement study; CKD, chronic kidney disease; CRIC, Chronic Renal Insufficiency Cohort; CTOT-04, Clinical Trials in Organ Transplantation 04 Study; CVD, cardiovascular disease; DKD, diabetic kidney disease; eGFR, estimated glomerular filtration rate; FSGS, focal segmental glomerulosclerosis; GCKD, German Chronic Kidney Disease; HDL, high-density lipoprotein; IDL, intermediate density lipoprotein; IgAN, IgA nephropathy; KORA, Cooperative Health Research in the Region of Augsburg; KRT, kidney replacement therapy; LDL, low-density lipoprotein; MDRD, Modification of Diet in Renal Disease; MESA, Multi-Ethnic Study of Atherosclerosis; mGFR, measured glomerular filtration rate; MGN, membranous glomerulonephritis; TCMR, T cell-mediated rejection; UACR, urinary albumin-to-creatinine ratio.

Study Design/Type	Study Question	Study Population and Investigated Biofluids	References	Detected Metabolites/Metabolic Biomarkers/Pathways
case-control study	AKI prediction	patients undergoing cardiac surgery, urine specimens collected before and after surgery	[18]	carnitine (elevated in AKI-free patients), tranexamic acid (elevated in AKI patients) and others
case-control study	AKI prediction	patients undergoing cardiac surgery, plasma specimens collected 24h after surgery	[19]	glucuronide conjugate of propofol, Mg ²⁺ , lactate and others
case-control study	indicators of AKI	hospitalized, newly diagnosed AKI patients, serum specimens	[20]	increases in acylcarnitines and amino acids and reduction of arginine and lysophosphatidylcholines in AKI patients
case-control study	distinct metabolic profile of ADPKD	54 patients with ADPKD, several control groups, urine specimens	[21]	on average 51 out of 701 NMR features could reliably discriminate ADPKD patients from other kidney disease patients and healthy controls
case-control study	non-invasive diagnosis of TCMR in pediatric kidney transplant recipients	pediatric kidney replacement recipients, urine specimens	[22]	proline, kynurenine, phosphatidylcholines, diacylglycerols elevated in TCMR patients
case-control study	identify metabolic pathways altered in CKD stage 3–4 non-diabetics	CKD patients from the Paricalcitol study; healthy controls: employees of study centers, urine and plasma specimens	[23]	27 urine and 33 plasma metabolites differed between CKD vs. controls; pathway analysis: citric acid cycle significantly affected: reduction of urinary excretion of citrate, cis-aconitate, isocitrate, 2-oxoglutarate, succinate; expression of genes regulating these metabolites were reduced
2 independent nested case-control studies (=analysis vs. replication cohort)	metabolites predicting CVD mortality in incident KRT patients	ArMORR study, plasma specimens	[24]	oleoylcarnitine, linoleoylcarnitine, palmitoylcarnitine, stearoylcarnitine, strongest association with CVD mortality: oleoylcarnitine

Table 1. Cont.

Study Design/Type	Study Question	Study Population and Investigated Biofluids	References	Detected Metabolites/Metabolic Biomarkers/Pathways
cross-sectional CKD study	plasma metabolite profile differences in CKD stages 2, 3, and 4	30 participants with differing CKD stages, plasma specimens	[25]	CKD stages 3 vs. 2: 62 differing metabolites (39 higher and 23 lower in CKD stage 3); CKD stages 4 vs. 2: 111 differing metabolites (66 higher and 45 lower in CKD stage 4); CKD stages 4 vs. 3: 11 differing metabolites (7 higher and 4 lower in CKD stage 4); major differences for higher CKD stages: altered arginine metabolism, elevated coagulation/inflammation, impaired carboxylate anion transport, decreased adrenal steroid hormone production
cross-sectional study (proof-of-concept study)	identification of serum metabolites to provide a more accurate GFR estimate	AASK study, MESA study: participants with mGFR, serum specimens	[26]	(1) serum metabolites from untargeted quantification: AASK—283 and MESA—387 significantly associated metabolites with mGFR; (2) targeted metabolites: 15 metabolites used for GFR estimation
2 cross-sectional observational studies of the general population	association of serum metabolites and their ratios with eGFR	KORA F4 study, TwinsUK registry, serum specimens	[15]	association with eGFR: 22 metabolites and 516 metabolite ratios; acylcarnitines were associated inversely, ratio with the lowest <i>p</i> -value: serine to glutaryl carnitine
differing study design per cohort	metabolites correlating with clinical markers of kidney disease	4 cohorts: training cohort, validation cohort, prospective cohort, drug treatment cohort	[27]	5 metabolites, e.g., 5-methoxytryptophan, correlate with markers of kidney function
nested case–control study	CKD progression	CRIC study, serum specimens	[14]	10 nominally associated metabolites; 6 higher in cases (uric acid, glucuronate, 4-hydroxy-mandelate, 3-methyladipate/pimelate, cytosine, homo-gentisate) and 4 lower in cases (threonine, methionine, phenylalanine, arginine)
prospective CKD cohort	risk of progression to KRT	GCKD study, plasma specimens	[17]	24 NMR features—highest weights: creatinine, high-density lipoprotein, valine, acetyl groups of glycoproteins, Ca ²⁺ -EDTA
prospective CKD cohort	urinary 6-bromotryptophan and incident ESKD	GCKD study, urine specimens	[28]	higher 6-bromotryptophan levels were associated with lower risk of ESKD

Table 1. Cont.

Study Design/Type	Study Question	Study Population and Investigated Biofluids	References	Detected Metabolites/Metabolic Biomarkers/Pathways
prospective CKD cohort	urine metabolites associated with adverse kidney outcomes and mortality	GCKD study, urine specimens	[29]	55 metabolites significantly associated with kidney failure, kidney failure + AKI or death; significant enrichment for phosphatidylcholine pathway
prospective CKD cohort	adverse cardiac events in CKD stage 3 patients	GCKD study, plasma specimens	[30]	association of trimethylamine <i>N</i> -oxide (TMAO) with cardiac arrhythmia and myocardial infarction
prospective CKD cohort, prospective population-based cohort	genetic studies of urinary metabolites	GCKD study, UK Biobank, urine specimens	[31]	240 unique metabolite-locus associations highlighting novel candidate substrates for transport proteins; genes identified are enriched in absorption, distribution, metabolism, and excretion (ADME) relevant tissues, potentially novel candidates for biotransformation and detoxification reactions
prospective diabetic cohort study	multimetabolite models of disease process from type 1 diabetic patients w/o CKD	Finnish Diabetic Nephropathy Study Group, serum specimens	[32]	cross-sectionally: patients w/o DKD complications: low lipids, less inflammation, better glycemic control vs. patients with advanced CKD: high sphingomyelin, cystatin-C; shared features: low unsaturated fatty acids (UFA), phospholipids; prospectively: progressive albuminuria: high UFAs, phospholipids, IDL, LDL; accelerated DKD progression: high saturated fatty acids, low HDL
prospective observational transplant recipient study	prediction of allograft status via urine metabolites	kidney graft recipients of the CTOT-04 study, urine specimens	[33]	best discrimination between acute cellular rejection vs. no rejection: ratio of urinary 3-sialyllactose to xanthosine
prospective population-based study	metabolite associations with eGFR; incident CKD	ARIC study, serum specimens	[16]	eGFR associations: 34 metabolites detected—strongest positive = creatinine, strongest negative = 3-indoxyl sulfate; lower risk of incident CKD: 5-oxoproline, 1,5-anhydroglucitol

Table 1. Cont.

Study Design/Type	Study Question	Study Population and Investigated Biofluids	References	Detected Metabolites/Metabolic Biomarkers/Pathways
prospective population-based study	kidney function decline, incident CKD	KORA S4/F4 study, serum specimens	[34]	kidney function decline: spermidine, phosphatidylcholine diacyl C42:5-to-phosphatidyl acyl-alkyl C36:0 ratio; incident CKD: kynerunine-to-tryptophan ratio
prospective population-based study; prospective twin cohort	metabolite association with eGFR, incident CKD	KORA F4 study, replication in TwinsUK registry, serum specimens	[35]	54 metabolites replicated and significantly associated with eGFR; 6 with pair-wise correlation with established kidney function measures (C-mannosyltryptophan, pseudouridine, <i>N</i> -acetylalanine, erythronate, myo-inositol, <i>N</i> -acetylcarnosine); incident CKD: C-mannosyltryptophan, pseudouridine, O-sulfo-L-tyrosine
prospective small patient sample	metabolic changes after kidney allograft transplantation	19 allograft recipients, serum specimens	[36]	hippurate, mannitol, and alanine associate with changes in transplant allograft function over time; hippurate/histidine are more sensitive to short-term changes in kidney activity than creatinine
two clinical trials	cross-sectional association of UACR with 637 known, non-drug, blood metabolites	AASK, MDRD study, serum specimens	[37]	58 metabolites associated with proteinuria; metabolites with lowest <i>p</i> -value: 4-hydroxychlorthalonil and 1,5-anhydroglucitol with all 6 metabolites of the phosphatidylethanolamine pathway being significant
review	DKD associated metabolites	multiple studies	[38]	early stages of DKD: association with tricarboxylic acid cycle, glucose metabolites; uremic toxins in DKD progression: phenyl sulfate and tryptophan derivatives

Table 1. Cont.

Study Design/Type	Study Question	Study Population and Investigated Biofluids	References	Detected Metabolites/Metabolic Biomarkers/Pathways
review	differential metabolites in MGN, FSGS, IgAN	multiple studies	[39]	amongst others—MGN: 13 urinary metabolites as most important (dopamine, fumarate, carnosine, nicotinamide D-ribonucleotide, pyridoxal, deoxyguanosine triphosphate, adenosine monophosphate, L-citrulline, nicotinamide, deoxyuridine, phenylalanine, tryptamine, succinate); FSGS: 10 prognostic urine metabolites (citrulline, proline, dimethylamine, acetoacetate, valine, alphaketoisovaleric acid, isobutyrate, histidine, D-palmitylcarnitine, N-methylnicotinamide); IgAN vs. controls: higher serum metabolite levels (phenylalanine, lactate, myo-Inositol, L6 lipids L5 lipids, L3 lipids) and lower serum metabolite levels (alpha-, beta-glucose, valine, phosphocholine, tyrosine, lysine, isoleucine, glycine, glycerolphosphocholine, glutamate, glutamine, alanine, acetate, 1-methylhistidine, 3-hydroxybutyrate)
perspectives, no study design	metabolomics in CKD research: metabolites and future risk of mortality	AASK study, serum specimens	[13]	number of associated metabolites reduced after adjustment for eGFR—metabolite classes detected: amino acid, carbohydrate, cofactors/vitamins, energy, lipid, nucleotide, peptide, xenobiotic, unknown

2. How to Get Started: A Priori Considerations for Metabolomics Cohort Studies

2.1. Possible Study Questions for Cohort Studies

Before getting started with a metabolomics study in the field of nephrology, the researcher has to determine a study question of interest. Several exemplary study questions can be found in Figure 2, where some of the asked questions have already been investigated in the past.

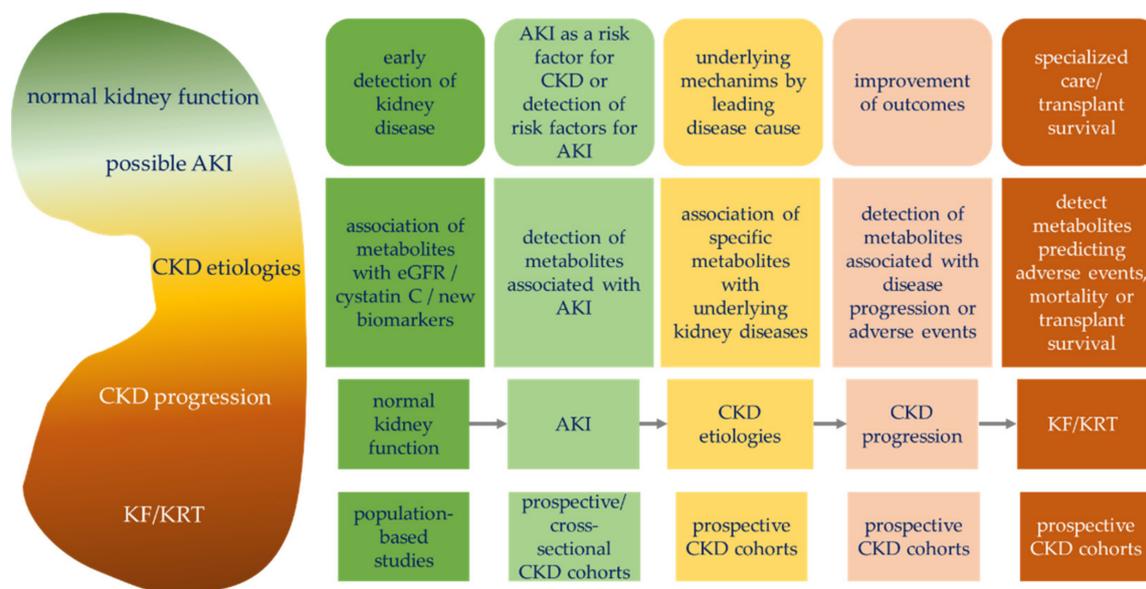


Figure 2. Possible research questions, kidney related prerequisites, and study cohort types used for metabolomics studies. Abbr.: AKI, acute kidney injury; CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; KF, kidney failure; KRT, kidney replacement therapy.

Since an eGFR decline only occurs after a patient has already lost a considerable amount of kidney function, an early detection of kidney disease is a relevant study question. However, it can only be tackled within a population-based study including participants prior to a manifested diagnosis of CKD. For example, Sekula et al. were able to apply a non-targeted metabolomics approach within a population-based study (KORA F4) by associating metabolites with eGFR and were able to replicate 54 significantly associated metabolites in an independent cohort (Twins UK) [35]. CKD patients are more prone to acute kidney injury (AKI) events, but to elucidate the mechanisms of AKI, these patients have to be monitored more tightly than what a usual study design for cohort studies will request. So, nesting a sub-cohort of AKI patients within a prospective CKD cohort should be the way to go. Examples of small pilot studies with a limited number of participants to detect metabolites associated with and/or diagnostic of AKI can be found in [18–20]. Large prospective CKD cohorts offer the possibility to not only elucidate the general mechanisms of CKD, but to also delve deep into the causal pathways of differing CKD etiologies. Some examples of single, specific kidney diseases can be found in the literature, but within CKD cohorts larger patient populations with more power to detect the associations between the metabolites and kidney diseases can be collected. Recent advances in the metabolomics field have, for example, identified dysregulated energy metabolism between early- and late-stage diabetic kidney disease (DKD), a well-defined etiological CKD group, and elucidated the interaction between metabolic stress, mitochondrial homeostasis, and organelle crosstalk in the kidney as being important for dynamics during DKD progression [38]. Other examples can be found for membranous nephropathy, focal segmental glomerulosclerosis or IgA nephropathy [39], as well as autosomal dominant polycystic kidney disease (ADPKD) [21]. One of the big study questions is the identification of metabolites associated with or predicting CKD progression. Here, prospective CKD cohorts with measurements

of kidney function markers and metabolites at multiple time-points during the study's follow-up phase pose an invaluable treasure of information that only needs to be launched. Studies that have detected metabolites associated with CKD progression or with a higher risk of adverse patient events are, e.g., studies by Rhee et al. [14], Kalim et al. [24], or Zacharias et al. [17]. Rhee et al. identified metabolite alterations associated with subsequent disease progression, Kalim et al. showed that acylcarnitine may be associated with a higher uremic cardiovascular risk, and Zacharias et al. identified a multivariate metabolite signature for end-stage kidney disease (ESKD) risk prediction. In prospective ESKD cohorts, interesting research questions will then move towards the detection of metabolites to predict adverse events for hemodialysis patients or towards transplant survival/early detection of allograft rejection in kidney transplant patients. Some smaller studies with few participants investigating allograft rejection have been carried out in the past, e.g., studies by Blydt-Hansen et al. [22] and Suhre et al. [33]. These studies concluded the utility of metabolomics for non-invasive diagnosis of allograft rejection.

Besides prospective CKD cohort studies, other common study types can be applied in metabolomics studies of CKD, which are briefly discussed in the following paragraphs.

2.2. Common Study Designs in Human Cohorts

2.2.1. Case Reports and Case Series

Historically, case reports, focusing on a single subject, or case series, reporting on a small group of phenotypically similar subjects, are a first step in identifying a new disease or adverse health effect from an exposure [40]. The possible association between the observed outcome and a specific exposure is described based on a small group of subjects. Such studies may be the first in identifying the value of a new scientific approach to clarify the pathophysiological background of a known disease [41]. Studies evaluating metabolomics in CKD started out small. For example, Shah et al. investigated only 30 participants with CKD and were able to show differences in the metabolic profiles for various CKD stages, reflecting alterations in arginine metabolism, elevated coagulation/inflammation, impaired carboxylate anion transport, and decreased adrenal steroid hormone production [25]. This study was a proof-of-concept study, setting the stage for large-scale prospective cohort studies in metabolomics of CKD. Another important small-scale proof-of-concept study of metabolic biomarker detection was conducted by Gronwald et al. [21]. Based on urinary nuclear magnetic resonance (NMR) metabolic fingerprints, the authors were able to discriminate ADPKD patients with moderately advanced disease from ADPKD patients with ESKD, patients with CKD of other etiologies, and healthy controls.

2.2.2. Cross-Sectional Study

In a cross-sectional study, outcome and exposures are analyzed at the same time. In comparison to case-control (participants selected based on the outcome status) or cohort studies (participants selected based on the exposure status), the participants in a cross-sectional study are only selected based on the inclusion and exclusion criteria. This type of study design can be easily implemented, is rather cheap, and can be started at enrollment or any later time point during the course of a cohort study. Since the inference of causal relationships between exposure and outcome by a one-time measurement is not possible, these studies are traditionally used to investigate disease prevalence or the influence of environmental factors such as drugs, toxins, or diet, as, e.g., demonstrated in [42].

Cross-sectional studies can further be employed to identify subgroups or stages in complex diseases. Luo et al., for instance, identified 58 serum metabolites associated with proteinuria in a cross-sectional study design, some of which were also associated with CKD progression [37]. Within a proof-of-concept study setting, multivariate metabolite signatures of measured GFR were used to improve GFR estimation [26]. Goek et al. found the serum concentrations of spermidine to be associated with kidney function change in the general population, and serum metabolites were able to predict incident CKD [34].

Another study by Barrios et al. investigated metabolic signatures of diabetic nephropathy combining four European cohorts [43].

2.2.3. Case–Control Study

Within case–control studies, the odds of an exposure within a predefined group with a characteristic trait of interest are compared to the odds of an exposure in a control group. When appropriately designed, case–control studies (1) can provide the same information as a cohort study; (2) are more rapid and efficient, because, unlike in cohort studies, only a minority of the population is included in the study; and (3) data on exposure are being collected in retrospect. Disadvantages include: (1) ‘general types of bias’; (2) specific sources of bias; and (3) selection of cases and controls can prove to be complex [40]. A special form of case–control study is the nested design of a case–control study, where cases and controls are drawn from within a prospective study. All cases who developed the outcome of interest during the follow-up are selected and compared with a random sample of the cohort [41]. A nested case–control study of metabolomics in a CKD population was, for instance, carried out within the Chronic Renal Insufficiency Cohort (CRIC) Study by Rhee et al. in 2016 [14]. The authors compared a subset of the CRIC Study population with rapid progression of kidney disease according to eGFR slopes to a subset with slow progression. For each case a control was selected that was categorized within the same eGFR and proteinuria category at study entry. Ten metabolite alterations were nominally associated with subsequent CKD progression and, cross-sectionally, six of the metabolites that were higher in the cases than controls were significantly associated with eGFR at baseline. The authors concluded that their results warrant further interest in arginine, methionine, and threonine as potential markers of kidney function and progression of kidney disease.

2.2.4. Prospective Cohort Study

As already outlined above, prospective cohort studies collect consecutive information on outcomes and exposures from the same participants within a specific time period. They allow, e.g., time-to-event analyses, time-course evaluations, and risk score development based on metabolite measures. To date, only few prospective CKD studies acquired metabolomics data, amongst them the German Chronic Kidney Disease (GCKD) study [44]. Within this study, Zacharias et al. developed a novel risk score based on NMR-derived plasma metabolic features, including creatinine, high-density lipoprotein, valine, acetyl groups of glycoproteins, and Ca^{2+} -EDTA, to predict the risk of ESKD within four years after the metabolomics measurements [17]. In another metabolomics study conducted in the GCKD cohort, Sekula et al. discovered a significant association between higher urinary 6-bromotryptophan levels and lower risk of kidney failure, both unadjusted and adjusted for kidney failure risk factors other than eGFR [28]. Similar results were obtained for investigations of serum 6-bromotryptophan levels. Steinbrenner et al. discovered 55 urinary metabolites that predict adverse kidney outcomes and/or mortality, including C-glycosyltryptophan, within a metabolome-wide association study [29]. Within the same cohort, Schlosser et al. identified the underlying molecular mechanisms related to the absorption, distribution, metabolism, and excretion (ADME) of metabolites in the kidney [31]. Comprehensive summaries of recent ongoing prospective CKD cohort studies within the International Network of Chronic Kidney Disease cohort studies (iNET-CKD) can be found in [45]. Another example from the CRIC Study by Kwan et al. detected a negative association of 3-hydroxyisobutyrate and 3-methylcrotonylglycine with eGFR slopes whereas citric acid and aconitic acid were positively associated.

2.2.5. Randomized Controlled Trial

Since RCTs, when carried out appropriately, are still the gold-standard for studying the effects of an intervention or any other type of therapy on an outcome, RCTs in a CKD metabolomics context might likewise be of interest. The first steps into this direction have

been taken in animal studies as well as for other disease entities. Hypertension in CKD patients, one of the leading underlying CKD causes, is mostly treated via several available drugs, but another treatment strategy could be to metabolically rewire the hypertensive kidney. Rinschen et al. were able to show promising results in animal models, leading to possible future dietary intervention studies [46]. These kinds of studies would constitute a metabolic challenge that can be supervised by measuring metabolites before and after the intervention of, e.g., a lipid-consuming, ketogenic diet in comparison to controls. Similar studies have been carried out with dietary interventions in patients diagnosed with rheumatoid diseases [47], as well as diabetes [48]. In the latter, postprandial metabolic alterations in healthy men with a high genetic risk of diabetes were evaluated after two meals with a varying macronutrient content, finding that modifications in intermediate lipid metabolism were induced by a high caloric meal.

2.3. Important Considerations for Sample Collection in Metabolomics Studies

In human studies involving non-deceased study participants, common specimen types include plasma, serum, urine, whole blood, saliva, cerebrospinal fluid, feces, and tissue. Plasma, serum, and urine, in particular, constitute the most suitable biofluids analyzed by metabolomics in large-scale cohorts due to easy sample collection, handling, and preparation, as well as being in high abundance, and, most importantly, involved in key renal regulatory mechanisms. The analysis of kidney tissue is likewise important, but requires more complex, invasive sample collection and extensive sample preparation.

Sample collection should, ideally, be carried out in a standardized fashion according to well-defined standard operating procedures (SOPs) across the whole study period to minimize unwanted technical and biological sample variation. The metabolite content of a blood and urine specimen is significantly influenced by an individual's fasting status, fluid intake, circadian rhythm, age, sex, body fat composition, comorbidities, and specific lifestyles, including smoking and alcohol intake, as well as a plethora of different medications [49–51], but also by numerous genetic factors [52]. The collection of such biofluid specimens should thus be carried out within a homogeneous time window across the whole study cohort, ideally after a well-defined fasting state period. To appropriately account for non-influenceable confounders, e.g., sex or comorbidities, matching or randomization strategies should be applied, and/or confounder adjustment and stratification during the statistical analysis. Here, accurate documentation of all important phenotypical, but also study protocol information, such as sample collection time, is warranted. To avoid bacterial growth in freshly collected urine samples, appropriate preservation steps, i.e., either filtration, centrifugation, or addition of bacteriostatics, should be carried out [53]. In the case of plasma samples, the use of only one specific type of anticoagulant, e.g., ethylenediaminetetraacetic acid (EDTA), heparin, or citrate, across the whole sample cohort is strongly recommended, since substance traces can appear in metabolic fingerprints and might complicate further analysis steps [19,54]. Ideally, appropriate sample volumes should be immediately aliquoted for subsequent metabolomics measurements to avoid unnecessary freeze–thaw cycles. To ensure metabolite stability, samples should be frozen at $-80\text{ }^{\circ}\text{C}$ immediately after collection until further processing. Especially in the case of large-scale cohort studies comprising hundreds to thousands of individual specimens, automatic sample handling and documentation is strongly recommended. A comprehensive review, including the SOPs for optimal pre-analytical handling of, e.g., urine, plasma, serum, and tissue specimens for subsequent metabolomics measurements, is provided in [53].

3. Metabolomics Data Acquisition

3.1. Common Analytical Platforms in Metabolomics Studies

Two main analytical platforms are commonly used for metabolomics studies: nuclear magnetic resonance (NMR) spectroscopy and hyphenated mass spectrometry (MS). The principle of NMR spectroscopy is based on the separation of different analyte signals by their resonance frequencies within a magnetic field. It is particularly well suited for

large-scale metabolomics studies, since instrumentation and data acquisition is highly stable across time and even across different lab facilities [55]. Only few, rather cheap sample preparation steps are mandatory and no metabolite derivatization is needed. Due to its non-destructive nature, NMR experiments allow the re-use of sample material after measurement and instrument cleaning is not required. NMR spectroscopic data allows, theoretically, the absolute quantification of all detectable metabolites with the use of only one internal standard. However, NMR spectroscopy suffers, in comparison to hyphenated mass spectrometry, from low sensitivity, resulting in lower metabolite coverage. Due to typically limited time resources, only one-dimensional (1D) NMR experiments are carried out for large-scale cohort studies. 1D NMR spectra, especially of urine and plasma/serum specimens, exhibit a high number of spectrally overlapping metabolite signals, which might complicate subsequent metabolite identification and accurate quantification. Two-dimensional (2D) NMR experiments are able to resolve these strongly overlapping signals into a second dimension and can provide further structural information about the detected metabolites, enhancing metabolite identification. Significantly longer acquisition times for 2D NMR experiments, however, preclude their wide application for large-scale cohort studies, although recent progress in the development of, e.g., non-uniform sampling techniques for 2D NMR [56] might overcome this obstacle soon. Although NMR experimental costs are, in general, low, the initial set-up of a well-operating NMR spectroscopy platform suitable for high-throughput metabolomics measurements is expensive and specific site requirements have to be fulfilled. Commercial NMR metabolomics platforms have been established in recent years and have proven their reliability in numerous studies [57]. The latest instrumental and analytical developments include the miniaturization of NMR spectrometers to a “benchtop” size [58–60] and the introduction of Bruker IVDr methods [61,62].

In contrast, hyphenated mass spectrometry, such as liquid chromatography (LC) or gas chromatography (GC)–MS offer much higher sensitivity and selectivity. MS identifies metabolites according to their mass-to-charge-ratios. It is typically coupled to an LC or GC, which separate analytes according to different physical and chemical properties, e.g., molecular size, charge, polarity, and affinity toward other molecules [63]. In contrast to NMR, which requires about 100–400 μL volume per biofluid specimen [62,64], MS experiments are typically carried out with much lower sample volumes of about 10 μL . MS sample preparation usually includes a derivatization step and the addition of individual internal standards for each absolutely quantified metabolite. These sample preparation steps, but also specific sample introduction systems and ionization techniques can prevent the detection of certain metabolite classes [65]. MS techniques are per se destructive and samples cannot be recovered after measurement. However, due to the low sample volume required, this hardly ever constitutes a serious limitation for MS in human cohort studies. The initial installation of a hyphenated MS system is, in comparison to an NMR spectrometer platform, cheaper, and less elaborate site requirements have to be fulfilled. On the other hand, hyphenated MS systems are, in general, less robust, and therefore data are less reproducible than when acquired on NMR systems; MS systems also require regular instrument cleaning. It has to be noted that the metabolome coverage of NMR spectroscopy and hyphenated MS, although displaying very good overlap between the different techniques, still exhibits distinct differences [66], and these analytical platforms should be rather considered as complementary than competing. Instrumental improvements in hyphenated mass spectrometry include the introduction of comprehensive two-dimensional (2D) gas chromatography (GC \times GC), displaying superior separation capacity for complex biological mixtures, high sensitivity, peak resolution, and reproducibility [67]. Likewise, comprehensive 2D LC \times LC substantially reduces peak overlap [68]. Imaging mass spectrometry (IMS) enables the *in vivo* or *in vitro* detection and 2D or 3D imaging of metabolites in tissues or cells and thus provides additional spatial information about metabolite distributions in these specimens [69].

Irrespective of the employed analytical platform, metabolomics analyses can be conducted in two different approaches: targeted or untargeted metabolomics. Targeted metabolomics constitutes the accurate detection and often absolute quantification of a preselected set of known metabolites. Commercially optimized kits for high-throughput quantitative analysis are readily available and several commercial contract research organizations offer targeted and/or untargeted metabolomics measurement services [70]. Such targeted MS protocols allow high-throughput measurements with excellent reproducibility. Non-targeted metabolomics, in contrast, aims at maximization of metabolome coverage without any a priori metabolite selection, i.e., hypothesis free. Analyte signals of interest, typically revealed by statistical data analysis, are then identified post hoc. For MS techniques, non-targeted metabolomics is only able to provide semi-quantitative metabolite measures since individual internal standards are naturally missing. NMR spectroscopy, however, still allows a posteriori absolute metabolite quantification after accurate identification of previously unknown metabolites measured in an untargeted approach. The choice of metabolomics approach for a nephrological study should be based on the particular research question: if the study aims at elucidating the role of one or several distinct metabolites in a phenotype, which are known based on previous research, a targeted, hypothesis-driven approach is recommended. If the study aims at uncovering yet unknown metabolic key players involved in a specific phenotype, an untargeted, hypothesis-generating approach should be employed.

3.2. Sample Preparation, Measurements, and Preprocessing in Metabolomics Studies

Numerous comprehensive protocols for both NMR and hyphenated MS, including elaborate sample preparation, measurement, and data preprocessing workflows, are available [8,63,64,71–74]. A selective summary of the key metabolomics data preprocessing steps, including the available software tools, is provided in Table 2. In brief, sample preparation for NMR-based metabolomics studies includes the addition of buffer solution, D₂O, and a spectral reference substance, such as 3-trimethylsilyl-2,2,3,3-tetradeuteropropionate (TSP), to the respective urine, plasma, serum, or tissue extract specimens [64]. Please note that the protein present in the specimen, as, for instance, in plasma or urine of patients suffering from proteinuria, gives rise to broad, unspecific NMR signals, which might obscure smaller metabolite signals, and severely binds to the reference substance TSP [64]. In this case, TSP can no longer be used as a reference for absolute quantification, and other reference substances, e.g., formic acid, have to be employed [75]. Alternatively, proteins can be removed prior to metabolomics data acquisition by, e.g., ultrafiltration or chemical protein precipitation [76], or a specific NMR pulse sequence, the Carr–Purcell–Meiboom–Gill (CPMG) sequence can be employed to suppress broad protein signals [75]. Likewise, suitable water suppression techniques are typically employed during NMR data acquisition for urine, plasma, and serum specimens to reduce the dominance of the strong water signals in the spectra [70]. Preprocessing of raw NMR data includes Fourier transformation of the NMR signal with the application of an exponential filter function, as well as phase and baseline correction [64]. To facilitate statistical evaluation of NMR spectra, the corresponding NMR signals need to be extracted beforehand. Various NMR signal extraction methods have been proposed, but a simple binning of the complete spectrum into equidistant sections of, e.g., 0.01 ppm width, is probably still the most popular technique [8]. Any metabolomics dataset is affected by unwanted technical and/or biological variances and biases, such as varying dilution of urine specimens [8]. These variances can be reduced by appropriate data normalization techniques, but subsequent statistical analysis results are inherently dependent on the specific, a priori chosen method [8,77]. To overcome this issue, Zacharias et al. proposed the application of (logistic) zero-sum regression [78,79] for the generation of normalization-invariant multivariate metabolic biomarker signatures, which proved to yield highly robust and predictive metabolic biomarker signatures of AKI after cardiac surgery [77]. Besides data normalization, both NMR and MS metabolomics data are typically transformed to approximately follow a multivariate normal distribution and

to exhibit constant variance, e.g., by application of a log transformation [8]. In the case of untargeted NMR metabolomics studies, statistical analysis steps are typically carried out with yet unidentified NMR spectral features, and subsequent metabolite identification only focuses on statistically relevant NMR peaks. This identification is achieved by manual comparison of the complex NMR spectrum of a biofluid specimen to the NMR reference spectra of pure compounds, available from either commercial or public data bases, e.g., the Human Metabolome Data Base (HMDB) [64,80]. Here, additional 2D NMR measurements can offer extremely valuable structural information to support this identification step. Following successful identification, these metabolites can then be absolutely quantified [64].

For LC-MS measurements, removal of protein during the sample preparation step, e.g., by methanol extraction, is mandatory to avoid signal suppression of the low-abundance analytes and protein precipitation under reversed-phase LC conditions [70]. The analysis of urine by LC-MS techniques is challenged by the high salt content, varying dilution, and the complex composition of the samples [70]. Various analytical pretreatment and data normalization strategies have been proposed to overcome these issues [81]. A systematic comparison of different protocols by Vogl et al. revealed that dilution of urine specimens to a fixed creatinine concentration yielded the least number of missing values and allowed reliable classification of urine specimens from healthy controls and CKD patients [81]. The urinary creatinine concentration is, however, significantly influenced by sex, age, muscle mass, diet, pregnancy, and renal pathology [82–85]. Alternative normalization approaches have thus been proposed, including a normalization to the urine volume, osmolality, and “total useful MS signal” [86,87]. Since, however, the choice of MS data normalization strategies, as described above in an analogous manner for NMR data normalization strategies, substantially influences subsequent statistical data analysis results [86], it is recommended to either employ a combination of different normalization strategies [86,87] or normalization-invariant data analysis methods, such as zero-sum regression [77,78]. The application of GC-MS for metabolomics analyses requires the volatilization of the analyzed compounds, which have to be thermally stable. Subsequent data preprocessing steps again include feature extraction and (automatic) identification, based on commercial or freely available databases, as well as absolute quantification of metabolites. One should keep in mind that the latter is only possible if the corresponding internal standard had been included in the measurement step [70]. Both untargeted and targeted MS datasets include certain amounts of missing data points due to failed peak detection, leading to incomplete data matrices. Since many statistical data analysis methods, however, require complete data matrices, these missing data points are typically imputed prior to statistical analysis. A combination of NMR and hyphenated MS experiments can significantly enhance the metabolite identification in untargeted metabolomics studies.

Table 2. Important preprocessing steps in targeted and/or untargeted metabolomics studies and selected commercial or freely available preprocessing software. Abbr.: JBA, pJRES binning algorithm; kNN, k-nearest neighbors; MCR-ALS, Multivariate Curve Resolution-Alternating Least Squares; MICE, multivariate imputation by chained equations; MS, mass spectrometry; NMR, nuclear magnetic resonance; RF, random forest; ROI, region of interest; SRV, statistical recoupling of variables.

Preprocessing Step	Goal		Available Methods		Commercially Available Software		Freely Available Software	
	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS
spectral preprocessing	transform spectral data from time to frequency domain, correct baseline and phase distortions	reproducible identification and quantification of peak features across multiple MS spectra	Fourier transformation, zero filling, apodization, phase correction, baseline correction, spectral alignment, removal of unwanted regions	deisotoping, retention time alignment, baseline and noise filtering, recalibration	TopSpin (BrukerBioSpin GmbH, Rheinstetten, Germany), AMIX (BrukerBioSpin GmbH, Rheinstetten, Germany), ACD (ACD labs)	ACD (ACD labs), AMIX (BrukerBioSpin GmbH, Rheinstetten, Germany), vendor-specific software, Mnova	Automics (Softpedia), NMRfX, NMRPipe [88], BAYESIL [89], R-package AlpsNMR [90], R-package speaq [91]	ChromA [92], Chromaligner [93], MetAlign [94], MZmine [95,96], MZmine 2 [97], OpenMS [98], XCMS [99], XCMS ² [100], MAVEN [101], eRah [102]
metabolic feature extraction	extract signal intensities in untargeted manner from spectra to perform subsequent statistical analysis, reduce dimensionality, minimize effects from peak position variations across different spectra		equidistant bucketing/binning, Gaussian binning [103], adaptive binning [104], intelligent binning [105], dynamic adaptive binning [106], SRV [107], JBA [108], peak picking, manual/automatic definition of ROIs	equidistant bucketing/binning, peak detection/picking, manual/automatic definition of ROIs	AMIX (BrukerBioSpin GmbH, Rheinstetten, Germany), Chenomx (Chenomx Inc. Edmonton, Canada) [109]	vendor-specific software	R-package mQTL [110], R-package MWASTools [111], R-package speaq [91], R-package speaq 2.0 [112], R-package AlpsNMR [90]	MetaboAnalyst [113], MZmine [95,96], MZmine 2 [97], XCMS [99], MetAlign [94], MAVEN [101], MSClust [114], ROIMCR [115]

Table 2. Cont.

Preprocessing Step	Goal		Available Methods		Commercially Available Software		Freely Available Software	
	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS
spectral deconvolution	deconvolute highly overlapping peak areas		curve fitting	MCR-ALS [116]	Chenomx (Chenomx Inc. Edmonton, Canada) [109]	vendor-specific software	BATMAN [117,118], decon1d [119], MetaboDecon1D [120], BAYESIL [89], non-linear peak fitting based on Voigt line shape model [121]	MetSign [122], DecoMetDIA [123], eRah [102]
missing value imputation	–	impute missing values to obtain full data matrix	–	half minimum imputation, mean value imputation, zero imputation, median value imputation, RF [124], MICE, kNN	–	vendor-specific software	–	MZmine [95,96], MetaboAnalyst [113], eRah [102], R-package mice [125], R-package VIM [126], R-package randomForest [127]
metabolite identification	identify metabolites in measured spectra		compare spectral features against reference spectra of pure compounds and/or query databases		Chenomx (Chenomx Inc. Edmonton, Canada) [109], AMIX (BrukerBioSpin GmbH, Rheinstetten, Germany) with BBIREFCODE database, Aldrich FT-NMR (Sigma-Aldrich)	vendor-specific software	COLMAR [128], KnowItAll Metabolomics (BioRad Corp.), MetaboHunter [129], MetaboMiner [130], BAYESIL [89], ASICS [131], R-package speaq 2.0 [112]	MZmine 2 [97], OpenMS [98], XCMS [99], XCMS ² [100], MZedDB [132], eRah [102]

Table 2. Cont.

Preprocessing Step	Goal		Available Methods		Commercially Available Software		Freely Available Software	
	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS	NMR Spectroscopy	Hyphenated MS
metabolite quantification	determine absolutely quantified concentrations of identified metabolites		accurately determine area under the curve of metabolite signal and reference with respect to known concentration of internal standard		Chenomx (Chenomx Inc. Edmonton, Canada) [109], AMIX (BrukerBioSpin GmbH, Rheinstetten, Germany)	vendor-specific software	BATMAN [117], [118], MetaboQuant [133], BAYESIL [89], AQuA [134], ASICS [131]	OpenMS [98]
metabolite data transformation	scaling of data in order to reduce data heteroscedasticity		e.g., log-transformation, variance stabilization transformation [135], auto-scaling, pareto scaling [136], mean centering		R Base, R-package vsn [137], R-package speaq 2.0 [112], Normalyzer [138], MetaPre [139]			
metabolite data normalization	minimize unwanted biological and/or technical variation between samples		e.g., creatinine normalization (for urine specimens), total spectral area normalization, normalization to internal standard, probabilistic quotient normalization [140], variance stabilization normalization [137], osmolality normalization, sample-specific normalization factors (e.g., volume), alternative: normalization-invariant zero-sum regression [77,78]		AMIX (BrukerBioSpin GmbH, Rheinstetten, Germany)	vendor-specific software	MetaboAnalyst [113], R-package AlpsNMR [90], R-package speaq 2.0 [112], Normalyzer [138], MetaPre [139], R-package zeroSum [77,78]	

4. Statistics and Bioinformatics Data Analysis

Probably the main goal of metabolomics analyses in biomedical research is the detection of powerful metabolic biomarkers for disease diagnosis or prognosis, response to therapeutic interventions, or, in general, response to external stimuli, e.g., nutrition or exercise. The NIH defines the term “biomarker” as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [141]. The search for novel metabolic biomarkers in the context of nephrology, as illustrated in Figure 2, is a highly emerging research area. From a statistical point of view, several different approaches can be distinguished (Table 3):

- **Hypothesis testing:** Univariate statistical differentiation between two or more predefined groups.
- **Multivariate biomarker signature detection:** Generation of multivariate regression scores to predict an outcome of an unknown test sample.
- **Subgroup identification:** Exploratory approach to identify biomedically different patient/sample subgroups.
- **Metabolome-wide association study:** Systematic analysis of the entire measured metabolome based on regression, including appropriate confounder adjustment to identify significant associations between metabolites and an outcome. A correction for multiple testing is essential for these comparisons.
- **Statistical network analysis:** Systematic analysis of interactions between different metabolites and/or patient parameters, other omics variables, etc., which are represented as a network. Allows a holistic view on the metabolome and its interaction with specific phenotypes, and can reveal molecular mechanisms or regulating processes.
- **Meta-analysis:** Combination of statistical results across multiple studies to increase statistical power and to gain more robust results.
- **Time-to-event analysis:** Time-to-event data contain information about if and when an event occurred, but typically also censored data. Survival analysis appropriately associates time-to-event data with, e.g., metabolite levels.
- **Time-course analysis:** Analysis of metabolite concentration changes across time and typically in response to external stimuli.
- **Pathway (enrichment) analysis:** Post-hoc mapping of differential metabolites to metabolic pathways, employing pathway databases, e.g., KEGG [142] or Gene Ontology [143], and subsequent testing if significantly differentiating metabolites are significantly enriched in a specific pathway.

The statistical analysis of high-dimensional metabolite data often includes multiple comparisons, which easily can result in a high number of false positives. To reduce this error and avoid misleading conclusions, the p -values have to be corrected for multiple testing by, e.g., adjustment of the false-discovery rate (FDR) as proposed by Benjamini and Hochberg [144]. Since the statistical analysis of metabolomics data requires a broad range of different methods, alongside the popular statistical analysis software *R* [145], several stand-alone software solutions exist, providing a collection of web-based tools with graphical user interfaces, e.g., MetaboAnalyst [113] and 3Omics [146], as reviewed by Cambiaghi et al. [147].

Table 3. Overview of different statistical research goals and corresponding statistics/bioinformatics tools, including freely available R software packages, for metabolomics data analysis. Abbr.: ANOVA, analysis of variance; ASCA, analysis of variance—simultaneous component analysis; GGM, gaussian graphical model; LASSO, least absolute shrinkage and selection operator; MGM, mixed graphical model; OPLS-DA, orthogonal projections to latent structures—discriminant analysis; ORA, over-representation analysis; PH, proportional hazards; PLS-DA, partial least squares—discriminant analysis; WGCNA; weighted gene co-expression network analysis.

Research Goal	Example	Literature Example	Common Statistics/ Bioinformatics Method	Popular Statistics/ Bioinformatics Tools	R Software Packages	Further Reading
hypothesis testing	compare metabolite levels in CKD patients and healthy controls	[18]	hypothesis testing	Student's <i>t</i> -test, ANOVA	>R Base: <i>t</i> .test, R Base: anova	[8,64]
multivariate bio-marker signature detection	multivariate metabolite signature to classify AKI vs. non-AKI patients	[19]	multivariate classification or linear regression	PLS-DA [148], OPLS-DA [149], support vector machine [150], Random Forest [124], LASSO regression [151], ridge regression [152], elastic net [153]	mixOmics [154], ropIs [155], e1071 [156], randomForest [127], glmnet [157]	[8,64,158]
subgroup identification	exploratory identify CKD patient subgroups with different survival outcomes based on metabolic profiles	[32]	supervised/unsupervised machine learning	PCA [159], Hierarchical Clustering, Self-organizing maps [160]	R Base: prcomp, ropIs [155], R Base: hclust, kohonen [161]	[8,64,160,162,163]
metabolome-wide association study	associations between all measured metabolites and eGFR, adjusted for age and sex	[35]	univariate/multivariate regression analysis (with confounder adjustment)	linear/logistic/Cox PH regression analysis	MWASTools [111]	[164]
statistical network analysis	exploratory identification of metabolite-metabolite associations	[30]	probabilistic graphical modeling, correlation networks	correlation network analysis, WGCNA [165], GGM [166], MGM [166]	corr, WGCNA [167], GeneNet [168], mgm [169]	[166,170,171]
meta-analysis	combining <i>p</i> -values for creatinine and eGFR metabolite associations across multiple studies	[35]	regression model	fixed-effects model	metafor [172], meta [173]	[174]

Table 3. Cont.

Research Goal	Example	LiteratureExample	Common Statistics/ Bioinformatics Method	Popular Statistics/ Bioinformatics Tools	R Software Packages	Further Reading
time-to-event analysis	estimate the mortality of CKD patients based on a set of metabolites	[17]	survival analysis	Cox PH regression analysis [175], LASSO Cox PH regression [151], random survival forest [176]	survival [177], glmnet [157], randomForestSRC [176]	[177–179]
time-course analysis	analyze metabolite intensity changes over time under different CKD treatment conditions	[36]	time-course analysis	ASCA [180,181]	MetStaT [182], DESeq2 [183]	[184]
pathway (enrichment) analysis	identify set of metabolites differentiating non-CKD and CKD patients with affiliation to a specific pathway	[23]	hypergeometric test, regression model	MSEA [185], ORA, global test [186]	FELLA [187], Lilikoi [188], globaltest [186]	[171,189]

5. Validation, Interpretation, and Beyond

Any statistical results of metabolomics studies have to be carefully validated. Ideally, novel statistically significant metabolic biomarkers should prove to be still significantly associated with the respective outcome in independent cohorts. Since metabolomics data in large-scale human CKD cohorts are still scarce, possibilities to replicate, e.g., a significant association between a set of metabolites and time-to-kidney-failure, are rather limited. Likewise, the predictive performance of novel metabolic risk scores for the diagnosis or prognosis of specific renal outcomes has to be validated on independent test sets to proof any clinical utility. Luckily, several strategies for unbiased performance assessment of novel classification or prediction scores within the same patient cohort are well established, e.g., cross-validation, where the complete data set is iteratively split into training and test data sets and, within each cross-validation run, the new predictive model is solely trained on the training set and solely tested on the test set, respectively (compare to Figure 1) [190].

Next to statistical replication, further exploration of metabolic findings, discovered in human cohorts, can be carried out in suitable animal models or cell lines, and vice versa. Animal and cell line studies offer the huge advantage of a very controlled experimental environment and suitability for extensive intervention studies, and they are able to further elucidate the underlying pathophysiologic mechanisms [191]. Chen et al., for example, identified 5-methoxytryptophan (5-MTP) as a potential marker of CKD in a human cohort, and subsequently examined the anti-inflammatory and anti-fibrotic effects of 5-MTP and the biological roles of its regulatory enzyme tryptophan hydroxylase-1 in cell and animal models [27]. A next step typically carried out in metabolomics studies is the interpretation of metabolic biomarkers in the context of their metabolic pathway environment, also known as pathway mapping, as well as in the context of already published research. Numerous open-source software for pathway mapping exist, including MetaboAnalyst [113], as extensively reviewed in [192]. The explosively growing amount of metabolomics data from many small studies and different analytical platforms, however, challenges the unified interpretation of metabolic findings across different studies. Abbiss et al. provide an extensive list of metabolites that have been reported as important for two or more kidney diseases [193].

The interplay of the microbiome and the metabolome in terms of the gut–kidney axis and its contribution to kidney diseases is reviewed in [194] and might help to highlight common biochemical processes in kidney diseases, such as the purine and tryptophan metabolism. The Human Metabolome Database (HMDB; <https://hmdb.ca/> (accessed on 10 June 2021)) offers a rich source of information on metabolites, their chemical properties, normal and abnormal abundances, biochemical/enzymatic/pathway data, as well as important literature references [80]. Kidney-specific web resources for different omics data, including Nephroseq (<https://www.nephroseq.org/> (accessed on 10 June 2021)), the Kidney and Urinary Pathway Knowledge Base (KUPKB; www.kupkb.org (accessed on 10 June 2021)) [195], and the Chronic Kidney Disease database (CKDdb; www.padb.org/ckdbd (accessed on 10 June 2021)) [196], which allow the unification of all available information from different sample origins and omics levels, are reviewed by [197].

While metabolomics studies represent a fascinating research field with huge potential that still needs to be launched on its own, an integration of multiple omics datasets will further help to elucidate CKD pathomechanisms. Multi-omics studies will make use of genome-wide association study (GWAS) data, whole exome or whole genome sequencing from DNA, messenger RNA (mRNA) as the product of gene transcription, as well as proteomics and metabolomics from the same patient or even the same sample. Genomic analyses can identify the risk factors/disease causing variants and can thereby enlighten regulatory networks. Together with proteomics and metabolomics measurements, one will be able to delve deeper into a functional/molecular basis of disease pathology [198]. Network analyses exploring the interconnectivity of genetic and molecular entities in CKD will provide additional information on the critical drivers of kidney diseases. Moreover, these networks will expand our understanding of how CKD affects different body systems

and how stimuli, such as diet, medication, and the microbiome, participate in this complex interplay [199]. Together, this will bring the field of omics research closer to possible clinical applications in order to improve patient treatment. Metabolomics especially has great potential for large-scale utilization in clinical practice; however, its current application in clinical routines is still limited. Current obstacles, which have to be resolved, include the development of small-scale measurement devices; extensive validation in external cohorts; introduction of SOPs for sample collection, storage, preparation, measurement, and preprocessing; data analysis and interpretation; and unambiguous metabolite identification as a key prerequisite for the development of targeted measurement kits [200,201].

6. Conclusions

The field of metabolomics already has been of unmeasurable value for nephrology research. Still, many questions remain and need to be addressed in the future. A first issue will be to understand the differing metabolite patterns across the diverse spectrum of kidney diseases, such as metabolic syndrome/diabetes mellitus, glomerular diseases, and many others; but, within similar phenotypic CKD etiologies, metabolomics also will help to unravel the mechanisms that differentiate, e.g., slow from fast CKD progressors. Translation of metabolomics research into routine CKD patient care will pave the way for novel metabolic biomarkers to evaluate and monitor the efficacy or safety of patient treatments. Thus, metabolomics studies will support clinical decision making. Eventually, metabolomics will become an integrated part of CKD diagnostics and will be able to inform the treating physicians on the rate of CKD progression, adverse risk evaluation, and other CKD-related comorbidities, such as the stage of metabolic syndrome vs. diabetes mellitus or others. Thereby, metabolomics will be a pioneering field for individualized patient treatment.

Funding: This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (grants no. 01ZX1912A, 01ZX1912B, and 01ZX1912C).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eckardt, K.-U.; Coresh, J.; Devuyst, O.; Johnson, R.J.; Köttgen, A.; Levey, A.S.; Levin, A. Evolving importance of kidney disease: From subspecialty to global health burden. *Lancet* **2013**, *382*, 158–169. [[CrossRef](#)]
2. Levin, A.; Lancashire, W.; Fassett, R.G. Targets, trends, excesses, and deficiencies: Refocusing clinical investigation to improve patient outcomes. *Kidney Int.* **2013**, *83*, 1001–1009. [[CrossRef](#)] [[PubMed](#)]
3. Wuttke, M.; Li, Y.; Li, M.; Sieber, K.B.; Feitosa, M.F.; Gorski, M.; Tin, A.; Wang, L.; Chu, A.Y.; et al.; Lifelines Cohort Study. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat. Genet.* **2019**, *51*, 957–972. [[CrossRef](#)] [[PubMed](#)]
4. Chambers, J.C.; Zhang, W.; Lord, G.; Van Der Harst, P.; Lawlor, D.A.; Sehmi, J.S.; Gale, D.; Wass, M.; Ahmadi, K.R.; Bakker, S.J.L.; et al. Genetic loci influencing kidney function and chronic kidney disease. *Nat. Genet.* **2010**, *42*, 373–375. [[CrossRef](#)]
5. Wilson, P.C.; Ledru, N.; Humphreys, B.D. Epigenomics and the kidney. *Curr. Opin. Nephrol. Hypertens.* **2020**, *29*, 280–285. [[CrossRef](#)]
6. Fan, Y.; Yi, Z.; D’Agati, V.D.; Sun, Z.; Zhong, F.; Zhang, W.; Wen, J.; Zhou, T.; Li, Z.; He, L.; et al. Comparison of Kidney Transcriptomic Profiles of Early and Advanced Diabetic Nephropathy Reveals Potential New Mechanisms for Disease Progression. *Diabetes* **2019**, *68*, 2301–2314. [[CrossRef](#)] [[PubMed](#)]
7. Assmann, N.; Dettmer, K.; Simbuerger, J.M.B.; Broeker, C.; Nuernberger, N.; Renner, K.; Courtneidge, H.; Klootwijk, E.D.; Duerkop, A.; Hallet, A.; et al. Renal Fanconi Syndrome Is Caused by a Mistargeting-Based Mitochondriopathy. *Cell Rep.* **2016**, *15*, 1423–1429. [[CrossRef](#)] [[PubMed](#)]
8. Zacharias, H.U.; Altenbuchinger, M.; Gronwald, W. Statistical Analysis of NMR Metabolic Fingerprints: Established Methods and Recent Advances. *Metabolites* **2018**, *8*, 47. [[CrossRef](#)] [[PubMed](#)]
9. Tofte, N.; Vogelzangs, N.; Mook-Kanamori, D.; Brahimaj, A.; Nano, J.; Ahmadizar, F.; Van Dijk, K.W.; Frimodt-Møller, M.; Arts, I.; Beulens, J.W.J.; et al. Plasma Metabolomics Identifies Markers of Impaired Renal Function: A Meta-analysis of 3089 Persons with Type 2 Diabetes. *J. Clin. Endocrinol. Metab.* **2020**, *105*, 2275–2287. [[CrossRef](#)] [[PubMed](#)]

10. Chen, Z.-Z.; Liu, J.; Morningstar, J.; Heckman-Stoddard, B.; Lee, C.G.; Dagogo-Jack, S.; Ferguson, J.F.; Hamman, R.F.; Knowler, W.C.; Mather, K.J.; et al. Metabolite Profiles of Incident Diabetes and Heterogeneity of Treatment Effect in the Diabetes Prevention Program. *Diabetes* **2019**, *68*, 2337–2349. [[CrossRef](#)]
11. Paynter, N.P.; Balasubramanian, R.; Giulianini, F.; Wang, D.D.; Tinker, L.F.; Gopal, S.; Deik, A.A.; Bullock, K.; Pierce, K.A.; Scott, J.; et al. Metabolic Predictors of Incident Coronary Heart Disease in Women. *Circulation* **2018**, *137*, 841–853. [[CrossRef](#)] [[PubMed](#)]
12. Lima, A.; Pinto, J.; Amaro, F.; Bastos, M.; Carvalho, M.; de Pinho, P.G. Advances and Perspectives in Prostate Cancer Biomarker Discovery in the Last 5 Years through Tissue and Urine Metabolomics. *Metabolites* **2021**, *11*, 181. [[CrossRef](#)]
13. Grams, M.E.; Shafi, T.; Rhee, E.P. Metabolomics Research in Chronic Kidney Disease. *J. Am. Soc. Nephrol.* **2018**, *29*, 1588–1590. [[CrossRef](#)]
14. Rhee, E.; Clish, C.; Wenger, J.; Roy, J.; Elmariah, S.; Pierce, K.; Bullock, K.; Anderson, A.; Gerszten, R.; Feldman, H. Metabolomics of Chronic Kidney Disease Progression: A Case-Control Analysis in the Chronic Renal Insufficiency Cohort Study. *Am. J. Nephrol.* **2016**, *43*, 366–374. [[CrossRef](#)]
15. Goek, O.-N.; Döring, A.; Gieger, C.; Heier, M.; Koenig, W.; Prehn, C.; Römisch-Margl, W.; Wang-Sattler, R.; Illig, T.; Suhre, K.; et al. Serum Metabolite Concentrations and Decreased GFR in the General Population. *Am. J. Kidney Dis.* **2012**, *60*, 197–206. [[CrossRef](#)] [[PubMed](#)]
16. Yu, B.; Zheng, Y.; Nettleton, J.A.; Alexander, D.; Coresh, J.; Boerwinkle, E. Serum Metabolomic Profiling and Incident CKD among African Americans. *Clin. J. Am. Soc. Nephrol.* **2014**, *9*, 1410–1417. [[CrossRef](#)] [[PubMed](#)]
17. Zacharias, H.U.; Altenbuchinger, M.; Schultheiss, U.T.; Samol, C.; Kotsis, F.; Poguntke, I.; Sekula, P.; Krumsiek, J.; Köttgen, A.; Spang, R.; et al. A Novel Metabolic Signature to Predict the Requirement of Dialysis or Renal Transplantation in Patients with Chronic Kidney Disease. *J. Proteome Res.* **2019**, *18*, 1796–1805. [[CrossRef](#)]
18. Zacharias, H.U.; Schley, G.; Hochrein, J.; Klein, M.S.; Köberle, C.; Eckardt, K.-U.; Willam, C.; Oefner, P.J.; Gronwald, W. Analysis of human urine reveals metabolic changes related to the development of acute kidney injury following cardiac surgery. *Metabolomics* **2013**, *9*, 697–707. [[CrossRef](#)]
19. Zacharias, H.U.; Hochrein, J.; Vogl, F.C.; Schley, G.; Mayer, F.; Jeleazcov, C.; Eckardt, K.-U.; Willam, C.; Oefner, P.J.; Gronwald, W. Identification of Plasma Metabolites Prognostic of Acute Kidney Injury after Cardiac Surgery with Cardiopulmonary Bypass. *J. Proteome Res.* **2015**, *14*, 2897–2905. [[CrossRef](#)] [[PubMed](#)]
20. Sun, J.; Shannon, M.; Ando, Y.; Schnackenberg, L.K.; Khan, N.A.; Portilla, D.; Beger, R.D. Serum metabolomic profiles from patients with acute kidney injury: A pilot study. *J. Chromatogr. B* **2012**, *893–894*, 107–113. [[CrossRef](#)]
21. Gronwald, W.; Klein, M.S.; Zeltner, R.; Schulze, B.-D.; Reinhold, S.W.; Deutschmann, M.; Immervoll, A.-K.; Böger, C.A.; Banas, B.; Eckardt, K.-U.; et al. Detection of autosomal dominant polycystic kidney disease by NMR spectroscopic fingerprinting of urine. *Kidney Int.* **2011**, *79*, 1244–1253. [[CrossRef](#)]
22. Blydt-Hansen, T.D.; Sharma, A.; Gibson, I.; Mandal, R.; Wishart, D.S. Urinary Metabolomics for Noninvasive Detection of Borderline and Acute T Cell-Mediated Rejection in Children After Kidney Transplantation. *Am. J. Transplant.* **2014**, *14*, 2339–2349. [[CrossRef](#)]
23. Hallan, S.; Afkarian, M.; Zelnick, L.; Kestenbaum, B.; Sharma, S.; Saito, R.; Darshi, M.; Barding, G.; Raftery, D.; Ju, W.; et al. Metabolomics and Gene Expression Analysis Reveal Down-regulation of the Citric Acid (TCA) Cycle in Non-diabetic CKD Patients. *EBioMedicine* **2017**, *26*, 68–77. [[CrossRef](#)] [[PubMed](#)]
24. Kalim, S.; Clish, C.B.; Wenger, J.; Elmariah, S.; Yeh, R.W.; Deferio, J.J.; Pierce, K.; Deik, A.; Gerszten, R.E.; Thadhani, R.; et al. A Plasma Long-Chain Acylcarnitine Predicts Cardiovascular Mortality in Incident Dialysis Patients. *J. Am. Hear. Assoc.* **2013**, *2*, e000542. [[CrossRef](#)] [[PubMed](#)]
25. Shah, V.O.; Townsend, R.R.; Feldman, H.I.; Pappan, K.L.; Kensicki, E.; Jagt, D.L.V. Plasma Metabolomic Profiles in Different Stages of CKD. *Clin. J. Am. Soc. Nephrol.* **2012**, *8*, 363–370. [[CrossRef](#)]
26. Coresh, J.; Inker, L.A.; Sang, Y.; Chen, J.; Shafi, T.; Post, W.S.; Shlipak, M.G.; Ford, L.; Goodman, K.; Perichon, R.; et al. Metabolomic profiling to improve glomerular filtration rate estimation: A proof-of-concept study. *Nephrol. Dial. Transplant.* **2019**, *34*, 825–833. [[CrossRef](#)]
27. Chen, D.Q.; Cao, G.; Zhao, Y.-Y. Identification of serum metabolites associating with chronic kidney disease progression and anti-fibrotic effect of 5-methoxytryptophan. *Nat. Commun.* **2019**, *10*, 1–15.
28. Sekula, P.; Tin, A.; Schultheiss, U.T.; Baid-Agrawal, S.; Mohny, R.P.; Steinbrenner, I.; Yu, B.; Luo, S.; Boerwinkle, E.; Eckardt, K.-U.; et al. Urine 6-Bromotryptophan: Associations with Genetic Variants and Incident End-Stage Kidney Disease. *Sci. Rep.* **2020**, *10*, 10018. [[CrossRef](#)] [[PubMed](#)]
29. Steinbrenner, I.; Schultheiss, U.T.; Kotsis, F.; Schlosser, P.; Stockmann, H.; Mohny, R.P.; Schmid, M.; Oefner, P.J.; Eckardt, K.-U.; Köttgen, A.; et al. Urine Metabolite Levels, Adverse Kidney Outcomes, and Mortality in CKD Patients: A Metabolome-wide Association Study. *Am. J. Kidney Dis.* **2021**. [[CrossRef](#)]
30. Altenbuchinger, M.; Zacharias, H.U.; Solbrig, S.; Schäfer, A.; Büyüközkan, M.; Schultheiß, U.T.; Kotsis, F.; Köttgen, A.; Spang, R.; Oefner, P.J.; et al. A multi-source data integration approach reveals novel associations between metabolites and renal outcomes in the German Chronic Kidney Disease study. *Sci. Rep.* **2019**, *9*, 13954. [[CrossRef](#)]

31. Schlosser, P.; Li, Y.; Sekula, P.; Raffler, J.; Grundner-Culemann, F.; Pietzner, M.; Cheng, Y.; Wuttke, M.; Steinbrenner, I.; Schultheiss, U.T.; et al. Genetic studies of urinary metabolites illuminate mechanisms of detoxification and excretion in humans. *Nat. Genet.* **2020**, *52*, 167–176. [[CrossRef](#)] [[PubMed](#)]
32. Mäkinen, V.-P.; Tynkkynen, T.; Soininen, P.; Peltola, T.; Kangas, A.J.; Forsblom, C.; Thorn, L.M.; Kaski, K.; Laatikainen, R.; Ala-Korpela, M.; et al. Metabolic Diversity of Progressive Kidney Disease in 325 Patients with Type 1 Diabetes (the FinnDiane Study). *J. Proteome Res.* **2012**, *11*, 1782–1790. [[CrossRef](#)] [[PubMed](#)]
33. Suhre, K.; Schwartz, J.E.; Sharma, V.K.; Chen, Q.; Lee, J.R.; Muthukumar, T.; Dadhania, D.M.; Ding, R.; Ikle, D.N.; Bridges, N.D.; et al. Urine metabolite profiles predictive of human kidney allograft status. *J. Am. Soc. Nephrol.* **2016**, *27*, 626–636. [[CrossRef](#)]
34. Goek, O.N.; Prehn, C.; Sekula, P.; Römisch-Margl, W.; Döring, A.; Gieger, C.; Heier, M.; Koenig, W.; Wang-Sattler, R.; Illig, T. Metabolites associate with kidney function decline and incident chronic kidney disease in the general population. *Nephrol. Dial. Transplant.* **2013**, *28*, 2131–2138. [[CrossRef](#)]
35. Sekula, P.; Goek, O.-N.; Quaye, L.; Barrios, C.; Levey, A.S.; Römisch-Margl, W.; Menni, C.; Yet, I.; Gieger, C.; Inker, L.A.; et al. A Metabolome-Wide Association Study of Kidney Function and Disease in the General Population. *J. Am. Soc. Nephrol.* **2015**, *27*, 1175–1188. [[CrossRef](#)]
36. Stanimirova, I.; Banasik, M.; Ząbek, A.; Dawiskiba, T.; Kościelska-Kasprzak, K.; Wojtowicz, W.; Krajewska, M.; Janczak, D.; Młynarz, P. Serum metabolomics approach to monitor the changes in metabolite profiles following renal transplantation. *Sci. Rep.* **2020**, *10*, 17223. [[CrossRef](#)]
37. Luo, S.; Coresh, J.; Tin, A.; Rebholz, C.M.; Appel, L.J.; Chen, J.; Vasani, R.S.; Anderson, A.H.; Feldman, H.I.; Kimmel, P.L.; et al. Serum Metabolomic Alterations Associated with Proteinuria in CKD. *Clin. J. Am. Soc. Nephrol.* **2019**, *14*, 342–353. [[CrossRef](#)]
38. Hasegawa, S.; Inagi, R. Harnessing Metabolomics to Describe the Pathophysiology Underlying Progression in Diabetic Kidney Disease. *Curr. Diab. Rep.* **2021**, *21*, 1–6. [[CrossRef](#)]
39. Taherkhani, A.; Farrokhi Yekta, R.; Mohseni, M.; Saidijam, M.; Arefi Oskouie, A. Chronic kidney disease: A review of proteomic and metabolomic approaches to membranous glomerulonephritis, focal segmental glomerulosclerosis, and IgA nephropathy biomarkers. *Proteome Sci.* **2019**, *17*, 1–18. [[CrossRef](#)]
40. Van Stralen, K.; Dekker, F.; Zoccali, C.; Jager, K. Case-Control Studies—An Efficient Observational Study Design. *Nephron Clin. Pr.* **2010**, *114*, c1–c4. [[CrossRef](#)] [[PubMed](#)]
41. Keogh, R.H.; Cox, D.R. *Case-Control Studies*; Cambridge University Press (CUP): Cambridge, UK, 2014.
42. Wu, I.-W.; Lee, C.-C.; Hsu, H.-J.; Sun, C.-Y.; Chen, Y.-C.; Yang, K.-J.; Yang, C.-W.; Chung, W.-H.; Lai, H.-C.; Chang, L.-C.; et al. Compositional and Functional Adaptations of Intestinal Microbiota and Related Metabolites in CKD Patients Receiving Dietary Protein Restriction. *Nutrients* **2020**, *12*, 2799. [[CrossRef](#)]
43. Barrios, C.; Zierer, J.; Würtz, P.; Haller, T.; Metspalu, A.; Gieger, C.; Thorand, B.; Meisinger, C.; Waldenberger, M.; Raitakari, O.; et al. Circulating metabolic biomarkers of renal function in diabetic and non-diabetic populations. *Sci. Rep.* **2018**, *8*, 15249. [[CrossRef](#)] [[PubMed](#)]
44. Eckardt, K.-U.; Bärthlein, B.; Baid-Agrawal, S.; Beck, A.; Busch, M.; Eitner, F.; Ekici, A.B.; Floege, J.; Gefeller, O.; Haller, H.; et al. The German Chronic Kidney Disease (GCKD) study: Design and methods. *Nephrol. Dial. Transplant.* **2011**, *27*, 1454–1460. [[CrossRef](#)]
45. Dienemann, T.; Fujii, N.; Orlandi, P.; Nessel, L.; Furth, S.L.; Hoy, W.E.; Matsuo, S.; Mayer, G.; Methven, S.; Schaefer, F.; et al. International Network of Chronic Kidney Disease cohort studies (iNET-CKD): A global network of chronic kidney disease cohorts. *BMC Nephrol.* **2016**, *17*, 1–9. [[CrossRef](#)]
46. Rinschen, M.M.; Palygin, O.; Guijas, C.; Palermo, A.; Palacio-Escat, N.; Domingo-Almenara, X.; Montenegro-Burke, R.; Saez-Rodriguez, J.; Staruschenko, A.; Siuzdak, G. Metabolic rewiring of the hypertensive kidney. *Sci. Signal.* **2019**, *12*, eaax9760. [[CrossRef](#)] [[PubMed](#)]
47. Winkvist, A.; Bärebring, L.; Gjerdtsson, I.; Ellegård, L.; Lindqvist, H.M. A randomized controlled cross-over trial investigating the effect of anti-inflammatory diet on disease activity and quality of life in rheumatoid arthritis: The Anti-inflammatory Diet In Rheumatoid Arthritis (ADIRA) study protocol. *Nutr. J.* **2018**, *17*, 44. [[CrossRef](#)] [[PubMed](#)]
48. Adamska-Patruno, E.; Samczuk, P.; Ciborowski, M.; Godzien, J.; Pietrowska, K.; Bauer, W.; Gorska, M.; Barbas, C.; Kretowski, A. Metabolomics Reveal Altered Postprandial Lipid Metabolism After a High-Carbohydrate Meal in Men at High Genetic Risk of Diabetes. *J. Nutr.* **2019**, *149*, 915–922. [[CrossRef](#)] [[PubMed](#)]
49. Beuchel, C.; Becker, S.; Dittrich, J.; Kirsten, H.; Toenjes, A.; Stumvoll, M.; Loeffler, M.; Thiele, H.; Beutner, F.; Thiery, J.; et al. Clinical and lifestyle related factors influencing whole blood metabolite levels—A comparative analysis of three large cohorts. *Mol. Metab.* **2019**, *29*, 76–85. [[CrossRef](#)] [[PubMed](#)]
50. Altmaier, E.; Fobo, G.; Heier, M.; Thorand, B.; Meisinger, C.; Römisch-Margl, W.; Waldenberger, M.; Gieger, C.; Illig, T.; Adamski, J.; et al. Metabolomics approach reveals effects of antihypertensives and lipid-lowering drugs on the human metabolism. *Eur. J. Epidemiol.* **2014**, *29*, 325–336. [[CrossRef](#)]
51. Chua, E.C.P.; Shui, G.; Tian-Guang Lee, I.; Lau, P.; Tan, L.-C.; Yeo, S.-C.; Lam, B.D.; Bulchand, S.; Summers, S.A.; Puvanendran, K.; et al. Extensive diversity in circadian regulation of plasma lipids and evidence for different circadian metabolic phenotypes in humans. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 14468–14473. [[CrossRef](#)] [[PubMed](#)]

52. Suhre, K.; Shin, S.Y.; Petersen, A.K.; Mohney, R.P.; Meredith, D.; Wagele, B.; Altmaier, E.; Gram, C.; Deloukas, P.; Erdmann, J.; et al. Human metabolic individuality in biomedical and pharmaceutical research. *Nature* **2011**, *477*, 54–60. [CrossRef]
53. Smith, L.; Villaret-Cazadamont, J.; Claus, S.P.; Canlet, C.; Guillou, H.; Cabaton, N.J.; Ellero-Simatos, S. Important Considerations for Sample Collection in Metabolomics Studies with a Special Focus on Applications to Liver Functions. *Metabolites* **2020**, *10*, 104. [CrossRef] [PubMed]
54. Barton, R.H.; Waterman, D.S.; Bonner, F.W.; Holmes, E.; Clarke, R.; Nicholson, J.; Lindon, J.; the PROCARDIS Consortium. The influence of EDTA and citrate anticoagulant addition to human plasma on information recovery from NMR-based metabolic profiling studies. *Mol. BioSyst.* **2009**, *6*, 215–224. [CrossRef] [PubMed]
55. Jiménez, B.; Holmes, E.; Heude, C.; Tolson, R.F.; Harvey, N.; Lodge, S.L.; Chetwynd, A.J.; Cannet, C.; Fang, F.; Pearce, J.T.M.; et al. Quantitative Lipoprotein Subclass and Low Molecular Weight Metabolite Analysis in Human Serum and Plasma by 1H NMR Spectroscopy in a Multilaboratory Trial. *Anal. Chem.* **2018**, *90*, 11962–11971. [CrossRef]
56. Von Schlippenbach, T.; Oefner, P.J.; Gronwald, W. Systematic Evaluation of Non-Uniform Sampling Parameters in the Targeted Analysis of Urine Metabolites by 1H,1H 2D NMR Spectroscopy. *Sci. Rep.* **2018**, *8*, 4249. [CrossRef] [PubMed]
57. Soininen, P.; Kangas, A.; Würtz, P.; Suna, T.; Ala-Korpela, M. Quantitative Serum Nuclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and Genetics. *Circ. Cardiovasc. Genet.* **2015**, *8*, 192–206. [CrossRef]
58. Percival, B.C.; Grootveld, M.; Gibson, M.; Osman, Y.; Molinari, M.; Jafari, F.; Sahota, T.; Martin, M.; Casanova, F.; Mather, M.L.; et al. Low-field, benchtop NMR spectroscopy as a potential tool for point-of-care diagnostics of metabolic conditions: Validation, protocols and computational models. *High Throughput* **2019**, *8*, 2. [CrossRef]
59. Leenders, J.; Grootveld, M.; Percival, B.; Gibson, M.; Casanova, F.; Wilson, P.B. Benchtop Low-Frequency 60 MHz NMR Analysis of Urine: A Comparative Metabolomics Investigation. *Metabolites* **2020**, *10*, 155. [CrossRef]
60. Edgar, M.; Percival, B.C.; Gibson, M.; Jafari, F.; Grootveld, M. Low-field benchtop NMR spectroscopy as a potential non-stationary tool for point-of-care urinary metabolite tracking in diabetic conditions. *Diabetes Res. Clin. Pr.* **2021**, *171*, 108554. [CrossRef]
61. Dona, A.C.; Jiménez, B.; Schäfer, H.; Humpfer, E.; Spraul, M.; Lewis, M.R.; Pearce, J.T.M.; Holmes, E.; Lindon, J.C.; Nicholson, J.K. Precision High-Throughput Proton NMR Spectroscopy of Human Urine, Serum, and Plasma for Large-Scale Metabolic Phenotyping. *Anal. Chem.* **2014**, *86*, 9887–9894. [CrossRef]
62. Lodge, S.; Nitschke, P.; Leng Loo, R.; Kimhofer, T.; Bong, S.-H.; Richards, T.; Begum, S.; Spraul, M.; Schaefer, H.; Lindon, J.C.; et al. Low Volume in Vitro Diagnostic Proton NMR Spectroscopy of Human Blood Plasma for Lipoprotein and Metabolite Analysis: Application to SARS-CoV-2 Biomarkers. *J. Proteome Res.* **2021**, *20*, 1415–1423. [CrossRef] [PubMed]
63. Tveite Bjerrum, J.T. Metabonomics: Methods and protocols. *Methods Mol. Biol.* **2015**, 1277. Available online: <https://mosys.univie.ac.at/publications/books/metabolomics-methods-and-protocols/> (accessed on 15 July 2021).
64. Zacharias, H.U.; Hochrein, J.; Klein, M.; Samol, C.; Oefner, P.; Gronwald, W. Current Experimental, Bioinformatic and Statistical Methods used in NMR Based Metabolomics. *Curr. Metab.* **2013**, *1*, 253–268. [CrossRef]
65. Dettmer, K.; Aronov, P.A.; Hammock, B.D. Mass spectrometry-based metabolomics. *Mass Spectrom. Rev.* **2006**, *26*, 51–78. [CrossRef] [PubMed]
66. Suhre, K.; Meisinger, C.; Döring, A.; Altmaier, E.; Belcredi, P.; Gieger, C.; Chang, D.; Milburn, M.V.; Gall, W.E.; Weinberger, K.; et al. Metabolic Footprint of Diabetes: A Multiplatform Metabolomics Study in an Epidemiological Setting. *PLoS ONE* **2010**, *5*, e13953. [CrossRef]
67. Almstetter, M.F.; Oefner, P.J.; Dettmer, K. Comprehensive two-dimensional gas chromatography in metabolomics. *Anal. Bioanal. Chem.* **2012**, *402*, 1993–2013. [CrossRef]
68. François, I.; Sandra, K.; Sandra, P. Comprehensive liquid chromatography: Fundamental aspects and practical considerations—A review. *Anal. Chim. Acta* **2009**, *641*, 14–31. [CrossRef]
69. Zhang, X.W.; Li, Q.H.; Di Xu, Z.; Dou, J.J. Mass spectrometry-based metabolomics in health and medical science: A systematic review. *RSC Adv.* **2020**, *10*, 3092–3104. [CrossRef]
70. Amberg, A.; Riefke, B.; Schlotterbeck, G.; Ross, A.; Senn, H.; Dieterle, F.; Keck, M. NMR and MS Methods for Metabolomics BT—Drug Safety Evaluation: Methods and Protocols. *Methods Mol. Biol.* **2017**, *1641*, 229–258.
71. Beckonert, O.; Keun, H.C.; Ebbels, T.; Bundy, J.; Holmes, E.; Lindon, J.; Nicholson, J. Metabolic profiling, metabolomic and metabonomic procedures for NMR spectroscopy of urine, plasma, serum and tissue extracts. *Nat. Protoc.* **2007**, *2*, 2692–2703. [CrossRef]
72. Van den Berg, R.A.; Hoefsloot, H.C.J.; Westerhuis, J.A.; Smilde, A.K.; van der Werf, M.J. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genom.* **2006**, *7*, 142. [CrossRef]
73. Vettukattil, R. Preprocessing of Raw Metabonomic Data. *Adv. Struct. Saf. Stud.* **2015**, 1277, 123–136. [CrossRef]
74. Tian, H.; Li, B.; Shui, G. Untargeted LC–MS Data Preprocessing in Metabolomics. *J. Anal. Test.* **2017**, *1*, 187–192. [CrossRef]
75. Wallmeier, J.; Samol, C.; Ellmann, L.; Zacharias, H.U.; Vogl, F.C.; Garcia, M.; Dettmer, K.; Oefner, P.J.; Gronwald, W.; GCKD Study Investigators. Quantification of Metabolites by NMR Spectroscopy in the Presence of Protein. *J. Proteome Res.* **2017**, *16*, 1784–1796. [CrossRef]
76. McHugh, C.E.; Flott, T.L.; Schooff, C.R.; Smiley, Z.; Puskarich, M.; Myers, D.D.; Younger, J.G.; Jones, A.E.; Stringer, K.A. Rapid, Reproducible, Quantifiable NMR Metabolomics: Methanol and Methanol: Chloroform Precipitation for Removal of Macromolecules in Serum and Whole Blood. *Metabolites* **2018**, *8*, 93. [CrossRef] [PubMed]

77. Zacharias, H.U.; Rehberg, T.; Mehrl, S.; Richtmann, D.; Wettig, T.; Oefner, P.J.; Spang, R.; Gronwald, W.; Altenbuchinger, M. Scale-Invariant Biomarker Discovery in Urine and Plasma Metabolite Fingerprints. *J. Proteome Res.* **2017**, *16*, 3596–3605. [[CrossRef](#)] [[PubMed](#)]
78. Altenbuchinger, M.; Rehberg, T.; Zacharias, H.U.; Stämmeler, F.; Dettmer, K.; Weber, D.; Hiergeist, A.; Gessner, A.; Holler, E.; Oefner, P.J.; et al. Reference point insensitive molecular data analysis. *Bioinformatics* **2017**, *33*, 219–226. [[CrossRef](#)] [[PubMed](#)]
79. Lin, W.; Shi, P.; Feng, R.; Li, H. Variable selection in regression with compositional covariates. *Biometrika* **2014**, *101*, 785–797. [[CrossRef](#)]
80. Wishart, D.S.; Feunang, Y.D.; Marcu, A.; Guo, A.C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Allison, P.; Karu, N.; et al. HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Res.* **2018**, *46*, D608–D617. [[CrossRef](#)] [[PubMed](#)]
81. Vogl, F.C.; GCKD Study Investigators; Mehrl, S.; Heizinger, L.; Schlecht, I.; Zacharias, H.U.; Ellmann, L.; Nürnberger, N.; Gronwald, W.; Leitzmann, M.F.; et al. Evaluation of dilution and normalization strategies to correct for urinary output in HPLC-HRTOFMS metabolomics. *Anal. Bioanal. Chem.* **2016**, *408*, 8483–8493. [[CrossRef](#)] [[PubMed](#)]
82. Waikar, S.S.; Sabbiseti, V.S.; Bonventre, J.V. Normalization of urinary biomarkers to creatinine during changes in glomerular filtration rate. *Kidney Int.* **2010**, *78*, 486–494. [[CrossRef](#)]
83. Stevens, L.A.; Levey, A.S. Measured GFR as a Confirmatory Test for Estimated GFR. *J. Am. Soc. Nephrol.* **2009**, *20*, 2305–2313. [[CrossRef](#)] [[PubMed](#)]
84. Zamora-Ros, R.; Rabassa, M.; Cherubini, A.; Urpi-Sarda, M.; Llorach, R.; Bandinelli, S.; Ferrucci, L.; Andres-Lacueva, C. Comparison of 24-h volume and creatinine-corrected total urinary polyphenol as a biomarker of total dietary polyphenols in the Invecchiare InCHIANTI study. *Anal. Chim. Acta* **2011**, *704*, 110–115. [[CrossRef](#)] [[PubMed](#)]
85. Curhan, G. Cystatin C: A Marker of Renal Function or Something More? *Clin. Chem.* **2005**, *51*, 293–294. [[CrossRef](#)] [[PubMed](#)]
86. Warrack, B.M.; Hnatyshyn, S.; Ott, K.-H.; Reily, M.; Sanders, M.; Zhang, H.; Drexler, D.M. Normalization strategies for metabonomic analysis of urine samples. *J. Chromatogr. B* **2009**, *877*, 547–552. [[CrossRef](#)]
87. Chen, Y.; Shen, G.; Zhang, R.; He, J.; Zhang, Y.; Xu, J.; Yang, W.; Chen, X.; Song, Y.; Abliz, Z. Combination of Injection Volume Calibration by Creatinine and MS Signals' Normalization to Overcome Urine Variability in LC-MS-Based Metabolomics Studies. *Anal. Chem.* **2013**, *85*, 7659–7665. [[CrossRef](#)]
88. Delaglio, F.; Grzesiek, S.; Vuister, G.W.; Zhu, G.; Pfeifer, J.; Bax, A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **1995**, *6*, 277–293. [[CrossRef](#)]
89. Ravanbakhsh, S.; Liu, P.; Bjordahl, T.C.; Mandal, R.; Grant, J.R.; Wilson, M.; Eisner, R.; Sinelnikov, I.; Hu, X.; Luchinat, C.; et al. Accurate, fully-automated NMR spectral profiling for metabolomics. *PLoS ONE* **2015**, *10*, e0124219. [[CrossRef](#)]
90. Madrid-Gambin, F.; Oller-Moreno, S.; Fernandez, L.; Bartova, S.; Giner, M.P.; Joyce, C.; Ferraro, F.; Montoliu, I.; Moco, S.; Marco, S. AlpsNMR: An R package for signal processing of fully untargeted NMR-based metabolomics. *Bioinformatics* **2020**, *36*, 2943–2945. [[CrossRef](#)] [[PubMed](#)]
91. Vu, T.N.; Valkenborg, D.; Smets, K.; Verwaest, K.A.; Dommissie, R.; Lemièrre, F.; Verschoren, A.; Goethals, B.; Laukens, K. An integrated workflow for robust alignment and simplified quantitative analysis of NMR spectrometry data. *BMC Bioinform.* **2011**, *12*, 405. [[CrossRef](#)]
92. Hoffmann, N.; Stoye, J. ChromA: Signal-based retention time alignment for chromatography–mass spectrometry data. *Bioinformatics* **2009**, *25*, 2080–2081. [[CrossRef](#)] [[PubMed](#)]
93. Wang, S.-Y.; Ho, T.-J.; Kuo, C.-H.; Tseng, Y.J. Chromaligner: A web server for chromatogram alignment. *Bioinformatics* **2010**, *26*, 2338–2339. [[CrossRef](#)] [[PubMed](#)]
94. Lommen, A. MetAlign: Interface-Driven, Versatile Metabolomics Tool for Hyphenated Full-Scan Mass Spectrometry Data Preprocessing. *Anal. Chem.* **2009**, *81*, 3079–3086. [[CrossRef](#)] [[PubMed](#)]
95. Katajamaa, M.; Oresic, M. Processing methods for differential analysis of LC/MS profile data. *BMC Bioinform.* **2005**, *6*, 179. [[CrossRef](#)]
96. Katajamaa, M.; Miettinen, J.; Orešič, M. MZmine: Toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* **2006**, *22*, 634–636. [[CrossRef](#)] [[PubMed](#)]
97. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Orešič, M. MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinform.* **2010**, *11*, 395. [[CrossRef](#)] [[PubMed](#)]
98. Röst, H.; Sachsenberg, T.; Aiche, S.; Bielow, C.; Weissner, H.; Aicheler, F.; Andreotti, S.; Ehrlich, H.-C.; Gutenbrunner, P.; Kenar, E.; et al. OpenMS: A flexible open-source software platform for mass spectrometry data analysis. *Nat. Methods* **2016**, *13*, 741–748. [[CrossRef](#)]
99. Smith, C.A.; Want, E.J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78*, 779–787. [[CrossRef](#)]
100. Benton, H.P.; Wong, D.M.; Trauger, S.; Siuzdak, G. XCMS2: Processing Tandem Mass Spectrometry Data for Metabolite Identification and Structural Characterization. *Anal. Chem.* **2008**, *80*, 6382–6389. [[CrossRef](#)]
101. Melamud, E.; Vastag, L.; Rabinowitz, J.D. Metabolomic Analysis and Visualization Engine for LC–MS Data. *Anal. Chem.* **2010**, *82*, 9818–9826. [[CrossRef](#)]
102. Domingo-Almenara, X.; Brezmes, J.; Vinaixa, M.; Samino, S.; Ramirez, N.; Ramon-Krauel, M.; Lerin, C.; Díaz, M.; Ibáñez, L.; Correig, X.; et al. eRah: A Computational Tool Integrating Spectral Deconvolution and Alignment with Quantification and Identification of Metabolites in GC/MS-Based Metabolomics. *Anal. Chem.* **2016**, *88*, 9821–9829. [[CrossRef](#)] [[PubMed](#)]

103. Anderson, P.E.; Reo, N.V.; Delraso, N.J.; Doom, T.E.; Raymer, M.L. Gaussian binning: A new kernel-based method for processing NMR spectroscopic data for metabolomics. *Metabolomics* **2008**, *4*, 261–272. [[CrossRef](#)]
104. Davis, R.A.; Charlton, A.J.; Godward, J.; Jones, S.A.; Harrison, M.; Wilson, J.C. Adaptive binning: An improved binning method for metabolomics data using the undecimated wavelet transform. *Chemom. Intell. Lab. Syst.* **2007**, *85*, 144–154. [[CrossRef](#)]
105. De Meyer, T.; Sinnaeve, D.; Van Gasse, B.; Tsiporkova, E.; Rietzschel, E.R.; De Buyzere, M.L.; Gillebert, T.; Bekaert, S.; Martins, J.; Van Criekinge, W. NMR-Based Characterization of Metabolic Alterations in Hypertension Using an Adaptive, Intelligent Binning Algorithm. *Anal. Chem.* **2008**, *80*, 3783–3790. [[CrossRef](#)] [[PubMed](#)]
106. AndersonDeirdre, P.E.; Mahle, D.A.; Doom, T.E.; Reo, N.V.; Delraso, N.J.; Raymer, M.L. Dynamic adaptive binning: An improved quantification technique for NMR spectroscopic data. *Metabolomics* **2011**, *7*, 179–190. [[CrossRef](#)]
107. Blaise, B.J.; Shintu, L.; Elena-Herrmann, B.; Emsley, L.; Dumas, M.-E.; Toulhoat, P. Statistical Recoupling Prior to Significance Testing in Nuclear Magnetic Resonance Based Metabonomics. *Anal. Chem.* **2009**, *81*, 6242–6251. [[CrossRef](#)]
108. Rodriguez-Martinez, A.; Ayala, R.; Posma, J.M.; Harvey, N.; Jiménez, B.; Sonomura, K.; Sato, T.-A.; Matsuda, F.; Zalloua, P.; Gauguier, D.; et al. pJRES Binning Algorithm (JBA): A new method to facilitate the recovery of metabolic information from pJRES 1H NMR spectra. *Bioinformatics* **2018**, *35*, 1916–1922. [[CrossRef](#)] [[PubMed](#)]
109. Weljie, A.M.; Newton, J.; Mercier, P.; Carlson, E.; Slupsky, C.M. Targeted profiling: Quantitative analysis of 1H NMR metabolomics data. *Anal. Chem.* **2006**, *78*, 4430–4442. [[CrossRef](#)]
110. Hedjazi, L.; Gauguier, D.; Zalloua, P.A.; Nicholson, J.K.; Dumas, M.-E.; Cazier, J.-B. mQTL.NMR: An Integrated Suite for Genetic Mapping of Quantitative Variations of 1H NMR-Based Metabolic Profiles. *Anal. Chem.* **2015**, *87*, 4377–4384. [[CrossRef](#)]
111. Rodriguez-Martinez, A.; Posma, J.M.; Ayala, R.; Neves, A.L.; Anwar, M.; Petretto, E.; Emanuelli, C.; Gauguier, D.; Nicholson, J.; Dumas, M.-E. MWASTools: An R/bioconductor package for metabolome-wide association studies. *Bioinformatics* **2017**, *34*, 890–892. [[CrossRef](#)]
112. Beirnaert, C.; Meysman, P.; Vu, T.N.; Hermans, N.; Apers, S.; Pieters, L.; Covaci, A.; Laukens, K. speaq 2.0: A complete workflow for high-throughput 1D NMR spectra processing and quantification. *PLoS Comput. Biol.* **2018**, *14*, e1006018. [[CrossRef](#)]
113. Xia, J.; Psychogios, N.; Young, N.; Wishart, D.S. MetaboAnalyst: A web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **2009**, *37*, W652–W660. [[CrossRef](#)]
114. Tikunov, Y.M.; Laptinok, S.; Hall, R.D.; Bovy, A.; De Vos, R.C.H. MSclust: A tool for unsupervised mass spectra extraction of chromatography-mass spectrometry ion-wise aligned data. *Metabolomics* **2012**, *8*, 714–718. [[CrossRef](#)] [[PubMed](#)]
115. Gorrochategui, E.; Jaumot, J.; Tauler, R. ROIMCR: A powerful analysis strategy for LC-MS metabolomic datasets. *BMC Bioinform.* **2019**, *20*, 256. [[CrossRef](#)] [[PubMed](#)]
116. Tauler, R. Multivariate curve resolution applied to second order data. *Chemom. Intell. Lab. Syst.* **1995**, *30*, 133–146. [[CrossRef](#)]
117. Hao, J.; Liebeke, M.; Astle, W.; De Iorio, M.; Bundy, J.G.; Ebbels, T. Bayesian deconvolution and quantification of metabolites in complex 1D NMR spectra using BATMAN. *Nat. Protoc.* **2014**, *9*, 1416–1427. [[CrossRef](#)] [[PubMed](#)]
118. Hao, J.; Astle, W.; De Iorio, M.; Ebbels, T.M.D. BATMAN—An R package for the automated quantification of metabolites from nuclear magnetic resonance spectra using a Bayesian model. *Bioinformatics* **2012**, *28*, 2088–2090. [[CrossRef](#)]
119. Hughes, T.; Wilson, H.D.; De Vera, I.M.S.; Kojetin, D.J. Deconvolution of Complex 1D NMR Spectra Using Objective Model Selection. *PLoS ONE* **2015**, *10*, e0134474. [[CrossRef](#)] [[PubMed](#)]
120. Häckl, M.; Tauber, P.; Schweda, F.; Zacharias, H.U.; Altenbuchinger, M.; Oefner, P.J.; Gronwald, W. An R-package for the Deconvolution and Integration of 1D NMR data: MetaboDecon1D. *Metabolites* **2021**, *11*, 452. [[CrossRef](#)]
121. Haslauer, K.E.; Schmitt-Kopplin, P.; Heinzmann, S.S. Data processing optimization in untargeted metabolomics of urine using voigt lineshape model non-linear regression analysis. *Metabolites* **2021**, *11*, 285. [[CrossRef](#)]
122. Wei, X.; Shi, X.; Kim, S.; Zhang, L.; Patrick, J.S.; Binkley, J.; McClain, C.; Zhang, X. Data preprocessing method for liquid chromatography-mass spectrometry based metabolomics. *Anal. Chem.* **2012**, *84*, 7963–7971. [[CrossRef](#)]
123. Yin, Y.; Wang, R.; Cai, Y.; Wang, Z.; Zhu, Z.-J. DecoMetDIA: Deconvolution of Multiplexed MS/MS Spectra for Metabolite Identification in SWATH-MS-Based Untargeted Metabolomics. *Anal. Chem.* **2019**, *91*, 11897–11904. [[CrossRef](#)]
124. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
125. van Buuren, S.; Groothuis-Oudshoorn, K. Mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **2011**, *45*, 1–67. [[CrossRef](#)]
126. Kowarik, A.; Templ, M. Imputation with the R Package VIM. *J. Stat. Softw.* **2016**, *74*, 1–16. [[CrossRef](#)]
127. Liaw, A.; Wiener, M. Classification and regression by randomForest. *R News* **2002**, *2*, 18–22.
128. Zhang, F.; Robinette, S.L.; Bruschiweiler-Li, L.; Bruschweiler, R. Web server suite for complex mixture analysis by covariance NMR. *Magn. Reson. Chem.* **2009**, *47*, S118–S122. [[CrossRef](#)]
129. Tulpan, D.; Léger, S.; Belliveau, L.; Culf, A.; Čuperlović-Culf, M. MetaboHunter: An automatic approach for identification of metabolites from 1H-NMR spectra of complex mixtures. *BMC Bioinform.* **2011**, *12*, 1–22. [[CrossRef](#)]
130. Xia, J.; Bjorn Dahl, T.C.; Tang, P.; Wishart, D.S. MetaboMiner—Semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. *BMC Bioinform.* **2008**, *9*, 507. [[CrossRef](#)]
131. Tardivel, P.J.C.; Canlet, C.; Lefort, G.; Tremblay-Franco, M.; Debrauwer, L.; Concordet, D.; Servien, R. ASICS: An automatic method for identification and quantification of metabolites in complex 1D 1H NMR spectra. *Metabolomics* **2017**, *13*, 109. [[CrossRef](#)]

132. Draper, J.; Enot, D.P.; Parker, D.; Beckmann, M.; Snowdon, S.; Lin, W.; Zubair, H. Metabolite signal identification in accurate mass metabolomics data with MZedDB, an interactive m/z annotation tool utilising predicted ionisation behaviour ‘rules’. *BMC Bioinform.* **2009**, *10*, 227. [[CrossRef](#)]
133. Klein, M.S.; Oefner, P.J.; Gronwald, W. MetaboQuant: A tool combining individual peak calibration and outlier detection for accurate metabolite quantification in 1D ¹H and 1H-13C HSQC NMR spectra. *Biotechniques* **2013**, *54*, 251–256. [[CrossRef](#)]
134. Röhnisch, H.E.; Eriksson, J.; Müllner, E.; Agback, P.; Sandström, C.; Moazzami, A.A. AQuA: An Automated Quantification Algorithm for High-Throughput NMR-Based Metabolomics and Its Application in Human Plasma. *Anal. Chem.* **2018**, *90*, 2095–2102. [[CrossRef](#)] [[PubMed](#)]
135. Purohit, P.V.; Rocke, D.M.; Viant, M.R.; Woodruff, D.L. Discrimination Models Using Variance-Stabilizing Transformation of Metabolomic NMR Data. *OMICS A J. Integr. Biol.* **2004**, *8*, 118–130. [[CrossRef](#)]
136. Eriksson, L.; Antti, H.; Gottfries, J.; Holmes, E.; Johansson, E.; Lindgren, F.; Long, I.; Trygg, J.; Wold, S. Using chemometrics for navigating in the large data sets of genomics, proteomics, and metabonomics (gpm). *Anal. Bioanal. Chem.* **2004**, *380*, 419–429. [[CrossRef](#)]
137. Huber, W.; Von Heydebreck, A.; Sültmann, H.; Poustka, A.; Vingron, M. Variance stabilization applied to microarray data calibration and to the quantification of differential expression. *Bioinformatics* **2002**, *18*, S96–S104. [[CrossRef](#)] [[PubMed](#)]
138. Chawade, A.; Alexandersson, E.; Levander, F. Normalyzer: A Tool for Rapid Evaluation of Normalization Methods for Omics Data Sets. *J. Proteome Res.* **2014**, *13*, 3114–3120. [[CrossRef](#)] [[PubMed](#)]
139. Li, B.; Tang, J.; Yang, Q.; Cui, X.; Li, S.; Chen, S.; Cao, Q.; Xue, W.; Chen, N.; Zhu, F. Performance Evaluation and Online Realization of Data-driven Normalization Methods Used in LC/MS based Untargeted Metabolomics Analysis. *Sci. Rep.* **2016**, *6*, 38881. [[CrossRef](#)] [[PubMed](#)]
140. Dieterle, F.; Ross, A.; Schlotterbeck, G.; Senn, H. Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in ¹H NMR Metabonomics. *Anal. Chem.* **2006**, *78*, 4281–4290. [[CrossRef](#)] [[PubMed](#)]
141. Strimbu, K.; Tavel, J.A. What are biomarkers? *Curr. Opin. HIV AIDS* **2010**, *5*, 463–466. [[CrossRef](#)]
142. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34. [[CrossRef](#)]
143. Ashburner, M.; Ball, C.A.; Blake, J.; Botstein, D.; Butler, H.; Cherry, J.M.; Davis, A.P.; Dolinski, K.; Dwight, S.S.; Eppig, J.T.; et al. Gene Ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)] [[PubMed](#)]
144. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **1995**, *57*, 289–300. [[CrossRef](#)]
145. R. Team. Development Core Team. *R A Lang. Environ. Stat. Comput.* **2013**, *55*, 275–286.
146. Kuo, T.C.; Tian, T.F.; Tseng, Y.J. 3Omics: A web-based systems biology tool for analysis, integration and visualization of human transcriptomic, proteomic and metabolomic data. *BMC Syst. Biol.* **2013**, *7*, 64. [[CrossRef](#)]
147. Cambiaghi, A.; Ferrario, M.; Masseroli, M. Analysis of metabolomic data: Tools, current strategies and future challenges for omics data integration. *Brief. Bioinform.* **2016**, *18*, 498–510. [[CrossRef](#)] [[PubMed](#)]
148. Wold, S.; Ruhe, A.; Wold, H.; Dunn, I.W.J. The Collinearity Problem in Linear Regression. The Partial Least Squares (PLS) Approach to Generalized Inverses. *SIAM J. Sci. Stat. Comput.* **1984**, *5*, 735–743. [[CrossRef](#)]
149. Trygg, J.; Wold, S. Orthogonal projections to latent structures (O-PLS). *J. Chemom.* **2002**, *16*, 119–128. [[CrossRef](#)]
150. Vapnik, V.; Chervonenkis, A. Theory of Pattern Recognition. *Nauka*. 1974. Available online: <https://www.bibsonomy.org/bibtex/936f556afc966ddda07ba175241d6924> (accessed on 15 July 2021).
151. Tibshirani, R. The lasso method for variable selection in the cox model. *Stat. Med.* **1997**, *16*, 385–395. [[CrossRef](#)]
152. Hoerl, A.E.; Kennard, R.W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. *Technometrics* **1970**, *12*, 55. [[CrossRef](#)]
153. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2005**, *67*, 301–320. [[CrossRef](#)]
154. Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K.-A. mixOmics: An R package for ‘omics feature selection and multiple data integration. *PLoS Comput. Biol.* **2017**, *13*, e1005752. [[CrossRef](#)] [[PubMed](#)]
155. Thévenot, E.A.; Roux, A.; Xu, Y.; Ezan, E.; Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **2015**, *14*, 3322–3335. [[CrossRef](#)] [[PubMed](#)]
156. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F.; Chang, C.C.; Lin, C.C.; e1071: Misc Functions of the Department of Statistics (e1071), TU Wien. R Package Version 1. 6–3. Available online: <https://rdrr.io/rforge/e1071/> (accessed on 15 July 2021).
157. Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **2010**, *33*, 1–22. [[CrossRef](#)]
158. Hochrein, J.; Klein, M.S.; Zacharias, H.U.; Li, J.; Wijffels, G.; Schirra, H.J.; Spang, R.; Oefner, P.J.; Gronwald, W. Performance Evaluation of Algorithms for the Classification of Metabolic ¹H NMR Fingerprints. *J. Proteome Res.* **2012**, *11*, 6242–6251. [[CrossRef](#)] [[PubMed](#)]
159. Pearson, K.F.R.S. LIII. On lines and planes of closest fit to systems of points in space. *Lond. Edinb. Dublin Philos. Mag. J. Sci.* **1901**, *2*, 559–572. [[CrossRef](#)]

160. Kohonen, T. *Self-Organizing Maps*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
161. Wehrens, R.; Buydens, L.M.C. Self- and Super-organizing Maps in R: The kohonen Package. *J. Stat. Software* **2007**, *1*. [CrossRef]
162. Abdi, H.; Williams, L.J. Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 433–459. [CrossRef]
163. James, G.; Witten, D.; Hastie, T.; Tibshirani, R. *An Introduction to Statistical Learning*; Springer: New York, NY, USA, 2013; Volume 103.
164. Bictash, M.; Ebbels, T.; Chan, Q.; Loo, R.L.; Yap, I.K.; Brown, I.J.; de Iorio, M.; Daviglius, M.L.; Holmes, E.; Stamler, J.; et al. Opening up the "Black Box": Metabolic phenotyping and metabolome-wide association studies in epidemiology. *J. Clin. Epidemiol.* **2010**, *63*, 970–979. [CrossRef]
165. Horvath, S. *Weighted Network Analysis: Applications in Genomics and Systems Biology*; Springer: Berlin/Heidelberg, Germany, 2011.
166. Lauritzen, S.L. *Graphical Models*; Clarendon Press: Oxford, UK, 1996.
167. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef] [PubMed]
168. Schaefer, R.; Opgen-Rhein, K. Strimmer, GeneNet: Modeling and Inferring Gene Networks, (2015) R Package Version 1.2.13. Available online: <https://cran.microsoft.com/snapshot/2014-09-09/web/packages/GeneNet/index.html> (accessed on 15 July 2021).
169. Haslbeck, J.M.B.; Waldorp, L.J. Mgm: Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data. *J. Stat. Softw.* **2020**, *1*. [CrossRef]
170. Altenbuchinger, M.; Weihs, A.; Quackenbush, J.; Grabe, H.J.; Zacharias, H.U. Gaussian and Mixed Graphical Models as (multi-)omics data analysis tools. *Biochim. Biophys. Acta BBA Bioenerg.* **2020**, *1863*, 194418. [CrossRef] [PubMed]
171. Rosato, A.; Tenori, L.; Cascante, M.; Carulla, P.R.D.A.; Dos Santos, V.A.P.M.; Saccenti, E. From correlation to causation: Analysis of metabolomics data using systems biology approaches. *Metabolomics* **2018**, *14*, 37. [CrossRef] [PubMed]
172. Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *J. Stat. Softw.* **2010**, *36*, 1–48. [CrossRef]
173. Schwarzer, G. Meta: An R package for meta-analysis. *R News* **2007**, *7*, 40–45.
174. Gurevitch, J.; Koricheva, J.; Nakagawa, S.; Stewart, G. Meta-analysis and the science of research synthesis. *Nat. Cell Biol.* **2018**, *555*, 175–182. [CrossRef]
175. Cox, D.R. Regression Models and Life-Tables. *J. R. Stat. Soc. Ser. B* **1972**, *34*, 187–202. [CrossRef]
176. Ishwaran, H.; Kogalur, U.B.; Blackstone, E.H.; Lauer, M. Random survival forests. *Ann. Appl. Stat.* **2008**, *2*, 841–860. [CrossRef]
177. Therneau, T.M.; Grambsch, P.M. The Cox Model. In *Statistics for Biology and Health*; Springer Science and Business Media LLC: Berlin/Heidelberg, Germany, 2000; pp. 39–77.
178. Jager, K.J.; Van Dijk, P.C.; Zoccali, C.; Dekker, F.W. The analysis of survival data: The Kaplan—Meier method. *Kidney Int.* **2008**, *74*, 560–565. [CrossRef] [PubMed]
179. Stel, V.S.; Dekker, F.; Tripepi, G.; Zoccali, C.; Jager, K.J. Survival Analysis II: Cox Regression. *Nephron Clin. Pr.* **2011**, *119*, c255–c260. [CrossRef] [PubMed]
180. Nueda, M.J.; Conesa, A.; Westerhuis, J.A.; Hoefsloot, H.C.J.; Smilde, A.K.; Talón, M.; Ferrer, A. Discovering gene expression patterns in time course microarray experiments by ANOVA-SCA. *Bioinformatics* **2007**, *23*, 1792–1800. [CrossRef]
181. Jansen, J.J.; Hoefsloot, H.C.J.; Van Der Greef, J.; Timmerman, M.E.; Westerhuis, J.A.; Smilde, A.K. ASCA: Analysis of multivariate data obtained from an experimental design. *J. Chemom.* **2005**, *19*, 469–481. [CrossRef]
182. Zwanenburg, G.; Hoefsloot, H.C.; Westerhuis, J.A.; Jansen, J.J.; Smilde, A.K. ANOVA-principal component analysis and ANOVA-simultaneous component analysis: A comparison. *J. Chemom.* **2011**, *25*, 561–567. [CrossRef]
183. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 1–21. [CrossRef]
184. Smilde, A.K.; Westerhuis, J.A.; Hoefsloot, H.C.J.; Bijlsma, S.; Rubingh, C.M.; Vis, D.J.; Jellema, R.; Pijl, H.; Roelfsema, F.; van der Greef, J. Dynamic metabolomic data analysis: A tutorial review. *Metabolomics* **2010**, *6*, 3–17. [CrossRef] [PubMed]
185. Xia, J.; Wishart, D.S. MSEA: A web-based tool to identify biologically meaningful patterns in quantitative metabolomic data. *Nucleic Acids Res.* **2010**, *38*, W71–W77. [CrossRef] [PubMed]
186. Goeman, J.J.; van de Geer, S.A.; de Kort, F.; van Houwelingen, H.C. A global test for groups of genes: Testing association with a clinical outcome. *Bioinformatics* **2004**, *20*, 93–99. [CrossRef]
187. Picart-Armada, S.; Fernández-Albert, F.; Vinaixa, M.; Yanes, O.; Perera-Lluna, A. FELLA: An R package to enrich metabolomics data. *BMC Bioinform.* **2018**, *19*, 538. [CrossRef] [PubMed]
188. Al-Akwaa, F.M.; Yunits, B.; Huang, S.; Alhajaji, H.; Garmire, L.X. Lilikoi: An R package for personalized pathway-based classification modeling using metabolomics data. *GigaScience* **2018**, *7*. [CrossRef]
189. Nguyen, T.-M.; Shafi, A.; Nguyen, T.; Draghici, S. Identifying significantly impacted pathways: A comprehensive review and assessment. *Genome Biol.* **2019**, *20*, 203. [CrossRef] [PubMed]
190. Varma, S.; Simon, R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinform.* **2006**, *7*, 91. [CrossRef]
191. Becker, G.J.; Hewitson, T. Animal models of chronic kidney disease: Useful but not perfect. *Nephrol. Dial. Transplant.* **2013**, *28*, 2432–2438. [CrossRef] [PubMed]
192. Booth, S.; Weljie, A.M.; Turner, R.J. Computational tools for the secondary analysis of metabolomics experiments. *Comput. Struct. Biotechnol. J.* **2013**, *4*, e201301003. [CrossRef] [PubMed]

193. Abbiss, H.; Maker, G.L.; Trengove, R.D. Metabolomics Approaches for the Diagnosis and Understanding of Kidney Diseases. *Metabolites* **2019**, *9*, 34. [[CrossRef](#)] [[PubMed](#)]
194. Chen, Y.Y.; Chen, D.-Q.; Chen, L.; Liu, J.-R.; Vaziri, N.D.; Guo, Y.; Zhao, Y.-Y. Microbiome-metabolome reveals the contribution of gut-kidney axis on kidney disease. *J. Transl. Med.* **2019**, *17*, 1–11. [[CrossRef](#)]
195. Klein, J.; Jupp, S.; Moulos, P.; Fernandez, M.; Buffin-Meyer, B.; Casemayou, A.; Chaaya, R.; Charonis, A.; Bascands, J.-L.; Stevens, R.; et al. The KUPKB: A novel Web application to access multiomics data on kidney disease. *FASEB J.* **2012**, *26*, 2145–2153. [[CrossRef](#)]
196. Fernandes, M.; Husi, H. Establishment of an integrative multi-omics expression database CKDdb in the context of chronic kidney disease (CKD). *Sci. Rep.* **2017**, *7*, 40367. [[CrossRef](#)]
197. Papadopoulos, T.; Krochmal, M.; Cisek, K.; Fernandes, M.; Husi, H.; Stevens, R.; Bascands, J.-L.; Schanstra, J.P.; Klein, J. Omics databases on kidney disease: Where they can be found and how to benefit from them. *Clin. Kidney J.* **2016**, *9*, 343–352. [[CrossRef](#)]
198. Breit, M.; Weinberger, K.M. Metabolic biomarkers for chronic kidney disease. *Arch. Biochem. Biophys.* **2016**, *589*, 62–80. [[CrossRef](#)]
199. Davies, R. The metabolomic quest for a biomarker in chronic kidney disease. *Clin. Kidney J.* **2018**, *11*, 694–703. [[CrossRef](#)]
200. Trivedi, D.K.; Hollywood, K.A.; Goodacre, R. Metabolomics for the masses: The future of metabolomics in a personalized world. *Eur. J. Mol. Clin. Med.* **2017**, *3*, 294–305. [[CrossRef](#)] [[PubMed](#)]
201. Pinu, F.R.; Goldansaz, S.A.; Jain, J. Translational Metabolomics: Current Challenges and Future Opportunities. *Metabolites* **2019**, *9*, 108. [[CrossRef](#)] [[PubMed](#)]