Network of vascular diseases, death and biochemical characteristics in a set of 4,197 patients with type 1 diabetes (The FinnDiane Study)

Ville-Petteri Mäkinen*1,2,3, Carol Forsblom1,2, Lena M Thorn1,2, Johan Wadén1,2, Kimmo Kaski3, Mika Ala-Korpela4,5 and Per-Henrik Groop*1,2,6

Address: 1Folkhälsoan Institute of Genetics, Folkhälsoan Research Center, Biomedicum Helsinki, Finland, 2Division of Nephrology, Department of Medicine, Helsinki University Central Hospital, Finland, 3Department of Biomedical Engineering and Computational Science, Helsinki University of Technology, Finland, 4Computational Medicine Research Group, Institute of Clinical Medicine, Faculty of Medicine, University of Oulu and Biocenter Oulu, Finland, 5Department of Internal Medicine and Biocenter Oulu, Clinical Research Center, University of Oulu, Finland and 6The Baker IDI Heart and Diabetes Institute, Melbourne, Victoria, Australia

Email: Ville-Petteri Mäkinen* - ville-petteri.makinen@finndiane.fi; Carol Forsblom - carol.forsblom@hus.fi; Lena M Thorn - lena.thorn@helsinki.fi; Johan Wadén - johan.waden@helsinki.fi; Kimmo Kaski - kimmo.kaski@hut.fi; Mika Ala-Korpela - mika.alakorpela@computationalmedicine.fi; Per-Henrik Groop* - per-henrik.groop@finndiane.fi

* Corresponding authors

Background: Cardiovascular disease is the main cause of premature death in patients with type 1 diabetes. Patients with diabetic kidney disease have an increased risk of heart attack or stroke. Accurate knowledge of the complex interdependencies between the risk factors is critical for pinpointing the best targets for research and treatment. Therefore, the aim of this study was to describe the association patterns between clinical and biochemical features of diabetic complications.

Methods: Medical records and serum and urine samples of 4,197 patients with type 1 diabetes were collected from health care centers in Finland. At baseline, the mean diabetes duration was 22 years, 52% were male, 23% had kidney disease (urine albumin excretion over 300 mg/24 h or end-stage renal disease) and 8% had a history of macrovascular events. All-cause mortality was evaluated after an average of 6.5 years of follow-up (25,714 patient years). The dataset comprised 28 clinical and 25 biochemical variables that were regarded as the nodes of a network to assess their mutual relationships.

Results: The networks contained cliques that were densely inter-connected ($r > 0.6$), including cliques for high-density lipoprotein (HDL) markers, for triglycerides and cholesterol, for urinary excretion and for indices of body mass. The links between the cliques showed biologically relevant interactions: an inverse relationship between HDL cholesterol and the triglyceride clique ($r < -0.3$, $P < 10^{-16}$), a connection between triglycerides and body mass via C-reactive protein ($r > 0.3$, $P < 10^{-16}$) and intermediate-density cholesterol as the connector between lipoprotein metabolism and albuminuria ($r > 0.3$, $P < 10^{-16}$). Aging and macrovascular disease were linked to death via working ability and retinopathy. Diabetic kidney disease, serum creatinine and potassium, retinopathy and blood pressure were inter-connected. Blood pressure correlations indicated accelerated vascular aging in individuals with kidney disease ($P < 0.001$).

Conclusion: The complex pattern of links between diverse characteristics and the lack of a single dominant factor suggests a need for multifactorial and multidisciplinary paradigms for the research, treatment and prevention of diabetic complications.
Background
A significant number of patients with type 1 diabetes suffer from severe microvascular complications such as diabetic kidney disease and proliferative retinopathy [1,2]. The pathogenetic mechanisms responsible for the degradation of the vascular system are not yet fully known, but a complex pattern of interactions between susceptibility genes and environmental factors is the likely cause. Kidney failure is not the primary cause of death, but these patients die mostly from cardiovascular complications at the later stages of the disease [1,3,4]. The risk factors have been extensively investigated [5-7]; however, only a handful of studies have focused on the statistical associations between biochemical and clinical variables from a multivariate perspective [8-10]. The biological variation at the individual level is substantial, which means that the phenotype cannot be compressed into a single variable. Albuminuria, for instance, is the most important clinical risk factor, but it alone provides only limited information on the systemic changes in the body.

Complex network analysis has gained popularity as new datasets and techniques have become available [11,12]. Recent examples include communication patterns in social networks [13,14], molecular interactions in proteomics and metabolomics [15,16] and the epidemiology of contagious diseases [17,18]. Visualization of the network structures helps to understand the complex phenomena and computerized applications are commonplace in network research [19,20].

This work illustrates the main modules of clinical and biochemical associations in type 1 diabetes. Our aim is to present the characteristics of diabetic complications as an inter-connected system, instead of focusing on any single variable at a time. We also discuss the biological processes that can be attributed to the observed network structures, and demonstrate the links between multiple chronic conditions, lifestyle, aging, and metabolic traits in their full context.

Methods
Type 1 diabetic patients were recruited by the Finnish Diabetic Nephropathy Study Group (N = 4,197). The design was cross-sectional (serum and urine samples), but with longitudinal records of albuminuria and clinical events before baseline and with all-cause mortality data available after an average of 6.5 years of follow-up from baseline (25,714 patient-years). Type 1 diabetes mellitus was defined as an age of onset below 35 years and transition to insulin treatment within a year of onset. Macrovascular disease (337 cases) was obtained from medical records and defined as a pooled end-point of coronary heart disease (224 cases), myocardial infarction (124 cases), stroke (100 cases), and peripheral vascular disease (91 cases). The classification of renal status was made centrally according to urinary albumin excretion rate (AER) in at least two out of three consecutive overnight or 24 h urine samples. Absence of diabetic kidney disease (DKD) was defined as AER within the normal range (AER <20 μg/min or <30 mg/24 h) and at least 15 years of type 1 diabetes. This kidney disease negative subset is denoted by KDNEG. Macroalbuminuria or overt kidney disease was defined as AER ≥ 200 μg/min or ≥ 300 mg/24 h. The intermediate range was defined as microalbuminuria (20 ≤ AER <200 μg/min or 30 ≥ AER <300 mg/24 h). Patients on renal replacement therapy (dialysis or transplantation) were classified as having end-stage renal disease (ESRD). An additional subset, denoted by DMDur<15', was formed from patients with less than 15 years of diabetes duration, and normal (1,004 individuals) or unknown AER (135 individuals). A total of 296 patients could not be classified.

The AER values that were used for the DKD diagnosis were measured in the local health care centers, but not used for statistical analyses. Instead, the continuous 24 h albumin excretion rate was estimated from a single 24 h-urine collection (available for 80% of patients) from which albumin was measured by a central laboratory.

Education level, smoking and alcohol dose, working status, asthma, rheumatoid arthritis and thyroid disease were determined by patient questionnaires. Education level was defined as the expected number of years in the educational system based on the current occupation, smoking exposure was calculated as the product of daily cigarettes and years of smoking, the daily dose of alcohol was estimated from the type and quantity of drinks consumed. Working status was compressed into a binary trait (disabled vs. employed or unemployed). Serum concentration of the soluble receptor for advanced glycation end-products (SRAGE) was measured by solid phase ELISA (Thomas et al. submitted). VLDL triglycerides and IDL and LDL cholesterol were estimated by neural network modeling [21]. Other details on the data sources, clinical definitions and patient characteristics have been published previously [22]. More information on the kidney disease subsets is available in [Additional file 1].

Statistical analysis
Many of the continuous variables had skewed distributions and it is typical for a large clinical study to have a small percentage (<5%) of outliers. Therefore, the continuous variables were sorted and converted to scaled ranks between -1 and 1 to prevent statistical artifacts. Two versions of the dataset were created: one with men and women pooled, the other with separate rank transforms for the sexes.
The network of continuous variables was based on pairwise Spearman’s correlation coefficients. Specifically, each variable is considered a node and the nodes are connected by links, the weights of which are quantified by the correlation coefficient. The full networks are too dense and have to be pruned in order to highlight the relevant patterns. There are numerous ways to reduce the network dimensionality [23], here we chose the spanning trees since they are computationally efficient and ensure the connectivity of the pruned network [24].

Direct graph-theoretic investigation of the networks provides little useful information since many of the variables are derived or otherwise non-biologically linked with each other. Therefore, structural considerations were made via comparisons between the kidney disease subsets to reduce the distraction from irrelevant connections. Statistical significance was estimated by random permutations of the subset labels [Additional file 2]. Although the individual links are not independent, they may be subject to multiple testing effects (less than 741 tests). P-values between 0.01 and 0.0001 are therefore considered suggestive. Topologically relevant links were chosen as follows: i) the link must belong to at least one of the spanning trees from difference networks between KDNEG and the other subsets and ii) the link must be one of the top 10 most significant (and \( P < 0.01 \)) in its spanning tree. This procedure was chosen to avoid selecting too many links for closer inspection, and yet ensuring that as many nodes as possible would be represented.

Correlation coefficient is not well suited for comparing binary and continuous variables. For this reason, a computationally intensive regression-correlation measure was applied to the full dataset to create the visualization [Additional file 2]. All statistical analyses were performed with in-house scripts in the Octave programming environment [25].

**Results**

Figure 1 depicts the correlation structure of the gender-adjusted dataset. The network is characterized by strong links between methodologically and biochemically dependent variables: markers of body mass (weight, BMI,
WHR, etc.), 24 h urine excretion (potassium, urea, sodium and creatinine), HDL-related biochemistry (HDL cholesterol, apolipoprotein A-I and A-II) and other lipoprotein quantities (triglycerides, total cholesterol, apolipoprotein B-100) form positively correlated cliques.

There are strong inverse associations between the HDL-clique and IDL cholesterol. Other connections include the links between triglycerides and body mass via C-reactive protein, and the central role of 24 h urine albumin as the connector of triglycerides, IDL cholesterol, hemoglobin A1c, blood pressure and serum creatinine. Adiponectin links HDL metabolism with 24 h urine metabolites and SRAGE is located between serum creatinine and body mass. Smoking and alcohol intake are correlated; alcohol consumption is also reflected in apolipoprotein A-I and A-II concentrations, and smoking dose is linked with lower education and - by definition - to higher age. Weight-adjusted insulin dose is inversely associated with aging in this dataset.

Network topology and diabetic kidney disease
The dataset was divided according to AER and diabetes duration (see Methods) and the subset networks were compared to detect relevant topological features. Table 1 shows the statistical significance of difference networks (based on the sets of pair-wise correlation coefficients) between the patient groups. The correlation structure for the KDNEG subset with 15 years or more duration is significantly different from the macroalbuminuria subset (P = 2.2 × 10^-16) and from the patients with short duration (P = 4.0 × 10^-32). On the other hand, the DMDur<15 subset is different from the macroalbuminuria subset (P = 1.2 × 10^-29).

Table 2 lists the (selected) significant changes in link weights with respect to the KDNEG subset. Age and blood pressure show a mixed trend: diastolic blood pressure has a negligible age-dependence in the KDNEG subset, but an inverse correlation in the macroalbuminuria subset (r = 0.02 vs. -0.20, P = 2.5 × 10^-5), whereas systolic blood pressure shows stronger dependence in the KDNEG subset (r = 0.43 vs. 0.28, P = 4.0 × 10^-4). Adiponectin is also age-dependent in the KDNEG subset, but uncorrelated in the macroalbuminuria subset (r = 0.32 vs. 0.08, P = 8.5 × 10^-7).

Serum creatinine is connected to adiponectin (r = 0.05 vs. 0.29, P = 6.3 × 10^-8) and SRAGE (r = 0.03 vs. 0.33, P = 1.5 × 10^-10) in the macroalbuminuria, but not in the KDNEG subset. The associations between albumin excretion and other variables are also negligible in the KDNEG subset. On the other hand, 24 h urine albumin is significantly correlated with total cholesterol (r = 0.02 vs. 0.23, P = 3.0 × 10^-6), IDL cholesterol (r = 0.06 vs. 0.30, P = 3.0 × 10^-8) and triglycerides (r = 0.08 vs. 0.23, P = 3.0 × 10^-4) in the macroalbuminuria subset.

Node strength measures the overall connectivity of a node: it is the sum of the correlation magnitudes that link the node to the rest of the network. A high strength indicates a structurally significant variable, although the value itself is less important and therefore not reported here. The strengths were not different between the microalbuminuria and KDNEG subsets (data not shown). Within the macroalbuminuria subset, the connections surrounding serum creatinine (P = 5.1 × 10^-7), 24 h urine albumin excretion (P = 0.0010), SRAGE (P = 0.0024) and apolipoprotein B-100 (P = 0.0046) were significantly changed. The ESRD group showed a similar structure: serum creatinine, 24 h urine albumin and SRAGE were significantly different (P < 0.00050). There were statistically significant differences also in the connectivity of adiponectin (P = 8.1 × 10^-6), age (P = 8.3 × 10^-5), 24 h urine urea (P = 0.00032), BMI (P = 0.0021), insulin dose (P = 0.0034) and apolipoprotein A-I (P = 0.0099) when compared with the KDNEG subset.

Regression-correlation network
Figure 2 depicts the network based on regression modelling of both the continuous and binary variables. The data were not adjusted for gender effects, since gender was

Table 1: Comparison of diabetic kidney disease networks

<table>
<thead>
<tr>
<th></th>
<th>Microalbuminuria (n = 586)</th>
<th>Macroalbuminuria (n = 558)</th>
<th>ESRD (n = 289)</th>
<th>DMDUR&lt;15 (n = 1,139)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KDNEG (n = 1,379)</td>
<td>0.0056</td>
<td>2.2 × 10^-16</td>
<td>4.6 × 10^-27</td>
<td>4.0 × 10^-32</td>
</tr>
<tr>
<td>Microalbuminuria</td>
<td>7.1 × 10^-4</td>
<td>7.3 × 10^-18</td>
<td>7.8 × 10^-20</td>
<td></td>
</tr>
<tr>
<td>Macroalbuminuria</td>
<td>6.2 × 10^-9</td>
<td>1.2 × 10^-29</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ESRD</td>
<td></td>
<td></td>
<td>8.3 × 10^-11</td>
<td></td>
</tr>
</tbody>
</table>

Statistical significance estimates (P-values) from permutation analysis of difference networks. The networks were formed from pair-wise Spearman correlation coefficients of 39 continuous clinical and biochemical variables.
Table 2: Correlations within diabetic kidney disease groups

<table>
<thead>
<tr>
<th></th>
<th>KDNEG r</th>
<th>Microalbuminuria r</th>
<th>Macroalbuminuria r</th>
<th>ESRD r</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age -- Diastolic blood pressure</td>
<td>0.02</td>
<td>-0.15*</td>
<td>-0.20**</td>
<td>-0.29**</td>
</tr>
<tr>
<td>Age -- Systolic blood pressure</td>
<td>0.43</td>
<td>0.35</td>
<td>0.28*</td>
<td>0.06**</td>
</tr>
<tr>
<td>Adiponectin -- Age</td>
<td>0.32</td>
<td>0.30</td>
<td>0.08**</td>
<td>-0.01**</td>
</tr>
<tr>
<td>Adiponectin -- HDL cholesterol</td>
<td>0.45</td>
<td>0.36</td>
<td>0.21***</td>
<td>0.10**</td>
</tr>
<tr>
<td>ApoA-II -- HDL2 cholesterol</td>
<td>0.13</td>
<td>0.17</td>
<td>0.34***</td>
<td>0.33*</td>
</tr>
<tr>
<td>ApoA-II -- Waist</td>
<td>0.16</td>
<td>-0.02*</td>
<td>0.02**</td>
<td>0.05</td>
</tr>
<tr>
<td>Total cholesterol -- Education</td>
<td>-0.03</td>
<td>-0.18*</td>
<td>-0.08</td>
<td>-0.01</td>
</tr>
<tr>
<td>Serum creatinine -- Adiponectin</td>
<td>0.05</td>
<td>0.03</td>
<td>0.29***</td>
<td>0.18</td>
</tr>
<tr>
<td>Serum creatinine -- Diabetes duration</td>
<td>0.07</td>
<td>0.22*</td>
<td>0.17</td>
<td>-0.06</td>
</tr>
<tr>
<td>Serum creatinine -- Insulin dose</td>
<td>-0.01</td>
<td>-0.17*</td>
<td>-0.15*</td>
<td>-0.13</td>
</tr>
<tr>
<td>Serum creatinine -- SRAGE</td>
<td>0.03</td>
<td>0.05</td>
<td>0.33***</td>
<td>0.40**</td>
</tr>
<tr>
<td>Serum creatinine -- 24 h-uAlb</td>
<td>0.06</td>
<td>0.07</td>
<td>0.15</td>
<td>0.44†</td>
</tr>
<tr>
<td>CRP -- Age</td>
<td>-0.10</td>
<td>0.09*</td>
<td>0.05*</td>
<td>-0.01</td>
</tr>
<tr>
<td>CRP -- Serum potassium</td>
<td>-0.05</td>
<td>0.12*</td>
<td>-0.01</td>
<td>-0.02</td>
</tr>
<tr>
<td>CRP -- Waist-hip ratio</td>
<td>0.18</td>
<td>0.34*</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>IDL cholesterol -- LDL cholesterol</td>
<td>0.72</td>
<td>0.63*</td>
<td>0.53***</td>
<td>0.53**</td>
</tr>
<tr>
<td>LDL cholesterol -- Education</td>
<td>-0.01</td>
<td>-0.17*</td>
<td>-0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>MBL -- 24 h-urine urea</td>
<td>0.08</td>
<td>-0.10*</td>
<td>-0.05</td>
<td>-0.02†</td>
</tr>
<tr>
<td>Serum potassium -- Diabetes duration</td>
<td>0.27</td>
<td>0.26</td>
<td>-0.02**</td>
<td>-0.02**</td>
</tr>
<tr>
<td>VLDL triglycerides -- 24 h-uAlb</td>
<td>0.07</td>
<td>0.12</td>
<td>0.22*</td>
<td>0.51†</td>
</tr>
<tr>
<td>24 h-uAlb -- ApoB</td>
<td>0.07</td>
<td>0.19</td>
<td>0.27***</td>
<td>0.31†</td>
</tr>
<tr>
<td>24 h-uAlb -- Total cholesterol</td>
<td>0.02</td>
<td>0.17*</td>
<td>0.23***</td>
<td>0.16†</td>
</tr>
<tr>
<td>24 h-uAlb -- HDL cholesterol</td>
<td>-0.06</td>
<td>-0.04</td>
<td>-0.12</td>
<td>-0.46†</td>
</tr>
<tr>
<td>24 h-uAlb -- IDL cholesterol</td>
<td>0.06</td>
<td>0.16</td>
<td>0.30***</td>
<td>0.50†</td>
</tr>
</tbody>
</table>
Table 2: Correlations within diabetic kidney disease groups (Continued)

<table>
<thead>
<tr>
<th>Comparison</th>
<th>24 h-uAlb -- Triglycerides</th>
<th>0.08</th>
<th>0.13</th>
<th>0.23*</th>
<th>0.50†</th>
</tr>
</thead>
<tbody>
<tr>
<td>24 h-uAlb -- 24 h-urine creatinine</td>
<td>0.11</td>
<td>-0.05*</td>
<td>0.02</td>
<td>-0.36†</td>
<td></td>
</tr>
<tr>
<td>24 h-uAlb -- 24 h-urine urea</td>
<td>0.04</td>
<td>-0.06</td>
<td>-0.06</td>
<td>-0.41†</td>
<td></td>
</tr>
</tbody>
</table>

Comparison of the KDNEG subset network against the micro-, macroalbuminuria and ESRD networks. The links were chosen by an automatic network topology algorithm (see Methods for details). The Spearman correlation coefficient (denoted by \( r \)) of continuous clinical and biochemical variables was used as the measure of association between the variables. Urine samples were not available from most patients with ESRD (72% missing); the \( r \) values presented were obtained from the imputed dataset. The links are sorted alphabetically. *\( P < 0.01 \), **\( P < 0.0001 \), comparison with KDNEG; † imputed.

Figure 2
Regression-correlation network of continuous and binary data. A pruned visualization of the correlation network from regression modeling. Unlike in Figure 1, the data were not adjusted for gender prior to the analysis. Each variable was converted to a surrogate linear predictor before computations. The symbols in the figure correspond to the source of information: directly observed variables are filled, whereas derived variables are denoted by open symbols. A circle is used for continuous quantities, and a diamond for binary traits. The width and color of the links indicate the association magnitude and type, as shown in the legend. The \( r \) denotes the correlation of the linear predictors and is not comparable with Figure 1. Abbreviations: history of macrovascular disease (MVD), systolic (SBP) and diastolic (DBP) blood pressure, anti-hypertensive treatment (AHT) and soluble receptor for advanced glycation end-products (SRAGE). Visualized with the Himmeli software [47].
Mortality and kidney disease were not connected directly, although part of the same node group. Instead, laser-treated retinopathy and anti-hypertensive treatment had a direct link with death, and there was also a strong inverse connection to working ability (Figure 2). Macrolambinuria is a powerful risk marker, but these results may indicate the later stages when the patient's health deteriorates to the point were normal life is severely interrupted (loss of sight and working ability) and death ensues. There were twice as many patients without ESRD in the kidney disease group (586 vs. 289). Furthermore, patients with ESRD suffer from secondary effects of kidney failure that disturb the metabolic patterns. This means that the DKD node in the regression-correlation network may be a more accurate estimate for persistent albuminuria (before kidney failure) than for ESRD and death.

The DCCT Study has established the beneficial effects of tight glycemic control on diabetic complications [26,27]. Here, hemoglobin A1c was not among the most structurally significant nodes, although it was suggestively positioned between insulin dose and triglycerides in Figure 1. This does not mean that better insulin treatment is useless; it most likely reflects the biological variability of the A1c measure in our observational data [28]. Advanced glycation end-products in general have been implicated in diabetic tissue damage [29]. The soluble receptor (SRAGE) was connected to complications (kidney function) in this study, although the result can also be explained by reduced clearance.

The patient material was extensive with detailed clinical characteristics and biochemical measurements from serum and urine. On the other hand, the dataset was not complete and special procedures had to be taken to impute the missing values. Many of the correlation coefficients were small (r < 0.3) and cannot be considered clinically significant. This is most likely due to the robust but less sensitive rank transform, the need to avoid linear artifacts in data imputation, and the observational nature of the study. In particular, 24 h urinary albumin excretion was weakly correlated with the other variables, which may be the result of the large biological variation in cross-sectional urine collections. There were only a few cases of asthma, rheumatoid arthritis or thyroid disease, which further reduced the power of the regression-correlation approach. Despite the problems, the negative associations with male gender were consistent with previous results [30].

Visual inspection of the networks was validated by additional analysis with alternate preprocessing and statistical comparisons of patient subsets. Nevertheless, the figures produced by the automatic graph drawing software are always simplifications of the true situation and should
not be used as a basis of inference without their original context. The selection and availability of variables is the critical determinant of the observed network structures, and should be taken into consideration when interpreting the results. Also, random fluctuations can change individual links, but the overall structure of a correlation network is usually resilient against sampling noise. Statistical significance estimates were not available for the regression-correlation network due to its method of construction. Nevertheless, the observations from Figure 2 were consistent with the correlation network in Figure 1, which suggests that the illustrations are reliable.

Skewed or highly variable biomarkers such as 24-h urine albumin and serum creatinine produce a correlation bias due to non-uniform signal-to-noise ratio in the subset comparisons. Low values have proportionally higher measurement errors than higher values, so the results in Table 2 also reflect changes in absolute concentrations. This is not necessarily an undesired effect; the network still reflects biologically relevant phenomena, albeit not pure associations.

The subset analyses were not matched for age. However, the association between age and diabetes duration was comparable between the groups (except within DMDur<15) and the age variances were similar [Additional file 1], which suggests that the differences in correlation coefficients were not produced by limited age ranges. The macroalbuminuria subset had longer duration than the KDNEG (29 vs. 26 years, \( P = 7.7 \times 10^{-11} \)) but, although statistically significant, the modest time gap did not interfere with the descriptive nature of the network approach.

Lipid abnormalities have been previously linked with diabetic kidney disease [31-33] and low HDL and high IDL have been implicated in cardiovascular risk [34-36]. In this study, the estimated IDL cholesterol had the strongest link with albuminuria, which is concordant with a mortality analysis of the same dataset [21]. The result may be partly explained by the reduction in the relative measurement noise after combining the three basic lipids (triglycerides, total and HDL cholesterol), but IDL was nevertheless the most important among the derived lipoprotein variables.

Low-grade chronic inflammation has emerged as a possible link between obesity and insulin resistance. For instance, when adipose tissue expands to accommodate excess lipids, macrophages therein are exposed to non-esterified fatty acids and respond by increased release of inflammatory cytokines [37,38]. The cytokines, in turn, disrupt the normal insulin signaling and fatty acid metabolism in the skeletal muscle. In the gender-adjusted network, C-reactive protein (a marker of inflammation) was positioned between the body-mass and the triglyceride cliques (Figure 1), thus reflecting the underlying biological mechanisms. In the regression-correlation network, C-reactive protein was connected with triglycerides only, probably due to the stronger link between gender and body mass (Figure 2).

Serum adiponectin is another signaling molecule that can be traced to adipose tissue - it is decreased in obesity and insulin resistance [39]. On the other hand, clinical research indicates that in kidney disease the concentration is increased [40,41] despite the simultaneous reduction in insulin sensitivity [22,42]. In this study, adiponectin was positioned as the connector between death and microvascular complications, and 24 h-urine metabolites (Figure 2), and a similar role in the middle of aging, serum creatinine and urine excretion remained in the gender-adjusted network (Figure 1). The results suggest that kidney function is a stronger determinant of adiponectin concentrations in these patients than the inverse correlation with obesity.

Aging and diastolic blood pressure are first positively correlated but then become negatively correlated at higher age. The process is accelerated in type 1 diabetes, most likely due to arterial stiffening [43,44]. The same phenomenon was also detected here from another perspective: significantly stronger negative correlations were observed within the macroalbuminuria and ESRD groups, which can indicate that the vascular aging is ahead of the KDNEG subset, even beyond the small chronological age difference. On the other hand, systolic blood pressure is less age-dependent in macroalbuminuria and ESRD groups, which is probably caused by interference from medication and the decline in the capacity of the heart to compensate for arterial stiffening [45,46].

Conclusion
The various clinical and biochemical risk factors that predispose to cardiovascular disease and diabetic complications share mutual connections that have overlapping origins in methodology, physiology and pathology. It may not be possible to fully isolate the effects of the various components in the traditional reductionist framework. Therefore, we think that the complex pattern of links between diverse characteristics such as working ability, life style, aging and biofluid chemistry is explicit evidence to develop multifactorial and multidisciplinary paradigms for the research, treatment and prevention of diabetic complications.

Competing interests
The authors declare that they have no competing interests.
Authors’ contributions
VPM designed and conceived the study, wrote the statistical software, analyzed the data and wrote the manuscript. CF collected the clinical data and reviewed the manuscript. LT collected the clinical data and reviewed the manuscript. MAK participated in conceiving the study and in writing the manuscript. PHG participated in conceiving the study, in collecting the clinical data, and in writing the manuscript.

Additional material

Additional file 1
Comparison of kidney disease groups. Median values (50% quantile) and 68% intervals (equivalent to ± SD for normally distributed variables) of the five patient subsets. The Kolmogorov-Smirnov test was used for continuous data. The P-values were obtained by comparing a given subset against the KDNEG group.
Click here for file

Additional file 2
Network methodology. A description of the statistical and visualization methods that were used in the study.
Click here for file

Additional file 3
The Finnish Diabetic Nephropathy Study Group. A listing of the hospitals and health care centers that have participated in the recruitment of patients.
Click here for file

Acknowledgements
Blood and urine samples and medical records were collected by the Finnish Diabetic Nephropathy Study Group [Additional file 3]. The study was supported by grants from the Folkhalsan Research Foundation, the Finnish Cultural Foundation, the Wilhelm and Else Stockmann Foundation, the Liv och Halsa Foundation and the European Commission (QLG2-CT-2001-01669, LSHB-CT-2003-503364, LSHB-CT-2006-037681). The work was also supported by the Center of Excellence (KK, VPM) and the SALVE (MAK) programs of the Academy of Finland.

References


37. Steinberg GR: Inflammation in obesity is the common link between defects in fatty acid metabolism and insulin resistance. Cell Cycle 2007, 6:888-894.


