Diagnostic yield of massively parallel sequencing in patients with chronic kidney disease of unknown etiology: rationale and design of a national prospective cohort study

Amber de Haan, Mark Eijgelsheim, Liffert Vogt, Bert van der Zwaag, Albertien M van Eerde, Nine V A M Knoers, Martin H de Borst

ABSTRACT

Introduction Chronic kidney disease (CKD) can be caused by a variety of systemic or primary renal diseases. The cause of CKD remains unexplained in approximately 20% of patients. Retrospective studies indicate that massively parallel sequencing (MPS)-based gene panel testing may lead to a genetic diagnosis in 12%–56% of patients with unexplained CKD, depending on patient profile. The diagnostic yield of MPS-based testing in a routine healthcare setting is unclear. Therefore, the primary aim of the VARIETY (Validation of algoRithms and IdEnTification of genes in Young patients with unexplained CKD) study is to prospectively address the diagnostic yield of MPS-based gene panel testing in patients with unexplained CKD and an estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m² before the age of 50 years in clinical practice.

Methods and analysis The VARIETY study is an ongoing, prospective, nationwide observational cohort study to investigate the diagnostic yield of MPS-based testing in patients with unexplained CKD in a routine healthcare setting in the Netherlands. Patients are recruited from outpatient clinics in hospitals across the Netherlands. At least 282 patients will be included to meet the primary aim. Secondary analyses include subgroup analyses according to age and eGFR at first presentation, family history, and the presence of extrarenal symptoms.

Ethics and dissemination Ethical approval for the study has been obtained from the institutional review board of the University Medical Center Groningen. Study findings should inform physicians and policymakers towards optimal implementation of MPS-based diagnostic testing in patients with unexplained CKD.

INTRODUCTION Chronic kidney disease (CKD) affects 11%–16% of the population worldwide and is associated with extensive comorbidity and an increased risk of premature mortality, and may ultimately result in end-stage kidney disease (ESKD) requiring dialysis or transplantation. CKD may be caused by a variety of systemic (eg, diabetes, hypertension) or primary renal diseases (eg, IgA nephropathy, membranous nephropathy). Current diagnostic approaches, including kidney biopsy, are often non-specific or inconclusive, contraindicated or omitted due to lack of clinical consequences. Therefore, the cause of CKD remains unknown in approximately 20% of patients with ESKD. However, knowledge of the underlying kidney disease can be pivotal as it may influence prognosis and medical treatment. In the setting of kidney transplantation, it may influence (living-related) donor selection and post-transplant recurrence risk. Approximately 27%–34% of patients with CKD report a positive family history of kidney disease (first or second degree relative with CKD) and a genetic cause can be identified in at least 10% of adults with CKD, indicating that in many cases a hereditary origin for the disease should be considered. Genetic testing could therefore be a valuable tool in the diagnostic process of CKD of unknown aetiology.
Recent studies suggest that massively parallel sequencing (MPS) techniques (previously referred to as next-generation sequencing)\(^ {15}\) could be used as diagnostic tool in adults with unexplained CKD and should even be considered as first mode of diagnostics in patients with ESKD prior to the age of 50 years.\(^ {16}\) Depending on patient selection, MPS led to a genetic diagnosis in 12%–56% of patients with unexplained CKD.\(^ {14} 17–19\) However, most of these studies have been performed in a research setting, and therefore little is known about the diagnostic utility of MPS for adults with unexplained CKD in a routine healthcare setting. Moreover, currently available studies have been commonly based on subgroups of larger retrospective cohorts, and are heterogeneous in design and selection of genes used in MPS.\(^ {20}\) For this reason, it is difficult to define profiles of patients (eg, based on age and severity at disease onset, extrarenal manifestations, positive family history) that should preferentially undergo genetic testing. A recent joint publication by the ERA Working Group on Inherited Kidney Disorders and the Molecular Diagnostics Taskforce of the European Rare Kidney Disease Reference Network called for further research to explore the diagnostic yield of genetic testing in CKD of unknown origin in a clinical setting.\(^ {21}\)

Therefore, the objective of this national prospective cohort study is to determine the diagnostic yield, that is, the percentage of participants with a genetic diagnosis, using a large MPS-based multigene panel for kidney diseases in young patients (first presentation at age <50) with unexplained CKD (estimated glomerular filtration rate (eGFR) <60 mL/min/1.73 m\(^ 2\)) in a routine healthcare setting. In addition, we aim to identify specific patient profiles with a high diagnostic yield. These findings can guide physicians and policymakers in implementing MPS-based diagnostic testing in patients with unexplained CKD.

**METHODS AND ANALYSIS**

**Study design**

The VARIETY (Validation of algoRithms and IdEnTiFication of genes in Young patients with unexplained CKD) study is a prospective nationwide observational cohort study designed to investigate the diagnostic yield of genetic testing in patients with unexplained CKD in a routine healthcare setting in the Netherlands. The study will collect and analyse data obtained during routine clinical practice and through a questionnaire. Participants will be included from both academic and non-academic hospitals throughout the Netherlands. The anonymised data are collected, stored and analysed in the University Medical Center Groningen (UMCG). All participants will give written informed consent on enrolment.

**Study population**

The targeted study population consist of all patients with unexplained CKD and an eGFR <60 mL/min/1.73 m\(^ 2\) before the age of 50 years. Unexplained CKD is defined as the absence of all the following criteria: a biopsy-proven diagnosis (eg, IgA nephropathy), a specific morphological renal diagnosis (eg, polycystic kidney disease suspected of autosomal/recessive polycystic kidney disease), or a specific or plausible renal diagnosis (eg, history of long-term insulin-dependent diabetes mellitus before the onset of CKD, lithium-induced nephropathy). Since hypertensive nephropathy is a nonspecific diagnosis and hypertension is also a very common consequence of CKD,\(^ {22}\) patients with hypertensive nephropathy in the absence of a clear underlying disorder such as renal artery stenosis are considered to have unexplained CKD. Patients with renal hypoplasia, renal atrophy, and nonspecific histological conditions (such as secondary focal segmental glomerulosclerosis, glomerulonephritis of unknown cause, or interstitial nephritis) are also considered to have unexplained CKD.

Patients with a current age >50 years, but who presented with an eGFR <60 mL/min/1.73 m\(^ 2\) before the age of 50 years, and renal transplant recipients who had a pre-transplant eGFR <60 mL/min/1.73 m\(^ 2\) before the age of 50 years are also eligible for inclusion. In addition, genetic testing with a specific MPS-based gene panel (see ‘Genetic testing’) is required for participation. Exclusion criteria for participation in the VARIETY study are: age <18 years at time of inclusion or patients who do not give or are unable to give informed consent for genetic testing or for the current study.

**Recruitment**

To ensure a representative sample of CKD patients, patients are recruited from outpatient clinics in both academic and non-academic hospitals across the Netherlands. Depending on the hospital, patients will be screened by the primary treating nephrologists or by trained study investigators. In case of a study investigator, a list with potential participants will be sent to the treating nephrologist to confirm the diagnosis of unexplained CKD. Eligible participants will be informed about the study by the investigators or their treating nephrologists aware of the study protocol. A study investigator or treating nephrologist will ask for informed consent for this study. Information for patients has been made available in the form of a patient information folder and a website (in Dutch): www.onbegreppennierziekte.nl.

**Data collection**

Detailed clinical and demographic data are collected from patients’ electronic health record (EHR) and through a questionnaire following informed consent. The data will subsequently be entered into a secure electronic case report form.

**Electronic health record**

The following information will be collected from the EHR: age at inclusion, sex, primary renal disease diagnosis, age at CKD onset/presentation, dialysis or kidney
transplantation, age at start dialysis or kidney transplantation, presence of extrarenal features, medication use at inclusion, medical history, family history (including three-generation pedigree), blood pressure at CKD onset and at inclusion, presence of haematuria and/or nephrotic syndrome, laboratory results (serum creatinine, eGFR, total cholesterol, high-density lipoprotein (HDL), low-density lipoprotein (LDL), triglycerides, 24-hour urine albumin excretion, 24-hour urine creatinine excretion, 24-hour urine total protein excretion, haematuria) at CKD onset/presentation and inclusion, renal histopathology and imaging of the native kidneys, and results of genetic testing in relation to kidney disease. We will also collect information regarding the clinical consequences of a genetic diagnosis and if genetic counselling was performed by a nephrologist or clinical geneticist. If participants are referred to a clinical geneticist, we will also collect the results of any additional genetic testing.

**Questionnaire**

Data collected from the EHR will be expanded with a questionnaire to collect additional data on family history, medical history, current health complaints and extrarenal manifestations (online supplemental table 1).

**Genetic testing**

We will include patients who have undergone MPS-based multigene panel testing, initiated by a clinical geneticist or nephrologist following pretest counselling as part of clinical care in patients with unexplained CKD, in accordance with guidelines in the Netherlands. Figure 1 shows the suggested flowchart for genetic testing in the VARIETY study, based on these recommendations. The criteria as shown in this flow chart are slightly more liberal than the published recommendations, which will help to define the optimal age and eGFR ranges where genetic testing is still of clinical benefit. To stimulate the implementation of the guideline, we made a website for the VARIETY study (www.onbegreppenierziekte.nl) where nephrologists can find information about genetic testing and pretest genetic counselling.

In order to reduce heterogeneity in the diagnostic approach, we will assess the diagnostic yield of a specific MPS-based gene panel, namely the ‘CKD-Y’ (‘CKD in Young patients’) targeted exome sequencing (ES) panel available at the University Medical Center (UMC) Utrecht, The Netherlands. The older version of this panel (v18) contains 141 different genes associated with early-onset CKD, including PKD1 and PKD2 (figure 2). On 8 March 2021, the CKD-Y panel was updated (v21) and the number of genes changed from 141 to 256 (figure 3). This panel was chosen as it is an ES-based panel and contains all the current genes associated with early-onset CKD. In addition, this panel can be ordered by nephrologists without referral to the clinical geneticist. Alternatively, the hereditary kidney disease panel from UMC Utrecht was allowed. This is another targeted ES panel, consisting of 379 genes in the v18 version and 495 genes in the updated v21 version (online supplemental figures 1-2). Since this panel contains some kidney cancer oncogenes, it may only be ordered by a clinical geneticist. The hereditary kidney disease panel includes all genes of the CKD-Y panel, making it possible to determine if a variant could also have been identified with the CKD-Y panel. Potential findings from the hereditary kidney disease

**Figure 1** Flow chart for inclusion in the VARIETY study. CKD, chronic kidney disease; eGFR, estimated glomerular filtration rate; VARIETY, Validation of algoRithms and IdEnTification of genes in Young patients with unexplained CKD.
panel that do not overlap with the CKD-Y panel will not be included in the primary analyses. Patients with older versions of the CKD-Y and hereditary kidney disease panels can be included. We will record which version of the CKD-Y and/or hereditary kidney disease panel was used. The procedures for ES and variant filtering have been described before. Copy number variation (CNV) detection was performed using an in-house adapted, diagnostically validated, version of the ExomeDepth CNV detection tool.

Primary and secondary analyses

The primary analysis will address the diagnostic yield of the CKD-Y panel, defined as the percentage of positive test results (ie, pathogenic variant(s) explaining the cause of the disease), in the overall cohort of patients with unexplained CKD and an eGFR <60 mL/min/1.73 m² before the age of 50 years. We will perform a sensitivity analysis in patients with onset eGFR <60 mL/min/1.73 m² between the age of 18 and 50 years. The pathogenicity of variants will be determined according to the standards and guidelines from the American College of Medical Genetics and Genomics. With this standard, variants are classified into five categories using several lines of evidence, such as available literature, patient databases and in silico prediction programmes. Class one variants are clearly not pathogenic; class five variants are clearly pathogenic. Class three are variants of uncertain significance/pathogenicity (VUS), these variants do not confirm or exclude the diagnosis (table 1). For the determination of the diagnostic yield, only class 4 and class 5 variants will be considered as a ‘positive test result’ to determine the diagnostic yield. In cases with two class 4/5 variants in an autosomal recessive gene, these will only be considered a ‘positive test result’ if testing in parents has confirmed the variants are positioned in trans.

Secondary analyses include subgroup analyses according to age and eGFR at first presentation, family history and the presence of extrarenal symptoms. A positive family history for CKD is recorded if the participant either has a first (parent or child), second (siblings, grandparents, grandchildren), third (aunts, uncles, nephews, nieces) or fourth (cousins) degree relative with CKD. Family history will be obtained from combining information present in the EHR with information obtained from the questionnaire. Other secondary analysis aims to define the percentage of genetic tests with a clinical consequence. A genetic diagnosis is considered to have a clinical consequence if it: (1) negated the need for kidney biopsy, (2) triggered or negated the need for immunosuppressive therapy.
therapy, (3) provides prognostic information, that is, the risk of post-transplantation anti-glomerular basement membrane (GBM) nephritis, (4) led to, or should lead to, referral to other specialties (eg, ophthalmologist), (5) led to targeted work-up for associated symptoms or extra-renal manifestations, (6) affected surveillance frequency, (7) led to, or should lead to, genetic testing in potential living related kidney donors, (8) enabled more precise (preconception) genetic counselling for the patient or family members or (9) led to more precise or extensive follow-up of potentially affected family members.

Tertiary outcomes will be the percentage of participants in which a VUS was identified and the number of incidental/secondary findings (results unrelated to the initial indication for genetic testing). In addition, if a molecular diagnosis is identified and a kidney biopsy from the native kidney is present, we will assess if the biopsy findings match the molecular diagnosis. Finally, we will perform health economic analyses to determine if MPS in patients with unexplained CKD is cost-effective.

We will report the number of participants who withdraw from participating in the VARIETY study after initial inclusion and the number of participants who have initially been included, but on further analysis by the study team did not match the inclusion criteria. Participants without results from genetic testing will be excluded from all analyses. If information is missing from the EHR, we will ask the general practitioner to deliver the missing data within participants’ consent. If the data cannot be retrieved, it will be regarded as ‘unknown’. In case eGFR at CKD onset is missing, the first available eGFR or serum creatinine measurement since the diagnosis of CKD will be used.

**Statistical analysis**

Statistical analysis will be performed with IBM SPSS statistics for Windows, V.23 (IBM). An overall significance level of 0.05 will be handled.

Continuous variables that are normally distributed will be presented as mean and standard deviation. Non-normally distributed variables will be expressed as median and interquartile range. Frequencies and percentages will be used to describe categorical variables such as gender, family history, renal replacement therapy, extrarenal manifestations, and diagnostic yield. The \( \chi^2 \) or Fisher’s exact test will be used to compare differences in categorical variables between the different subgroups of the secondary analysis. Logistic regression will be performed to identify characteristics associated with a genetic diagnosis.

**Sample size calculation**

The minimal sample size was calculated using the following formula,\(^2^{7}\) based on the study’s primary endpoint:

\[
n = \frac{Z^2 \times \text{P}(1-P)}{d^2} 
\]

Based on the literature, the expected percentage of positive test results is 17%.\(^4^{1}\) Assuming a level of confidence (\( Z \)) of 1.96 and precision (\( d \)) of 0.05,\(^2^{7}\) a minimum of 217 participants are required for a reliable assessment of the primary outcome. In order to be clinically and politically significant, we aim to increase the sample size of this prospective cohort study beyond the largest currently available retrospective study, that is, to include at least 282 patients in the current study.\(^2^{0,27}\)

**Data management**

Study data will be recorded digitally using the secure REDCap (Research Electronic Data Capture) web application (REDCap, Nashville, Tennessee, USA) hosted at the UMCG.\(^2^{8,29}\) Data collection and entry is performed by trained investigators from the UMCG. To minimise differences and errors in data entry, investigators from the UMCG will travel to other participating centres for data collection and entry in REDCap. Data validation in REDCap will be performed according to a data validation plan, which has been made in collaboration with the UMCG Research Data Support and approved by the Institutional Review Board. Data analysis will take place on validated and anonymised data. On study closure, data will be extracted from REDCap and exported to SPSS for analysis.

**Patient and public involvement**

Patients and/or the public were not involved in the design of this study.

**ETHICS AND DISSEMINATION**

Ethical approval for the study has been obtained from the institutional review board of the UMCG (METc 2019/106). The study is conducted in accordance with the WMA Declaration of Helsinki. The results of the study will be presented at (inter)national congresses and submitted for open access publication in peer-reviewed journals. In addition to the primary results, related to the main research questions as defined above, case reports/series may be submitted for publication in case of unique or interesting findings and these will also be submitted for publication in peer-reviewed journals. In accordance with the information sheet for participants, the main results and any publications from the VARIETY study will also be made available on the study website. After completion of the study and publication of the main results, request for re-use of the data can be submitted to the corresponding author.

**CONCLUSION AND STUDY STATUS**

Genetic testing shows promising results as a diagnostic tool in adults with CKD and it has the potential to resolve CKD cases with an unknown aetiology. However, further research is needed in a clinical setting to define the position of MPS-based diagnostics in clinical practice.
248 patients have been included. Inclusion started on 31 July 2019. As of September 2021, physicians and policymakers involved in implementational (guideline) committees or institutions may slightly explained CKD has not been uniformly defined by international core curriculum 2019. 


Author affiliations
1Department of Internal Medicine, Division of Nephrology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands
2Department of Internal Medicine, Section Nephrology, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, The Netherlands
3Department of Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
4Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Acknowledgements
Several authors of this publication (AMVE, NVAMK and MdB) are members of the European Reference Network for Rare Kidney Diseases (ERKNet) - Project ID no. 739 532.

Contributors
AdH and MdB wrote the first draft of the manuscript. ME, LV, BvdZ, AMvE and NVAMK gave feedback and contributed to manuscript revision. All authors read and approved the submitted version.

Funding
This collaboration project is cofinanced by Sanofi Genzyme and the Dutch Ministry of Economic Affairs and Climate Policy by means of the PPP Allowance made available by the Top Sector Life Sciences & Health to stimulate public-private partnerships (grant numbers RVO/6320 and IMAGEN/LSHM20009).

Competing interests
MdB and LV have received research support and lecture fees (all to institution) from Sanofi Genzyme related to the current study. NVAMK has received reimbursement of travel expenses for lectures related to the current study from Sanofi Genzyme.

Patient and public involvement
Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication
Not applicable.

Provenance and peer review
Not commissioned; externally peer reviewed.

Supplemental material
This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have beenPeer reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access
This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD
Amber de Haan http://orcid.org/0000-0003-4228-6927

REFERENCES
