KidCAD: An Interactive Cohort Analysis Dashboard of Patients with Chronic Kidney Diseases

M. Höhn1, S. Schwindt2, S. Hahn2, S. Patyna3, S. Büttner1,5, J. Kohlhammer1,2

1Fraunhofer IGD, Darmstadt, Germany 2TU Darmstadt, Germany 3Department of Internal Medicine IV, Nephrology, University Hospital, Goethe-University, Frankfurt am Main, Germany 4Medical Clinic III - Department of Nephrology, University Hospital Frankfurt, Frankfurt am Main, Germany 5Medical Clinic I - Cardiology, Pneumology, Nephrology and Intensive Care Medicine, Klinikum Aschaffenburg-Alzenau, Aschaffenburg, Germany

Abstract
Chronic Kidney Diseases (CKD) are a prominent health problem. With an ongoing process, CKD leads to impaired kidney function with decreased ability to filter the patients’ blood, concluding in multiple complications, like heart disease and finally death. We developed a prototype to support nephrologists to gain an overview of their CKD patients. The prototype visualizes the patients in cohorts according to their pairwise similarity. The user can interactively modify the similarity by changing the underlying weights of the included features. The prototype was developed in response to the needs of physicians due to a context of use analysis. A qualitative user study shows the need and suitability of our new approach.

CCS Concepts
• Human-centered computing → Visualization; Interface design prototyping; Visual analytics;

© 2023 The Authors. Proceedings published by Eurographics - The European Association for Computer Graphics. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
1. Introduction

Chronic kidney disease (CKD) is a major driver of secondary diseases such as high blood pressure and cardiovascular diseases (including heart attacks and strokes) and thus places a burden on the healthcare system. CKD and end-stage renal disease (ESRD) with the need for kidney replacement procedure and renal replacement therapy are among the most cost-intensive chronic diseases worldwide [SRA∗20] and their incidence is increasing [RMVA∗19]. Furthermore, a not inconsiderable proportion is attributable to associated secondary diseases, such as the increased cardiovascular risk [SRA∗20, GCRH∗13]. This also leads to physical and cognitive limitations and drastically reduces life expectancy. It is estimated that deaths related to kidney disease will increase 2.6-fold by 2040 [FMD∗18]. CKD as progressive organ dysfunction finally results in ESRD with the need for dialysis or kidney transplantation. The average cost of dialysis therapy is up to €58,000/patient/year and a significant increase in dialysis patients is predicted by 2040 [KR07, HKS21].

New e-Health interventions like electronic health records or tele-health are already well-established in medicine [WK20]. Due to its multifactorial onset and complex course of disease, CKD exhibit numerous application possibilities not only for e-Health systems, but especially for clinical decision support systems (CDSS) to improve CKD management [FC20]. The relevant health records of CKD patients consist of static and temporal attributes that nephrologists would like to utilize for a finer-grained cohort definition and the overuse of reminders and alerts, so that the level of adherence to treatment guidelines [AKFL∗18, FKS20] nor the systems already on the market are focused on nephrological treatment or use elements of visual analytics to process treatment-relevant parameters.

Currently, there are no comparable approaches that apply visual analytics to the nephrology context. Neither existing patents [AS16, CHK∗18, FKS20] nor the systems already on the market are focused on nephrological treatment or use elements of visual analytics to process treatment-relevant parameters.

3. Visual Analytics Approach in KidCAD

In this section, we describe the data basis and the various features of our approach.

Data

The used data is assembled from various sources. The basis consists of a realistic data set validated by our medical co-authors and collaborators. This data provides all relevant demographic information of the patients as well as nephrologic data, underlying conditions and haemodialysis data. Furthermore, this data set contains a detailed list of all important measured values for blood and urine samples. With the help of our professional co-authors, we acquired the initial weights that are used for the similarity calculation. The data also provides short time series for single patients of blood and urine samples. Since the time series play an important role in the similarity calculation we generated a large synthetic patient population with the Synthea Patient Generator [syn17]. We selected five synthetic patients with typical conditions for different real patient groups and used the synthetic time series to extend our data sets with realistic patient data for the development of our visualizations and similarity definition.

3.1. Similarity Definition

A patient is described by a feature vector, consisting of static properties and several vectors with time series. The static properties are patient-related data like age, sex and underlying medical conditions as well as nephrological-related properties like CKD stage, nephrological primary diseases or [years since] start of dialysis. The time series are divided into three categories: laboratory serum (e.g., creatinine), laboratory urine (e.g., total protein) and haemodialysis.

The distance between two patients $i$ and $j$ is defined as the sum $ω_p \times δ_p(i, j) + ω_s \times δ_s(i, j)$ with the distance of the properties of patient $i$ and $j$ defined as $δ_p(i, j)$ and the distance between time series of patient $i$ and $j$ defined as $δ_s(i, j)$. $ω_p$ and $ω_s$ are weights used to set the importance of the properties and time series in the calculation.

Distance between patient properties

The distance $δ_p(i, j)$ between two patients $i$ and $j$ according to their properties is described as the weighted sum of the single attribute differences $δ_p(i, j) = \frac{1}{n} \sum_p [α_p \times |x_{i, p} - x_{j, p}|]$, where $α_p$ is the set weight for this property and $x_{i, p}$ and $x_{j, p}$ are the values for this property from patient $i$ and patient $j$, respectively. For numeric values
like age, or CDK class a measurement based on a Euclidean distance and used for sets like the nephrological primary diseases and underlying conditions a measurement based on a Jaccard distance [Jac12] is used. As shown in Figure 1 (C), the weight $a_p$ for a property can be interactively set via the user interface.

**Distance between time series** The distance $\bar{d}(i, j)$ between two patients $i$ and $j$ according to their time series is defined as a weighted sum of a similarity function $\phi(x_{(i)}(t), x_{(j)}(t))$ as follows:

$$
\bar{d}(i, j) = \frac{1}{\alpha_p} \sum \alpha_x \cdot \phi(x_{(i)}(t), x_{(j)}(t))
$$

where $\alpha_x$ is the set weight for the time series item and $x_{(i)}(t)$ and $x_{(j)}(t)$ are the time series of patient $i$ and patients $j$, respectively. The weight $\alpha_x$ for a time series item can be set by the user via the interactive user interface (see Figure 1 (C)). The similarity function $\phi$ takes several derived statistic measurements of the underlying time series $x_{(i)}(t)$ and $x_{(j)}(t)$ into consideration. The derived measurements are Average, Correlation, Interquartile range, Last [value], Max [value], Median [value], Min [value], N [quantity], Peaks, lower quartile (Q1), upper quartile (Q3), Range and Standard Deviation. The correlation is calculated as the sample correlation coefficient of a single time series to estimate whether a series is stable. A peak will occur, if the difference between adjacent items is greater than a multiple of the standard deviation. This parameter is used to identify very steep sections of the time series, i.e. large variation of the data in a short period of time.

3.2. Views

Our KidCAD dashboards consists of several views that are combined to give medical experts the relevant information on single patients and cohorts, and allow the adaptation of similarity settings.

**Main View** The Main View (Figure 2a) applies a scatter plot to visualize the cohorts. This approach uses multidimensional scaling to project the calculated similarity matrix $\bar{d}^{n \times n}$ onto a 2D plane with $n$ points. The view is designed as an easy-to-use entry point and the clustering of patients with similar disease patterns provides an overview of patients and potential groupings. The medication of a patient is displayed via a tool tip. The user is also able to select a person or group of interest with a lasso to open further views to investigate the measured values for blood and urine samples.

**Data basis panel** The data basis panel (Figure 1 (C)) lists all attributes and their weights $\alpha_q$ and $\alpha_t$ included in the calculation of the similarity. It is grouped into the four categories patient properties, laboratory serum, laboratory urine and haemodialysis. This module offers the possibility to set the weights of the attributes interactively. It also shows the correlation of a single attribute concerning the calculated similarity of all patients both as gradient and number. The attributes of a group can be sorted alphabetically or by correlation value.

**Patient Cards** The patient cards (Figure 1 (A)) give detailed information on preconditions, comorbidities and the current medical treatment of a patient. Besides the demographic data like name, sex and date of birth, the card provides nephrologie information in different categories. The classification (Figure 1 (A) P1) provides information on the nephrologic primary diseases and the current CKD stage, whether the patient has haemodialysis, and the number of transplantations if any kidney transplantations occurred. Another category lists all known physical conditions and comorbidities that affect the kidney disease and the treatment (Figure 1 (A) P2). The last category contains the information on haemodialysis (Figure 1 (A) P3). It shows the time on dialysis, the number of treatments per week, the time per treatment, and the dialysis vascular access.

**Patient View** Along with the patient cards the physician can use the Patient View (Figure 2b) to become familiar with a patient. The Patient View visualizes the time series of the measured blood and urine samples in three ways. First, the values are displayed in chronological order with a line chart augmented with the median value. During examination, user orientation is supported by a crosshair. Furthermore, the distribution of the values is visualized both as a box plot and as a histogram. Users can select various samples according to their choice to obtain an individual view of the most important time series.

**Comparison View** To match samples from multiple patients, the physician is assisted with a Comparison View (Figure 1 (B)). The user can either use the lasso selection in the Main View or select the desired patients as a subset of the Patient Cards. As in the Patient View, the samples are visualized as time series and distributions. However, instead of visualizing different samples of a patient at once, the physician can use the Comparison View to examine one sample of multiple patients to identify similarities within a cohort or differences between cohorts. In order to retain the familiar visualization of a time series with a line chart, a boxplot and a histogram, the Comparison View is based on juxtaposition with the time axis aligned. This approach is similar to the evaluated multipatient time series view by Gschwandtner et al. [GAK*11].

4. Human Centered Design and Evaluation

The design process of the dashboard was aligned with the human-centered design process for interactive systems [DIN20]. First, we conducted a context of use analysis with 30 nephrologists to identify the requirements. Then, the results from this analysis informed the design of the dashboard, which is presented in section 3. Finally, we evaluated the design solution with five nephrologists through an online survey. In the following sections, the user studies will be explained in more detail.

4.1. Context of Use and Requirement Analysis

In order to identify the requirements for a dashboard to support nephrologists, the context of use was first analyzed. Using an online survey and structured interviews, both for the use cases “Patient with Chronic Kidney Failure” and “Dialysis Patient”, it was identified who would use such a dashboard and under which circumstances, as well as which treatment steps the dashboard could support in which manner. In particular, the required information and its presentation were discussed. 15 nephrologists participated in both the online survey and the interviews. Of the 30 nephrologists aged between 29 and 82 years ($Md = 52.4, SD = 12.6$), 16 identified themselves as male and 14 as female. 12 of the respondents worked in a clinic, two in a private medical practice, five
in an outpatient dialysis centre and eleven in both a private medical practice and an outpatient dialysis centre.

The results of the surveys showed for both use cases in particular, that nephrologists can be supported by the provision and visualization of vital and laboratory parameters, as well as the aggregated presentation of information on diagnosis, CKD stage, transplants, medication, comorbidities and information on the patient’s dialysis. A clear presentation of the time course of vital and laboratory parameters is particularly important. It has also been found that categorizing patients according to different parameters can be helpful in their daily work when making diagnoses and identifying interdependencies. This is especially relevant for clinical use, as one nephrologist stated.

4.2. Evaluation

The usability evaluation of the dashboard with domain experts was conducted with another online survey that consisted of two scenarios. Both scenarios were presented by a short video showing and explaining the features of the dashboard. In scenario 1 “Main and Comparison View” (Figure 2a, Figure 1 (B)) the features for comparing patients and similarity-based grouping were explained while in scenario 2 “Patient View” (Figure 2b) the features for individual patients were presented. The videos were followed by open-ended questions about the first impressions of the dashboard, its support in daily work, possible future features, and (in addition for “Main and Comparison View”) an open-ended question about the comparison of the patients. Afterwards, different scales were used to assess usability, perceived usefulness, perceived ease of use, and compliance with user expectations. Therefore, the German version of the System Usability Scale [B'96], of the Perceived Usefulness and Perceived Ease of Use Scale [Dav89] and of the subscale ‘conformity with user expectations’ of the IsoMetrics Scale [WHG97] were used. Five nephrologists who identified themselves as male were recruited exclusively to participate in this online survey. The average length of practice as a nephrologist was six years (min = two years, max = 17 years).

In general, the features presented in “Main and Comparison View” were rated as helpful and the participants requested additional features as described below. Participants also wished for more features in “Patient View” and rated the added value of the dashboard to be diverse for “Patient View”. The results of the evaluation are presented in detail below.

First impression Regarding the first impressions of the features presented in “Main and Comparison View”, participants described the grouping as meaningful and stated that the dashboard facilitates medical decisions. One participant said, ‘Grouping patients with similar disease patterns certainly makes sense.’ First impressions of the features presented in “Patient View” were rather mixed. The dashboard was described as ‘helpful’, ‘intuitive’, and ‘clear’, but also ‘unclear’ and with little added value to existing software.

Support in daily work The features shown in “Main and Comparison View” would support the nephrologists in their daily work if they sent grouped orders to the laboratory and nursing staff and if the dashboard visualized deviations. In “Patient View”, one participant saw little advantage over existing software and another participant stated that the dashboard would support daily work if guideline-based therapy failed. One other participant said that ‘The progressions can be recorded and traced well. This is also helpful in patient consultations to explain the ... medical progression.’ The visualization of laboratory parameters was described as ‘helpful’.

Future features For “Main and Comparison View”, additional features mentioned by the participants included: ‘grouping by presumed diagnosis’, ‘setting a pivot point’, ‘visualization of the medication plan’, and other visualization options. In “Patient View”, the participants wanted features such as ‘standard medication for the medication plan’, ‘monitoring of therapy success’, ‘listing for transplantation’ or ‘comparison with previous values’.

Comparison View Regarding the visualization of the comparison of patients in “Main and Comparison View”, one participant stated that there was more added value for scientific use than for daily work in the clinic. One participant said, ‘This would make the se-
The participants were shown the ‘Comparison View’ and the ‘Patient View’, and the analysis was carried out. Finally, a qualitative user study was conducted. Furthermore, the similarity visualization updates immediately to gain a data basis with all relevant features. The user can vary features as well as one time series of multiple patients for comparison reasons. The main view is used to show the left side, the user can gather information on demographics as well as disease-related data like the ‘Comparison View’ and the ‘Patient View’. To assess subjective perceptions, the participants were asked to complete a System Usability Scale [B*96*] and one questionnaire each to measure perceived usefulness and perceived ease of use. The SUS score shows, that the system can be classified as ‘good’, whereas the perceived usability and ease of use for their daily work was rated as ‘rather likely’ and ‘quite likely’, respectively.

6. Future Work

One of the next steps will be the enlargement of the data set. Besides the data that we already used to enrich the time series of the patients, the Synthea [syn17] data set also contains data of patients with different CKD stages as well as patients with other diagnoses that we can use as control groups. Along with a larger data set, the main visualization showing the cohorts must be renewed. The key change is to find a visualization that allows for an interactive clustering of patients, so that a specific therapy out of a variety of possible treatment options can easily be identified for a new patient at first consultation. With significantly more data, it is not possible to give each patient their own color. Thus, we are considering to use color for distinct cohorts instead of individual patients. Furthermore, the list of patients must be modified so that the patient card will be an on-demand feature for further information. We also plan to integrate usability enhancements such as filters, or an estimation of the configuration quality. Furthermore, aligning by sentinel events and comparing the time series before or after these could be an interesting new feature. In addition to feedback from the domain experts mentioned during the evaluation, qualitative studies of usefulness with clinicians and real data of their patient populations will be important impulses for future developments.

References


Nierenversagen in Deutschland: Vergleich der Dialyseversorgung inner-
erised decisions support system for renal risk drugs targeting primary
life lost, and all-cause and cause-specific mortality for 250 causes of
support systems change.
International journal
centred design for interactive systems (ISO 9241-210:2019); German
user acceptance of information technology.
sion support systems: improving patient safety in nephrology.

BMJ open 5
IEEE, pp. 43–50. 3

The Lancet 392

International journal of health care finance and economics
Germany. 
International journal of health care finance and economics (2007), 185–200. 2

tilayer perception method for supporting the medical decision making. 
Pattern Recognition Letters 151 (2020), 15–22. 2

and nephrology 45 (2013), 1033–1038. 2

[MPU2014*14] Martínez-Pérez B., de la Torre-Díez I., López-Coronado M., Sainz-de Abaro B., Robles M., García-
Gómez J. M.: Mobile clinical decision support systems and applica-
tions: a literature and commercial review. Journal of medical systems 38
(2014), 1–10. 2

[PMR*96] Plaisant C., Milash B., Rose A., Wijffels S., Shenber-
pp. 221–227. 2

[PSS*19] Pape L., Schneider N., Schleef T., Junius-Walker U., Haller H., Brunkhorst H., Hellrung N., Prokosch H., 

[RVM*19] Roberts N. L., Mountjoy-Vennin W. C., An-
jomshoa M., Banoub J. A. M., Yasin Y. J.: GBD 2017 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, 
and national incidence, prevalence, and years lived with disability for 354 
diseases and injuries for 195 countries and territories, 1990-2017: a 
systematic analysis for the Global Burden of Disease Study (vol 392, pg 
1789, 2018). Lancet 393, 10190 (2019), E44–E44. 2

[SBRE15] Shemekka T., Bastholm-Rahmner P., Elinder C.-
G., Veg A., Tornvqvist E., Cornelius B., Korkmaz S.: A health record integrated clinical decision support system to support prescrip-
tions of pharmaceutical drugs in patients with reduced renal function: 
design, development and proof of concept. International journal of med-
ical informatics 84, 6 (2015), 387–395. 2

P. P.: Dimensionality reduction using genetic algorithm for improving 
accuracy in medical diagnosis. International Journal of Intelligent Sys-
tems and Applications 8, 1 (2016), 67. 2

[SRA*20] Saran R., Robinson B., Abbott K. C., Bragg-
Gresham J., Chen X., Gipson D., Gu H., Hirth R. A., Hutton 
D., Jin Y., et al.: US renal data system 2019 annual data report: epi-
demiology of kidney disease in the United States. American Journal of 
Kidney Diseases 75, 1 (2020), A6-A7. 2


Fragebogen zur Evaluation von graphischen Benutzungsschnittstellen. 
4

Nephrology 16, 7 (2020), 368–370. 2

© 2023 The Authors.
Proceedings published by Eurographics - The European Association for Computer Graphics.