Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Original Research Article

Early stage of chronic kidney disease by using statistical evaluation of the previous measurement results

Selahaddin Batuhan Akben *
Bahce Vocational School, Osmaniye Korkut Ata University, Turkey

ARTICLE INFO

Article history:
Received 13 April 2016
Received in revised form
27 July 2016
Accepted 7 August 2016
Available online xxx

Keywords:
Chronic kidney disease
Chronic renal failure
Statistical evaluation
Signal processing
CKD

ABSTRACT

Chronic kidney disease (CKD) that causes the progressive losses in kidney functions is one of the major public health problems. Expert medical doctors can diagnose the CKD through symptoms, blood and urine tests in its early stages. However, the diagnosis of CKD might be difficult for expert medical doctors in case of the questionable measurement result. Therefore a new mathematical method that would be helpful to the expert medical doctors is required. It can be said that there is no studies related with automatic diagnosis of CKD in the literature. This study aims to remedy this shortcoming in the literature. In this study, for each of test and symptom values, averages of measurement results were calculated separately for healthy subjects and patients. Then the measured values were marked as "0" or "1" (healthy or patient) according to closeness to average values. Finally, the classification was performed by averaging the values marked for each subject. The success rate of the proposed method is between 96% and 98% according to the age ranges. In conclusion section of the study, how to implement the proposed method in real life is offered.

1. Introduction

Chronic kidney disease (CKD), also known as chronic renal disease, is a major public health problem [1,2]. Since CKD is a heterogeneous disorder affecting kidney structure and functions, it causes the progressive losses in kidney functions [3]. As a result, end stages of the CKD may require dialysis or kidney transplantation for patients. Therefore the disease should be diagnosed by expert medical doctors in its early stages [4].

Diagnosis methods of the CKD are generally based on assessing the glomerular filtration rate (GFR), routine laboratory tests, and symptoms (answer/question) [5]. However, since main purpose of GFR is to determine the stages of disease, GFR is not usually used primarily to diagnose the disease in the early stages [6]. Instead of it, the CKD is most frequently diagnosed through symptoms, blood and urine tests in its early stages [7,8].

However, the test and symptoms results may be deceptive than the results of glomerular filtration rate [9]. The reason for

* *Corresponding author at*: Bahce Vocational School, Osmaniye Korkut Ata University, Turkey.
E-mail address: batu130@hotmail.com
http://dx.doi.org/10.1016/j.bbe.2016.08.004

Please cite this article in press as: Akben S.B. Early stage of chronic kidney disease by using statistical evaluation of the previous measurement results. Biocybern Biomed Eng (2016), http://dx.doi.org/10.1016/j.bbe.2016.08.004

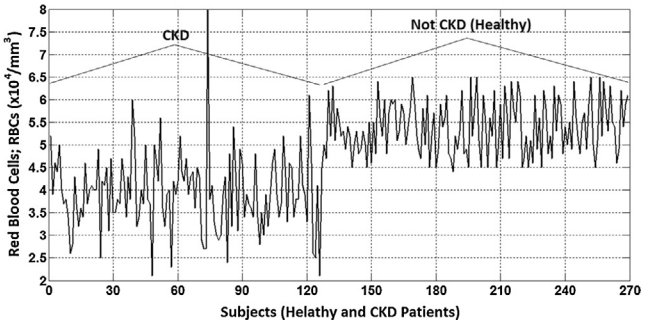

Fig. 1 – RBCs for the CKD patients and healthy ones.

this problem: To diagnose the CKD in early stage, result of the tests and symptoms is compared with the normal-ranges determined by previous experiences [10,11]. But, while some results may indicate disease, others may be in normal-ranges. For example, normal-range of red blood cells (RBCs) is 4.3–6 $\times$ 10^4/mm^3. Fig. 1 shows an example of this problem.

As seen in Fig. 1, RBCs value of some CKD patients is in normal range. These misleading results may be due to many different reasons.

Expert medical doctors can overcome this problem with their experience. However, some exceptions related with the measurement results can also be misleading for expert medical doctors. For example, while half of measurement results indicate the disease the other half not.

Therefore a new mathematical method that would be helpful to the expert medical doctors is required (method can be able to simulate the doctor decision process). Moreover, accurate diagnosis decisions of medical doctors can also be supported mathematically thanks to the new method. However, it should be noted that the new method should be understandable in order to applicability by medical doctors.

Studies related to the diagnosis of CKD by using statistical and signal processing methods are not much in the literature [12,13]. Even, it can be said that there is no studies related with automatic diagnosis of CKD in the literature, since previous studies are mostly related with improve and predict the normal ranges [14,15]. Few studies related with automatic diagnosis of CKD in the literature are difficult to be understood by medical doctors, because are not suitable for use in real life and include complex mathematical equations [16,17]. Therefore, this study aims to remedy this shortcoming in the literature by proposing a simple and effective method understandable by medical doctors.

In this study, for each of test and symptom values, averages of measurement results were calculated separately for healthy subjects and patients. Then the measurement result values were digitized (marked) as healthy or patient according to closeness to average values. Finally, the classification was performed by averaging the values marked (measurement results marked before) for each subject. Thus, a new method easily understandable was presented. In addition, result was compared and approved by signal processing methods. Moreover, the statistical average values of both the healthy and the CKD patients have also been presented for each of the tests and symptoms.

2. Materials

The dataset used in this study was collected from Apollo Hospital in India. It consists of tests (blood and urine tests) and symptoms results of 400 subjects. The subjects were diagnosed by nephrologist whether CKD patient or healthy. Namely, with using this dataset it is aimed to analyze the measurement results based on whether the subjects are CKD patient or not. The diagnostic step was performed as follows.

In the first step, the personal and family histories of subjects were evaluated by nephrologists. Then, according to the national kidney foundation criteria, the nephrologists asked to subjects whether they had symptoms. Next, some certain tests (blood and urine tests) have been performed to subjects by nephrologists. Again, blood and urine attributes measured in these tests are recognized by national kidney foundation also. Then, the disease classification was made by nephrologists according to symptoms and test results. Finally, the doctor experience based diagnostic results using symptoms and laboratory tests were compared with GFR results, to be sure and diagnosis was made absolute. Essentially, these diagnostic procedures mentioned above are adopted by the national kidney foundation. Therefore the attributes (test and symptom attributes) used by doctors and in this study are adopted also by the national kidney foundation. Moreover, dataset attributes measured by laboratory tests are used also for GFR.

Various attributes are measured during the laboratory tests. However, some attributes only associated with early stage of CKD are used according to the previous experiences and criteria by expert nephrologists. Therefore age, 16 measurement result, 6 symptom and blood pressure were used as attributes in this study. Since the diagnostic results obtained using these attributes were approved comparing with GFR results, use of these attributes is reliable.

Measured values (attributes used in this study) are respectively: age, blood pressure (bp), specific gravity (sg), albumin (al), sugar (su), red blood cells (rbc), pus cell (pc), pus cell clumps (pcc), bacteria (ba), blood glucose random (bgr), blood urea (bu), serum creatinine (sc), sodium (sod), potassium (pot), hemoglobin (hemo), packed cell volume (pcv), white blood cell count (wc), red blood cell count (rc), hypertension (htn), diabetes mellitus (dm), coronary artery disease (cad), appetite (appet), pedal edema (pe), anemia (ane).

In addition, units of attributes can be seen in Table 1. Since expert medical doctors evaluate the albumin and sugar values by levels, these values were assigned between 0 and 5 (from very low to very high).

The non-numeric values were marked as ‘‘0’’ or ‘‘1’’ to analyze numerically: Non-numeric values indicating the presence of chronic kidney disease was assumed as ‘‘1’’ and others were assumed as ‘‘0’’. Because class labels of CKD patients are ‘‘1’’ and class labels of healthy subjects are ‘‘0’’.

This data set was taken from publicly available UCI Machine Learning Repository. It is reusable by taking from referenced website [18].

Please cite this article in press as: Akben S.B. Early stage of chronic kidney disease by using statistical evaluation of the previous measurement results. Biocybern Biomed Eng (2016), http://dx.doi.org/10.1016/j.bbe.2016.08.004