# A Univariate Analysis of Risk Factors for Diabetic Nephropathy 

Anthony Shannon* ${ }^{1}$, Yee Hung Choy ${ }^{2}$<br>${ }^{1}$ Warrane College, University of New South Wales<br>PO Box 123, Kensington, NSW 1465, Australia<br>Email: tony@warrane.unsw.edu.au<br>${ }^{2}$ Hong Kong Polytechnic University<br>Hung Hom, Kowloon, Hong Kong<br>Email: mayhchoy@inet.polyu.edu.hk<br>* Corresponding author

Received: August 22, 2006
Accepted: November 12, 2006
Published: December 8, 2006


#### Abstract

This paper uses actual data from 267 patients with non-insulin-dependent (Type 2) diabetes mellitus in order to see how the various risk factors can affect the progression of diabetic nephropathy. Examination of each independent variable individually can only provide a preliminary idea of how important each variable is by itself. The relative importance of all the variables has to be examined simultaneously by multivariate methods. The approach succeeds in identifying preliminary risk factors such as smoking for males, although the females had higher fasting blood glucose at diagnosis. Not surprisingly, hypertension is common among patients of both sexes and it has an association with proteinuria in female patients in the sample.


Keywords: Multivariate model, Univariate analysis, Glomerular filtration rate, Microalbuminuria, Proteinuria.

## Introduction

The aim of this paper is to identify those factors which are related closely to the development of nephropathy in patients with diabetes mellitus. In the present case the presence of proteinuria in the patients is considered to be the hallmark for the progression of diabetic nephropathy.

Furthermore, a large number of patient characteristics is under consideration. These include demographic variables, such as age; behavioural, such as smoking; or clinical variables, such as body mass index, renal functions and blood pressure.

The approach to the present data analysis is through univariate analysis, in which different contingency tables will be set up among possible risk factors. The general procedure of the univariate analysis is to relate various patient characteristics (or independent variables) to the occurrence of the event (dependent variable) on the basis of data collected from those individuals for whom the event (that is, nephropathy) occurred (group 1) and those for whom the event did not occur (group 2). For example, in order to relate variables such as age, sex, and body mass index to the development of diabetic nephropathy, one needs to analyse the information provided by the data about these variables from a group of patients with nephropathy as well as from a group of patients without the disease. The analysis will be discussed in detail in the next section.

Another approach to the problem is to utilize multivariate methods to examine the possible risk factors simultaneously. We shall apply linear logistic regression method in a later paper. This method can be used to identify risk factors and predict the probability of developing diabetic nephropathy for individuals as a function of the risk factors.

## Proteinuria

Research suggests that measurement of proteinuria is the most accurate way for the screening and diagnosis of overt diabetic nephropathy. Moreover, protein measurement in spot urine is a reliable and simple method for the screening and diagnosis of overt diabetic nephropathy [11].

Persistent proteinuria is defined as a protein excretion $>0.5 \mathrm{~g} / 24 \mathrm{~h}$ in at least four consecutive urine samples with an interval of at least 1 month in patients without renal infection. Persistent proteinuria is strongly associated with increased mortality in insulin-dependent diabetes mellitus (IDDM), and risk of this condition can be predicted many years in advance by subclinical increases in albumin excretion rate (microalbuminuria) [2]. It was found that the reduction in albumin excretion rate was accompanied by a significant fall in median glomerular filtration rate (GFR) and a fractional renal clearance of albumin. Kidney volume remained unchanged.

Microalbuminuria, the early phase of diabetic nephropathy, is associated with increased cardiovascular morbidity and mortality, but the reason for this is not clear [4]. Patrick et al., [6] assessed the prevalence of microalbuminuria, and its associations with other clinical features. The study showed that persistent microalbuminuria was found in a significant number of non-insulin-dependent diabetes (NIDDM) patients at the time of diagnosis.

Similarly, the impact of microalbuminuria on mortality among a large cohort of NIDDM and other risk factors was investigated by Schmitz et al., [7]. They found that age, urine albumin concentration (UAC), known duration, and serum creatinine were the only significant risk factors. More specifically, Turtle [8] explained that patients with microalbuminuria have an increased risk of developing diabetic nephropathy, hypertension, large vessel disease and retinopathy. Hence epidemiological studies have focused on the identification of risk factors for the development of microalbuminuria. In the same way, the United Kingdom Prospective Diabetes Study Group [9] concluded that urinary albumin excretion was associated with hyperglycaemia and hypertension, whereas urinary N -acetyglycaeminidase was primary associated with hyperglycaemia.

Gall et al., [3] also found that hypertension plays an important role in the increased excretion of albumin and immunoglobulin G (IgG)'s observed in the study. They suggested that impaired barrier size selectivity, probably due to an increase in large pore area in the glomerular capillary wall, and systemic hypertension are the major pathogenetic mechanisms of proteinuria in NIDDM patients with diabetic nephropathy. These results were also consistent with those of Mattock et al., [5] who found an association between overnight urinary albumin excretion rate and prevalent coronary heart disease and concluded from the study that albumin excretion rate is significantly associated with coronary heart disease mortality (after taking into account the confounding effects of raised blood pressure and other cardiovascular risk factors).

In order to evaluate the effects of dietary protein restriction Ciavarella et al., [1] evaluated the effects of dietary protein restriction (low protein diet) on the progression of renal disease in

IDDM patients with clinical nephropathy. They found that early restriction of dietary protein contents, together with an early aggressive treatment of arterial hypertension would be expected to delay the progression of diabetic nephropathy.

## Data

The longitudinal data of the diabetic patients have been collected by Professor David Owens CBE and his team at the Diabetes Research Unit, University of Cardiff School of Medicine, with which the authors have worked over the years. The data set, displayed in Table 1, consists of measurements of age, sex, weight, height, blood glucose, cholesterol, high and low density lipoprotein, blood pressure, urea and creatinine, together with various demographic and biochemical data for 267 NIDDM patients ( 75 females, 192 males).

Table 1. Summarized statistics of all the metric variables

| Variable | N | Minimum | Maximum | Mean | Std. deviation |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Age (year) | 267 | 17.00 | 78.00 | 52.73 | 10.97 |
| BMI (kg/m²) | 267 | 17.15 | 48.73 | 28.66 | 5.08 |
| Cholesterol(mmol/l) | 248 | 2.80 | 11.40 | 5.33 | 1.24 |
| Creatinine ( $\mu \mathrm{mol} / \mathrm{l}$ ) | 257 | 44.00 | 176.00 | 86.00 | 17.25 |
| Dia_BP (mm Hg) | 262 | 58.00 | 120.00 | 85.08 | 10.00 |
| $\mathrm{HBA}_{1}$ (\%) | 265 | 6.70 | 19.30 | 11.20 | 2.40 |
| HDL ( $\mathrm{mmol} / \mathrm{l}$ ) | 243 | 0.40 | 2.20 | 1.08 | 0.32 |
| Height (m) | 267 | 1.49 | 1.95 | 1.69 | 0.09 |
| LDL (mmol/l) | 239 | 0.70 | 7.50 | 3.25 | 1.06 |
| MTT ( $\mathrm{mmol} / \mathrm{l}$ ) | 267 | 5.60 | 20.50 | 11.75 | 3.37 |
| OGTT ( $\mathrm{mol} / \mathrm{l}$ ) | 267 | 5.80 | 22.80 | 12.00 | 3.50 |
| Sys_BP (mmHg) | 262 | 92.00 | 210.00 | 138.80 | 20.40 |
| Triglyceride (mmol/l) | 248 | 0.40 | 10.70 | 2.10 | 1.30 |
| Urea (mmol/l) | 261 | 2.20 | 11.50 | 5.40 | 1.30 |
| Weight (kg) | 267 | 45.00 | 135.50 | 81.50 | 14.70 |
| Valid N(listwise) | 232 |  |  |  |  |

However, it is more informative to compare the means of different variables of the patients' characteristics classified by sex. The results are tabulated in Table 2.

Now from Table 2 it can be seen that apart from the systolic blood pressure, cholesterol, diastolic blood pressure, $\mathrm{HbA}_{1}$ and triglyceride, all other variables shows a significant mean difference ( $p<0.05$ ) between the males and females. This shows that the patients' characteristics are, in many aspects, different between opposite sexes. For this reason we must be very careful when analysing the data as gender may be an important confounding variable that distorts the existing real relationship.

## Age

In general the mean age of the female patients is lower than that of the males (that is, mean age $50(\mathrm{~F})$ vs $53(\mathrm{M})$ ). In order to investigate whether there is any association between age of patients and the presence of proteinuria, a chi-square test for the $2 \times 4$ contingency tables will be adopted. This test is particularly applicable in our present situation as the risk variable (age, in this case) is classified into four categories. Using a chi-square test one can test
whether there is any significant association between the risk variable (age) and the outcome variable (proteinuria). Hence the chi-square tests will be used for other categorical risk variables as well.

Table 2. Means of patient characteristics

| Variable | Female <br> mean | SD | $\mathbf{N}$ | Male <br> mean | SD | $\mathbf{N}$ | $p$ Value |
| :--- | ---: | :---: | :---: | ---: | ---: | ---: | ---: |
| Age | 50.38 | 11.25 | 75 | 53.64 | 10.29 | 192 | 0.029 |
| BMI | 31.10 | 6.13 | 75 | 27.71 | 4.27 | 192 | $<0.000$ |
| Cholesterol | 5.56 | 1.30 | 71 | 5.23 | 1.35 | 177 | 0.059 |
| Creatinine | 73.02 | 12.48 | 71 | 90.95 | 16.65 | 186 | $<0.000$ |
| Dia_BP $_{\text {HBA }_{1}}$ | 83.22 | 10.43 | 73 | 85.79 | 9.41 | 189 | 0.062 |
| HDL $_{\text {Height }}$ | 11.69 | 2.42 | 73 | 11.06 | 2.27 | 192 | 0.056 |
| LDL | 1.15 | 0.34 | 71 | 1.04 | 0.32 | 172 | 0.016 |
| MTT | 1.59 | 0.066 | 75 | 1.72 | 0.079 | 192 | $<0.000$ |
| OGTT | 12.45 | 1.03 | 71 | 3.15 | 1.07 | 168 | 0.048 |
| Sys_BP | 12.70 | 3.35 | 75 | 11.49 | 3.36 | 192 | 0.042 |
| Triglyceride | 139.30 | 22.31 | 75 | 11.68 | 3.466 | 192 | 0.033 |
| Urea | 1.97 | 1.18 | 71 | 138.58 | 19.46 | 189 | 0.800 |
| Weight | 79.53 | 1.53 | 72 | 5.18 | 1.19 | 177 | 0.278 |

After careful examination of the results derived from the chi-square tests the number of possible variables can be reduced and eventually one can identify the preferable risk factors which can then be the subject of further investigations. For the present case initially the age group was divided into four categories: (1) child 0-15; (2) young > 15-30; (3) middle age $>30-55$; (4) old > 55. Then the SPSS printout shown in Table 3-1, 2 indicates that there is significant association between age and proteinuria ( $p=0.0138$ ). On careful examination of Fig. 1 we also note that among all age groups the proportion of patients developing proteinuria is highest with the > 30-55 group (33.8\%) compared with no-one developing proteinuria among the young (> 15-30 group) and only $15.4 \%$ in the $>55$ group. Hence we have evidence to conclude that age may be a risk factor for nephropathy.

Table 3. Patients classified by age

## 1. Coded age * PROTEINURIA Crosstabulation

|  |  | PROTEINURIA |  | Total |  |
| :--- | :--- | :--- | :---: | :---: | :---: |
|  |  | absence | presence |  |  |
| coded <br> age | young > 15-30 | Count | 6 |  | 6 |
|  |  | \% within coded age | $100.0 \%$ |  | $100.0 \%$ |
|  | $>30-55$ | Count | 53 | 27 | 80 |
|  |  | \% within coded age | $66.3 \%$ | $33.8 \%$ | $100.0 \%$ |
|  |  | Count | 55 | 10 | 65 |
|  |  | \% within coded age | $84.6 \%$ | $15.4 \%$ | $100.0 \%$ |
|  |  | Count | 114 | 37 | 151 |
|  |  | \% within coded age | $75.5 \%$ | $24.5 \%$ | $100.0 \%$ |

## 2. Chi-Square Tests

|  | Value | df | Asymp. sig. <br> (2-sided) |
| :--- | :---: | :---: | :---: |
| Pearson Chi-Square | $8.567^{\mathrm{a}}$ | 2 | 0.014 |
| Likelihood Ratio | 10.047 | 2 | 0.007 |
| Linear-by-Linear | 2.225 | 1 | 0.136 |
| Association |  |  |  |
| N of Valid Cases | 151 |  |  |

${ }^{\text {a }} 2$ cells (33.3\%) have expected count less than 5.
The minimum expected count is 1.47 .
Moreover, when the patients are classified by sex (Table 4-1,2), it can be shown that for male patients there is significant association $(p=0.010)$ between age and proteinuria, while for the female patients no such significant association is observed ( $p=0.333$ ).

## Body mass index (BMI)

The mean body mass index (BMI) for the females is larger than that for the males in the sample. This can be shown by the boxplot in Fig. 1.


Fig. 1 Boxplots of BMI
The mean BMIs for females and males are 31.09 and 27.72 respectively. This shows that the proportion of the female patients with obesity is larger than that of male patients, though actually both of them are obese. This is consistent with what we expected as we know that obesity has a strong association with diabetes.

An attempt has been made to classify those patients with BMI > 26.5 as obese and tests were carried out to investigate any association between BMI and the presence of proteinuria. However, no significant association was observed. Even when the patients are classified by gender the result is still not significant. But when retinopathy is taken into consideration, Table $5-1$, 2 shows a significant association between males and BMI ( $p=0.024$ ). However, it remains a problem to see how BMI and retinopathy are related and this problem needs further investigation.

Table 4. Patients classified by sex

1. Coded age * PROTEINURIA * SEX Crosstabulation

| SEX |  |  |  | PROTEINURIA |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | absence | presence |  |
| Female | coded <br> age | young > 15-30 | Count \% within coded age | $\begin{gathered} 6 \\ 100.0 \% \\ \hline \end{gathered}$ |  | $\begin{gathered} 6 \\ 100.0 \% \\ \hline \end{gathered}$ |
|  |  | > 30-55 | Count \% within coded age | $\begin{aligned} & \hline 16 \\ & 72.7 \% \\ & \hline \end{aligned}$ | $\begin{gathered} \hline 6 \\ 27.3 \% \\ \hline \end{gathered}$ | $\begin{gathered} 22 \\ 100.0 \% \\ \hline \end{gathered}$ |
|  |  | > 55 | Count $\%$ within coded age | $\begin{gathered} 9 \\ 81.8 \% \end{gathered}$ | $\begin{gathered} 2 \\ 18.2 \% \end{gathered}$ | $\begin{gathered} 11 \\ 100.0 \% \end{gathered}$ |
|  | Total |  | Count \% within coded age | $\begin{aligned} & 31 \\ & 79.5 \% \\ & \hline \end{aligned}$ | $\begin{gathered} 8 \\ 20.5 \% \\ \hline \end{gathered}$ | $\begin{aligned} & 139 \\ & 100.0 \% \\ & \hline \end{aligned}$ |
| Male | coded age | > 30-55 | Count \% within coded age | $\begin{aligned} & \hline 37 \\ & 63.8 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 21 \\ & 36.2 \% \\ & \hline \end{aligned}$ | $\begin{gathered} \hline 58 \\ 100.0 \% \\ \hline \end{gathered}$ |
|  |  | > 55 | Count \% within coded age | $\begin{aligned} & \hline 46 \\ & 85.2 \% \\ & \hline \end{aligned}$ | $\begin{gathered} 8 \\ 14.8 \% \\ \hline \end{gathered}$ | $\begin{gathered} 54 \\ 100.0 \% \\ \hline \end{gathered}$ |
|  | Total |  | Count $\%$ within coded age | $\begin{aligned} & 83 \\ & 74.1 \% \end{aligned}$ | $\begin{aligned} & 29 \\ & 25.9 \% \end{aligned}$ | $\begin{aligned} & 112 \\ & 100.0 \% \end{aligned}$ |

## 2. Chi-Square Tests ${ }^{\text {a }}$

| SEX |  | Value | df | Asymp. sig. <br> (2-sided) | Exact. sig. <br> (2-sided) | Exact. sig. <br> (2-sided) |
| :--- | :--- | :---: | :---: | :---: | :---: | :---: |
| Female | Pearson Chi-Square | $2.202^{\mathrm{b}}$ | 2 | 0.333 |  |  |
|  | Likelihood Ratio | 3.367 | 2 | 0.186 |  |  |
|  | Linear-by-Linear | 0.347 | 1 | 0.556 |  |  |
|  | Association |  |  |  |  |  |
|  | N of Valid Casses | 39 |  |  |  |  |
| Male | Pearson Chi-Square | $6.69^{\mathrm{C}}$ | 1 | 0.010 |  |  |
|  | Likelihood Ratio | 5.601 | 1 | 0.18 |  | 0.008 |
|  | Fisher's Exact Test | 6.875 |  |  | 0.017 | 0.009 |
|  | Linear-by-Linear | 6.610 | 1 | 0.009 |  |  |
|  | Association |  | 1 | 0.010 |  |  |
|  | N of Valid Casses | 112 |  |  |  |  |

${ }^{\text {a }}$ Computed only for $2 \times 2$ table.
${ }^{\mathrm{b}} 4$ cells $(66.7 \%)$ have expected count less than 5 . The minimum expected count is 1.23.
${ }^{\mathrm{c}} 0$ cells $(0.0 \%)$ have expected count less than 5 . The minimum expected count is 13.98.

## Blood glucose

According to the World Health Organization [10] a patient is classified as having diabetes if the fasting blood glucose in OGTT is $\geq 7.8 \mathrm{mmol} / \mathrm{L}$. Now from Table 2 the mean fasting blood glucose (F-BG) in OGTT of female patients is higher than that of the male patients ( 12.7 vs 11.68 respectively). Similar results are also obtained from the meal tolerance test (MTT). Now for both tests the F-BG of both male and female patients are around 11 to $12 \mathrm{mmol} / \mathrm{L}$. Hence there is no doubt that they have been correctly diagnosed as diabetic patients. For reference both the boxplots of MTT and OGTT classified by sex are listed below in Figs. 2 and 3 respectively.

Table 5. Associations between BMI and retinopathy for patients classified by sex

1. Fat * RETINOPATHY * SEX Crosstabulation

| SEX |  |  |  | RETINOPATHY |  |  |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | absence | presence | moderate | serious |  |
| Female | Fat <br> Total | normal | Count \% within Fat | $\begin{aligned} & 12 \\ & 75.0 \% \end{aligned}$ | $\begin{aligned} & \hline 1 \\ & 6.3 \% \end{aligned}$ | $\begin{gathered} 2 \\ 12.5 \% \\ \hline \end{gathered}$ | $\begin{aligned} & 1 \\ & 6.3 \% \end{aligned}$ | $\begin{array}{\|c\|} \hline 16 \\ 100.0 \% \\ \hline \end{array}$ |
|  |  | fat > 26.5 | Count \% within Fat | $\begin{aligned} & \hline 49 \\ & 83.1 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 4 \\ & 6.8 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 3 \\ & 5.1 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 3 \\ & 5.1 \% \\ & \hline \end{aligned}$ | $\begin{array}{\|l\|} \hline 59 \\ 100.0 \% \\ \hline \end{array}$ |
|  |  |  | Count \% within Fat | $\begin{aligned} & \hline 61 \\ & 81.3 \% \end{aligned}$ | $\begin{aligned} & \hline 5 \\ & 6.7 \% \end{aligned}$ | $\begin{aligned} & \hline 5 \\ & 6.7 \% \end{aligned}$ | $\begin{aligned} & \hline 4 \\ & 5.3 \% \end{aligned}$ | $\begin{array}{\|l\|} \hline 75 \\ 100.0 \% \\ \hline \end{array}$ |
| Male | Fat | normal | Count \% within Fat | $\begin{aligned} & \hline 58 \\ & 68.2 \% \end{aligned}$ | $\begin{aligned} & \hline 5 \\ & 5.9 \% \end{aligned}$ | $\begin{aligned} & 16 \\ & 18.8 \% \end{aligned}$ | $\begin{aligned} & \hline 6 \\ & 7.1 \% \end{aligned}$ | $\begin{array}{\|l\|} \hline 85 \\ 100.0 \% \end{array}$ |
|  |  | fat > 26.5 | Count \% within Fat | $\begin{aligned} & \hline 92 \\ & 86.0 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 3 \\ & 2.8 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 7 \\ & 6.5 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 5 \\ & 4.7 \% \end{aligned}$ | $\begin{array}{\|l\|} \hline 107 \\ 100.0 \% \\ \hline \end{array}$ |
|  | Total |  | Count \% within Fat | $\begin{aligned} & \hline 150 \\ & 78.1 \% \end{aligned}$ | $\begin{aligned} & \hline 8 \\ & 4.2 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 23 \\ & 12.0 \% \\ & \hline \end{aligned}$ | $\begin{gathered} \hline 11 \\ 5.7 \% \\ \hline \end{gathered}$ | $\begin{array}{\|l\|} \hline 192 \\ 100.0 \% \\ \hline \end{array}$ |

2. Chi-Square Tests

| SEX |  | Value | df | Asymp. sig. <br> (2-sided) |
| :--- | :--- | :---: | :---: | :---: |
| Female | Pearson Chi-Square <br>  <br> N of Valid Casses | $1.176^{\mathrm{a}}$ <br> 75 | 3 | 0.759 |
| Male | Pearson Chi-Square <br>  <br>  <br> N of Valid Casses | $9.422^{\mathrm{b}}$ <br> 192 | 3 | 0.024 |

${ }^{\mathrm{a}} 6$ cells ( $75.0 \%$ ) have expected count less than 5 .
The minimum expected count is 0.85 .
${ }^{\mathrm{b}} 3$ cells ( $37.5 \%$ ) have expected count less than 5 . The minimum expected count is 3.54 .


SEX
Fig. 2 Boxplots of meal tolerance tests


SEX
Fig. 3 Boxplots of oral glucose tolerance tests

## Blood pressure

From Table 2 the systolic and diastolic blood pressures of both sexes are about the same. Comparing these with the criteria for hypertension (sys/dia $\geq 140 / 90$ ), the female: male ratio for mean systolic BP is 139.3:138.59, and for the mean diastolic BP, the mean value for female: male ratio is 83.22 : 85.79 . So apparently hypertension is quite common among these diabetic patients.

In order to investigate whether the presence of proteinuria among diabetic patients has any association with blood pressure, the data for the systolic and diastolic BP are coded according to the 140/90 (sys BP/dia BP) criteria above which are classified as hypertension. Now we divide the patients into two groups: those whose BPs are above the set criterion are classified as having hypertension and vice versa.

From Table $6-1$, 2 it is observed that since there are only 37 female patients whose proteinuria level is recorded, the number of patients falling in each cells is small. Nevertheless, Fisher's exact test ( $p<0.05$ ) shows that hypertension in terms of diastolic BP for female patients has an association with proteinuria. However, similar testing for males does not indicate any significant association. Furthermore, it is interesting to note that if the whole group of patients is taken into consideration irrespective of sex, then the result is also non-significant. Hence sex may be a confounding factor. However, results indicate that there is no significant association between hypertension in terms of systolic blood pressure and proteinuria for both male and female patients.

## Lipid levels

By referring to Table 2, and also by comparing the boxplots in Figs. 4 - 7, except for triglycerides (TG), all the other measures of lipid levels, (that is, cholesterol (CHOL), high density lipoprotein (HDL), and low density lipoprotein (LDL)) show that the female patients have higher values than those of males.

Table 6. Association between proteinuria and age group by sex

1. Coded diabp * PROTEINURIA * SEX Crosstabulation

| SEX |  |  |  | PROTEINURIA |  | Total |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  | absence | presence |  |
| Female | coded <br> diabp | normal | Count \% within coded diabp | $\begin{aligned} & \hline 25 \\ & 86.2 \% \end{aligned}$ | $\begin{gathered} 4 \\ 13.8 \% \end{gathered}$ | $\begin{gathered} \hline 29 \\ 100.0 \% \end{gathered}$ |
|  |  | high diabp | Count \% within coded diabp | $\begin{gathered} 4 \\ 50.0 \% \end{gathered}$ | $\begin{gathered} 4 \\ 50.0 \% \end{gathered}$ | $\begin{gathered} 8 \\ 100.0 \% \end{gathered}$ |
|  | Total |  | Count \% within coded diabp | $\begin{aligned} & 29 \\ & 78.4 \% \\ & \hline \end{aligned}$ | $\begin{gathered} \hline 8 \\ 21.6 \% \end{gathered}$ | $\begin{gathered} \hline 37 \\ 100.0 \% \end{gathered}$ |
| Male | coded diabp | normal | Count \% within coded diabp | $\begin{aligned} & \hline 57 \\ & 72.2 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 22 \\ & 27.8 \% \\ & \hline \end{aligned}$ | $\begin{gathered} \hline 79 \\ 100.0 \% \end{gathered}$ |
|  |  | high diabp | Count \% within coded diabp | $\begin{aligned} & 24 \\ & 80.0 \% \end{aligned}$ | $\begin{gathered} 6 \\ 20.0 \% \end{gathered}$ | $\begin{gathered} 30 \\ 100.0 \% \end{gathered}$ |
|  | Total |  | Count \% within coded diabp | $\begin{aligned} & 81 \\ & 74.3 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 28 \\ & 25.7 \% \\ & \hline \end{aligned}$ | $\begin{aligned} & \hline 109 \\ & 100.0 \% \end{aligned}$ |

## 2. Chi-Square Tests

| SEX |  | Value | df | Asymp. sig. (2-sided) | Exact. sig. (2-sided) | Exact. sig. (2-sided) |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Female | Pearson Chi-Square Fisher's Exact Test N of Valid Casses | $\begin{aligned} & 4.850^{b} \\ & 37 \\ & \hline \end{aligned}$ | 1 | 0.028 | 0.049 | 0.049 |
| Male | Pearson Chi-Square Fisher's Exact Test N of Valid Casses | $\begin{array}{\|c\|} \hline 0.702^{\mathrm{c}} \\ 109 \end{array}$ | 1 | 0.402 | 0.470 | 0.281 |

${ }^{\mathrm{b}} 1$ cells $(25.0 \%)$ have expected count less than 5 . The minimum expected count is 1.73 . ${ }^{\text {c }} 0$ cells $(0.0 \%)$ have expected count less than 5 . The minimum expected count is 7.71 .

In comparison with the normal ranges, the ratio of mean cholesterol for female: male patients is $5.56: 5.23$, while the normal range is $3.5-7.0 \mathrm{mmol} / \mathrm{L}$. Thus the NIDDM patients under consideration have normal cholesterol levels. For TG, the ratio of mean TGs for female: male is $1.973: 2.181$. Comparison of this with the normal range of $0.3-1.8 \mathrm{mmol} / \mathrm{L}$ shows that both female and male patients have higher TG values than the healthy subjects. The normal ranges for LDL and HDL are $1.55-4.4 \mathrm{mmol} / \mathrm{L}$ and $0.9-1.93 \mathrm{mmol} / \mathrm{L}$ respectively. While the ratios of means for female: male (LDL) are 3.45:3.15, and for that of (HDL) are 1.15 and 1.04. Hence both mean values of LDL and HDL of patients are within the normal ranges for both sexes.


Fig. 4 Boxplots of triglycerides


SEX
Fig. 6 Boxplots of LDL

Fig. 5 Boxplots of HDL


SEX
Fig. 7 Boxplots of cholesterol

## Renal functions

In this case for both the 'urea' and 'creatinine' levels, the values of the female patients are lower than those of the males (that is, for urea, the female: male ratio is $5.08: 5.57$ ), for
creatinine, the female: male ratio is 73.03:90.95) in Figs. 8 and 9. Comparing these with the normal range for urea, $3.0-6.5 \mathrm{mmol} / \mathrm{L}$, and the normal range for creatinine $60-120 \mu \mathrm{~mol} / \mathrm{L}$, we can see that the mean values of urea and creatinine of the male/female patients are all within the normal range though the urea values are towards the upper end of normality, particularly for the males. As these renal functions are considered to be closely related to the development of renal disease, the next step is to investigate the association of these renal functions with other variables, in particular, their relationship with the presence of proteinuria.


Fig. 8 Boxplots of creatinine


Fig. 9 Boxplots of urea

## Conclusion

In order to identify those risk factors which are closely related to the development of nephropathy, we choose the presence/absence of proteinuria as a landmark to discriminate those patients with/without nephropathy, respectively. Using proteinuria status as an index we can test all the variables considered so far and check whether their means are significantly different (that is, $p<0.05$ ) for different proteinuria status. However, if we consider the set of patients as a whole we cannot observe any significant difference between the means of the variables of different proteinuria status. For further investigation we subdivide the patients into different categories, namely, classified by sex, classified by smoking habit and by state of hypertension (SBP/DBP). After detailed analysis, those variables which have significant difference in means (that is, $p$ value less than or close to 0.05 ) with different proteinuria status are given in Table 7. Here the definition for hypertension follows the rule as $\mathrm{SBP} \geq 140$ or DBP $\geq 90$.

Table 7. Summary results

| Patient characteristics | Variable | Proteinuria status |  |  |  | $\boldsymbol{p}$ value |
| :--- | :--- | ---: | ---: | ---: | :---: | :---: |
|  |  | No |  | Yes |  |  |
|  |  | SD | Mean | SD |  |  |
| Male |  | 56.13 | 10.35 | 50.00 | 8.84 | 0.003 |
| Male | Triglyceride | 2.08 | 1.32 | 1.648 | 0.567 | 0.027 |
| Smoker | Triglyceride | 2.54 | 1.72 | 1.77 | 0.51 | 0.031 |
| Male non-smoker | Age | 56.40 | 11.18 | 51.19 | 8.06 | 0.046 |
| Male non-smoker | MTT | 11.59 | 3.55 | 9.71 | 2.39 | 0.019 |
| Male smoker | Age | 55.61 | 8.65 | 48.54 | 9.84 | 0.038 |
| Male smoker | Triglyceride | 2.48 | 1.84 | 1.70 | 0.45 | 0.055 |
| Male hypertension (SBP) | Age | 57.37 | 10.60 | 52.55 | 8.91 | 0.059 |
| Male hypertension (SBP) | LDL | 3.11 | 0.87 | 2.37 | 1.26 | 0.047 |

From Table 7 we can also see that age appears as an important risk factor among different categories, namely, male patients, male non-smoker, male smoker, and male with high SBP. On the other hand, triglyceride also shows significant difference among those patient groups who are male or male smokers. Similarly, MTT for male smokers also shows significantly different values for different proteinuria status. So these variables will need to be taken into further investigation. Furthermore, LDL for male with high DBP also indicates a significant difference in values for different proteinuria status.

To summarize, we can observe from Table 7 that though we cannot find any significant difference in values for all the variables by simply taking all the diabetic patients into consideration, we did, however, when subdividing the patients into different categories according to status of sex, smoking habit and hypertension, identify some of the risk factors which behave significantly differently for different proteinuria states. So at this stage we have obtained some preliminary identification of the risk factors by analysing the variables individually under different classification of patients' characteristics.

## References

1. Ciavarella A., D. M. Mizio, S. Stefoni, C. Campieri, L. C. Borgino, P. Vannini (1987). Earlt Restriction of Dietary Protein Intake Reduces Albuminuria, Frontiers in Diabetes, 8, 210-212.
2. Cohen D., R. Dodds, G. Viberti (1987). Effect of Protein Restriction in Insulin Dependent Diabetics at Risk of Nephropathy, British Medical Journal, 294, 795-798.
3. Gall M-A., A. Kofoed-Enveoldsen, F. S. Nielsen, H. H. Parving (1994). Glomerular Size and Charge Selectivity in Type 2 (Non-Insulin Dependent) Diabetic Patients with Diabetic Nephropathy, Diabetologia, 37, 195-201.
4. Gruden G., P. Cavallo-Perin, R. Romagnoli, C. Olivetti, D. Frezel, G. Pagano (1994). Prothrombin Fragment 1+2 Antithrombin III-thrombin Complex in Microalbumineric Type 2 Diabetic Patients, Diabetic Medciine, 11, 485-488.
5. Mattock M. B., H. Keen, G. C. Viberti, M. R. El-Gohari, T. J. Murells, G. S. Scott, J. R. Wing, P. G. Jackson (1988). Coronary Heart Disease and Urinary Albumin Excretion Rate in Type 2 (Non-Insulin Dependent) Diabetes Patients, Diabetologia, 31, 82-87.
6. Patrick A. W., P. J. Leslie, B. F. Clarke, B. M. Frier (1990). The Natural History and Associations of Microalbuminuria in Type 2 Diabetes During the First Year After Diagnosis, Diabetic Medicine, 7, 902-908.
7. Schmitz A., M. Veth (1988). Microalbuminuria: a Major Risk Factor in Non-InsulinDependent Diabetes. A 10-year Follow-up Study of 503 Patients, Diabetic Medicine, 5, 126-134.
8. Turtle J. R. (1995). Healthcare Models in Diabetes, Dialogue, 1, 5-7.
9. United Kingdom Prospective Diabetes Study Group (1993). Reltionships of Urinary Albumin and N -acetylglucosaminidase to Glycaemia and Hypertension at Diagnosis of Type 2 (Non-Insulin Dependent) Diabetes Melitus and after 3 Months Diet Therapy, Diabetologia, 36, 835-842.
10. World Health Organisation (1985). Diabetes Mellitus. Technical Report 727, WHO, Geneva.
11. Zelmanovitz T., J. L. Gross, J. Oliveira, M. J. de Azevedo (1998). Proteinuria is Still Useful for the Screening and Diagnosis of Overt Diabetic Nephropathy, Diabetes Care, 21, 1076-1079.
