The Biobank of Nephrological Diseases in the Netherlands cohort: the String of Pearls Initiative collaboration on chronic kidney disease in the university medical centers in the Netherlands

Gerjan J. Navis1, Peter J. Blankestijn2, Jeroen Deegens3, Johan W. De Fijter4, Jaap J. Homan van der Heide5, Ton Rabelink4, Raymond T. Krediet8, Arjan J. Kwakernaak1, Gozewijn D. Laverman1, Karel M. Leunissen6, Pieter van Paassen6, Marc G. Vervloet7, Pieter M. Ter Wee7, Jack F. Wetzels3, Robert Zietse8 and Frans J. van Ittersum7 on behalf of the BIND-NL investigators

1Departments of Nephrology, University Medical Center Groningen (UMCG), Groningen, The Netherlands, 2University Medical Center (UMC), Utrecht, The Netherlands, 3University Medical Center Nijmegen (UMCN), Nijmegen, The Netherlands, 4Leiden University Medical Center (LUMC), Leiden, The Netherlands, 5Academic Medical Center Amsterdam (AMC), Amsterdam, The Netherlands, 6Maastricht University Medical Center (MUMC), Maastricht, The Netherlands, 7Free University Medical Center (VUMC), The Netherlands and 8Erasmus Medical Center (Erasmus MC), Rotterdam, The Netherlands

*Correspondence and offprint requests to: Gerjan J. Navis; E-mail: g.j.navis@umcg.nl

ABSTRACT

Despite advances in preventive therapy, prognosis in chronic kidney disease (CKD) is still grim. Clinical cohorts of CKD patients provide a strategic resource to identify factors that drive progression in the context of clinical care and to provide a basis for improvement of outcome. The combination with biobanking, moreover, provides a resource for fundamental and translational studies. In 2007, the Dutch government initiated and funded the String of Pearls Initiative (PSI), a strategic effort to establish infrastructure for disease-based biobanking in the University Medical Centres (UMCs) in the Netherlands, in a 4-year start-up period. CKD was among the conditions selected for biobanking, and this resulted in the establishment of the Biobank of Nephrological Diseases-NL (BIND-NL) cohort. Patients with CKD Stages 1–4 are eligible. The data architecture is designed to reflect routine care, with specific issues added for enrichment, e.g. questionnaires. Thus, the collected clinical and biochemical data are those required by prevailing guidelines for routine nephrology care, with a minimal dataset for all patients, and diagnosis-specific data for the diagnostic categories of primary and secondary glomerular disorders and adult dominant polycystic kidney disease, respectively. The dataset is supplemented by a biobank, containing serum, plasma, urine and DNA. The cohort will be longitudinally monitored, with yearly follow-up for clinical outcome. Future linking of the data to those from the national registries for renal replacement therapy is foreseen to follow the patients’ lifeline throughout the different phases of renal disease and different treatment modalities. In the design of the data architecture, care was taken to ensure future exchangeability of data with other CKD cohorts by applying the data harmonization format of the Renal DataSHaPER, with a dataset based upon standardized indicator sets to facilitate collaboration with other CKD cohorts. Enrolment started in 2010, and over 2200 eligible patients have been enrolled in the different UMCs. Follow-up of enrolled patients has started, and enrolment will continue at a slower rate. The aggregation and standardization of clinical data and biosamples from large
numbers of CKD patients will be a strategic resource not only for clinical and translational research, but also by its basis in routine clinical care for clinical governance and quality improvement projects.

**Keywords:** biobank, BIND-NL, renal data shaper, String of Pearls Initiative, PSI; chronic kidney disease

### Introduction

Despite advances in detection and therapy, the prognosis in chronic kidney disease (CKD) is still grim. Moreover, the incidence and prevalence of CKD are increasing worldwide, which is associated with a high burden of disease, both in terms of human suffering and health care expenditure [1]. The costs of treatment for end stage renal disease alone have been estimated at 0.7–1.8% of national expenditures in European countries [2], and the costs of earlier stages of CKD may amount to similar figures [3]. Better understanding of the forces driving CKD, its progression and complications, is therefore urgently needed to improve management and prognosis. Clinical cohorts of CKD patients, providing longitudinal data on the clinical course in CKD patients are a strategic resource to this purpose. Moreover, combining clinical datasets with a biobank will allow the possibility to benefit from the advances in fundamental science and technology, including genetics, genomics [4], proteomics [5], metabolomics and more, thus providing a resource to translate fundamental progress into clinical benefit for CKD patients. As epidemiology of CKD is geographically diverse, it is important to have well-documented biobanks from different regions, as extrapolation from one cohort to another may not be warranted. Here, we describe the background and organization of the recently established PSI BIND-NL cohort, a large national CKD cohort plus biobank from the Netherlands.

### The Parelnoer-string of Pearls Initiative: PSI

In 2007, the Dutch government provided a start-up grant for the National Parelnoer-String of Pearls Initiative (PSI), with the metaphor of Parelnoer, Dutch for String of Pearls, symbolizing the collaboration between the different University Medical Centres (UMCs). ([www.parelnoer.org](https://www.parelnoer.org); [www.string-of-pearls.org](https://www.string-of-pearls.org)) PSI was funded as a strategic project of the Netherlands Federation of UMCs (NFU), the network of the eight Dutch UMCs, to establish a series of collaborative disease-based biobanks, for selected medical conditions, including CKD, with participation of all UMCs. The other selected disease conditions include cerebrovascular accident, diabetes mellitus, inflammatory bowel disease, malignant lymphoma, neurodegenerative disease, rheumatoid arthritis, artherosclerosis and hereditary colon carcinoma and more are in the process of being added. The overall aim of the PSI start-up project was to establish a durable collaborative infrastructure for clinical and translational scientific purposes for the UMCs, and more specifically to (i) provide a resource for translational science, by up-scaling and harmonization of biobanking, thus fuelling collaboration between UMCs; (ii) provide infrastructure for quality of care improvement, by use of data extracted from the clinical care process and (iii) stimulate standardization of clinical care processes by use of standardized electronic patients files and harmonization of clinical protocols.

The disease-based biobanks share a central infrastructure that serves to connect the de-centralized data collections into the respective disease-based data banks by an integrated Information and Communication Technology (ICT) system that supports the development of the specific technical as well as ethical and legal expertise required for large-scale biobanking. The collection of data and biosamples is based on Standard Operating Procedures (SOP) within the routine clinical care of the different UMCs, which will allow the use of the data collections for quality development systems, as well as for registry purposes. All data are pseudonymized. The structure and workflow of the PSI cohorts, including the BIND-NL cohort, is shown in Figure 1, depicting how clinical and biochemical data from the different UMCs are being entered from local sources into the central database infrastructure, whereas the biobank sampling as well as processing and storage take place locally.

By this initiative, the NFU and the government acknowledge the enormous potential of rapidly expanding expertise on large-scale biobanking that was mainly developed in general population cohorts, and aim to extend this dynamic field of expertise to the domain of patient cohorts. The importance of well-documented long-term follow-up of cohorts from the general population and patient cohorts has stood out for decades. The recent developments in genetics, requiring sample sizes that were unheard of so far, have provided considerable impetus to the domain of data banking and biobanking, with specific emphasis on the requirements to allow meaningful large-scale collection and pooling or exchange of data and biosamples from different sources, such as harmonization of data description and biosample processing [6]. Carrying this expertise to the realm of clinical nephrology is an enormous challenge, but will be of great potential value by enabling the generation of new insights from large-scale datasets generated from clinical practice.

The start-up funding of the PSI program, covering the period of 2007–11 allowed the establishment of the local and central infrastructure, and the inclusion of eight disease-based baseline cohorts. After a review by the NFU in autumn 2011, a 4-year extension period was granted. During this extension period, the PSI infrastructure should be fully integrated into the routine care processes of the UMCs, thus denoting state-of-the-art biobanking as a distinguishing feature of clinical care in UMCs. Moreover, this integration should ensure future inclusion and follow-up at minimal costs.

### BIND-NL: Design and Current Status

Patients with CKD Stages 1–4 attending outpatient nephrology clinics of all eight UMCs in the Netherlands are invited to participate. All patients with Stage 1–4 are eligible, the main exclusion criteria being inability to provide informed consent, non-CKD and circumstances in which of loss of follow-up is
very likely, such as a life expectancy <6 months. The nephrology departments of the UMCs in the Netherlands generally function as a second- and third-line referral centre, providing basic as well as sophisticated patient care, depending on the regional circumstances. The distribution of the CKD patients over primary care, private practice nephrology and the UMCs in the Netherlands has not been quantified so far, but it can safely be assumed that the majority of patients with early stages of CKD are under care of their general practitioner. The BIND-NL cohort, by its broad inclusion criteria, thus will contain a diversity of patients, generally those referred to the nephrologist either because of a complex primary condition (e.g. glomerulonephritis), or because the CKD turned out to be a progressive condition during medical supervision in primary care. Whereas most patients at UMC outpatient clinics are also being followed at the UMC, also a proportion of the outpatients returns to their referring nephrologists in private practice or primary care after the work-up in the UMC. This will have to be accounted for in the collection of follow-up and outcome data.

The study is conducted according to the principles of Good Clinical Practice. Clinical data are collected (see Table 1), according to a core dataset for all patients which includes all data and clinical indicators required by clinical guidelines and hence corresponds to clinical routine. The core dataset is complemented by diagnosis-specific items for specific predefined diagnostic categories (profiles), namely primary glomerular disease (PGD), secondary glomerular disease (SGD), and adult polycystic kidney disease (APKD), respectively. For PGD and SGD, the renal disease must be proven by biopsy and in APKD patients must comply with the Ravine criteria. For the baseline cohort, the data bank is enriched by a number of relevant issues not part of the clinical routine, such as the Charlson index for comorbidity, the SF-12 health index and ankle-brachial index. Follow-up data are collected yearly based on routine clinical care.

Follow-up data include updated medical history, including cardiovascular events, current medication use and morphometric and laboratory measures (Table 1). Main outcome parameters are start and mode of renal replacement therapy, the (cardiovascular) complications of CKD,
hospitalizations and hospitalization diagnoses, mortality and cause of death, as well as health care costs. For patients remaining under UMC care, clinical and biochemical data and biosamples are collected at the outpatient clinic. For patients that have returned to private practice or routine care, only clinical data will be collected during follow-up, wherever possible. To reliably assess outcome data in patients no longer under UMC care, links are being established with other databases and registries available in the Netherlands with full national coverage on hospitalization and hospitalization diagnoses (www.dutchhospitaldata.nl), entry in renal replacement therapy (www.renine.nl) and death (www.cbs.nl), respectively.

To facilitate future collaboration with cohorts in the Netherlands and abroad, interoperability and exchange of data, the data architecture is designed according to the format of the Renal DataSHaPER (www.regenet.eu) (Data Schema and Harmonization Platform for Epidemiological Research). The Renal DataSHaPER is a data harmonization format for renal cohorts, developed as part of the GENECURE FP6 project (www.genecure.eu) 7 in collaboration with P3G (Public Population Project in Genomics; www.p3g.org), as the first disease-specific elaboration of the generic Data-SHaPER, a data harmonization tool for general population cohorts developed by P3G to enable the (very) large-scale collaboration between multiple cohorts required to achieve sufficient power for genome-wide association study (GWAS) studies [6].

**Biobanking**

Biobanking is performed locally at the participating centres, with de-central sampling, sample tracking and storage, linked together by SOPs across the different UMCs. All centres use the same SOPs for collecting, handling, identification and storage of biomaterial. Stored materials include EDTA-plasma, citrate-plasma, heparin-plasma and serum. All these materials are aliquoted in several (minimal five per type of material) 0.5 mL tubes and stored at −80°C within specified time-frames. In addition a minimum of five aliquots of 1 mL urine taken from a stirred and timed 24-h urine collection is stored as well. In dedicated centres, urine is collected and stored after addition of protease inhibitors for exosome studies. Finally, a separate EDTA sample is stored at −80°C for DNA isolation of those consenting subjects.

**Current status**

At the time of this writing, over 2100 patients have been included under the start-up funding. During the transition to incorporation in regular care, the inclusion will continue at a lower rate. Depending on the success of transition to regular care, and available future funding, inclusion will continue. Follow-up is likewise aimed to be ongoing (as it is in clinical care), and not limited to a fixed period of time.

**PERSPECTIVES**

Over the last two decades substantial insights on the factors driving CKD and its complications have been obtained from data from large randomized clinical trials (RCT), demonstrating the effect of pharmacological interventions, thus providing a basis for current evidence-based treatment in CKD [7, 8]. Yet, despite these treatments and despite their proven efficacy, prognosis in CKD is still grim. This could well relate to interaction of therapy with lifestyle factors, such as sodium and phosphate intake, on the therapeutic efficacy [9, 10], but may also be inherent to the differences between the standardized setting of the RCTs and the much more heterogeneous setting of daily clinical practice. Data from non-selected cohorts of CKD patients are therefore of utmost importance to increase our understanding of the factors driving progression of CKD and its complications.

Whereas several very large-scale studies and meta-analyses on renal phenotypes are available, with data on renal phenotypes obtained mainly from the general population [11, 12], it has been pointed out that the number of true CKD patients available in well-documented cohorts is still modest, in particular with regard to the availability of biosamples [13]. In the USA, the Chronic Renal Insufficiency Cohort (CRIC) includes over 3000 patients of diverse ethnic background [14] and, in Japan, the Chronic Kidney Disease Japan Cohort (CKD-JAC) has included 3000 patients [15]. As the epidemiology of CKD is geographically diverse, with respect to genetic and environmental background as well as the organization of health care, extrapolation of one national setting to the other is unlikely to be justified. It would be important to have data from different cohorts from different settings, first, to provide data for the specific geographic setting, and second, to be able to investigate for generalizability and independent replication of data by combined analyses of data from different cohorts. In Europe, recently, the German Chronic Kidney Disease Cohort (G-CKD) was established, which included 5000 CKD patients with moderate reduction of GFR and/or proteinuria at enrolment, with a design resembling the BIND-NL cohort [13]. Moreover, in France, the CKD-REIN cohort, striving for representativeness of the CKD population in France is about to start its inclusion, aiming for 3600 patients with eGFR <60 mL/min [16]. Taken together with already available European CKD cohorts, such as the Masterplan cohort (n = 793) [17] and the NephroTest cohort (n=1038) [18], this can provide substantial critical mass for collaborative clinical studies on CKD in Europe.

Considering the rapid developments, in particular in genetics and proteomics, the availability of such resources is highly relevant to test the prognostic and clinical significance of novel findings in the setting of CKD in clinical practice. For instance, GWAS identified numerous loci associated with differences in eGFR in the general population [4, 19]. To prioritize for follow-up studies among the many loci, it would be important to assess whether loci are also associated with CKD, progressive renal function loss and an increased risk for end-stage renal disease. As to the latter, it is relevant that excess cardiovascular mortality acts as a competing risk (see Figure 2): the true relationships can thus be unravelled only by longitudinal data. Availability of environmental data, including nutrition and medication, is not only relevant in themselves, but also to assess modifiability of genetically conferred risk. In this respect, the value of 24-h urine as a source of non-biased
information on pivotal nutrition factors, such as sodium and phosphate, cannot be overestimated.

The BIND-NL group will take up studies investigating the determinants of progressive renal function loss and its complications in this particular cohort, and is designed to be open to adjunct studies, both in its own right and in a collaborative setting. For long-term follow-up of the CKD patients, moreover, collaboration will be established with the registries for dialysis (RELINE: www.reline.nl) and transplantation (NOTR) in the Netherlands, respectively, to enable a lifelong follow-up of the patients across the different stages of their renal condition.

The scientific program addresses several domains, i.e. biomedical, psychosocial, quality of care/clinical governance and health economics. Main topics within these domains are provided in Table 2.

### ACCESSIBILITY OF THE DATA

The PSI BIND-NL cohort seeks to promote collaborative projects. Interested investigators, be it or not participants of the BIND-NL can submit project proposals. A scientific committee with representatives of all participating centres will review applications for access to the data and biosamples. The applications will be reviewed for compliance with the overriding aim of promoting benefit for the renal patient, and for the match with our scientific program, as well as for scientific quality, novelty, feasibility and the suitability of the PSI BIND-NL cohort for the study question.

### LIMITATIONS

The data are obtained from routine clinical care from the UMCs: this poses the limitations of the routine care setting as well as the issue of patient selection, with allegedly more severe cases of CKD, and over-representation of the primary renal disorders. Whereas this selected population is worthwhile investigating in its own right because of the burden of disease it represents, for better interpretation of the data from the cohort, it will be important to assess the impact of selection by proper patient characterization, and comparison to the available data on CKD in the Netherlands, i.e. data on the patients entering renal replacement therapy (www.reline.nl) and on CKD in large, unselected general population cohorts, such as Lifelines [20], and moreover, with data from abroad. Albeit not perfect, this will allow to assess the effects of selection for the BIND-NL cohort.
The strength of the BIND-NL cohort is its basis in clinical care for a broad range of CKD patients, with its longitudinal follow-up, embedded in state of the art expertise on biobanking, as provided by PSI. The aggregation and standardization of clinical data and biosamples from large numbers of CKD patients will be a strategic resource for translational research, identifying novel pathways of disease, allowing better predictive models, and identifying and prioritizing novel targets for intervention. Moreover, it will provide data for clinical governance initiatives, including evaluation of coherence of clinical decision-making, and evaluation of implementation of established guidelines and its association with outcome. Thus, by multifaceted studies rooted in clinical care, the BIND-NL cohort will provide a long-term resource for improving outcome in CKD [21–23].

ACKNOWLEDGEMENT

PSI was funded by the Dutch Government (FES-funds), the NFU and the 8 UMCs of The Netherlands.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES


Received for publication: 28.2.2013; Accepted in revised form: 6.6.2013