

Exploring medical data concerning patients with urolithiasis treated by the extracorporeal shock wave lithotripsy (ESWL)

November 12, 2008

1 Introduction

Physicians examining patients from the point of view of final diagnosis and choice of the proper treatment have to take into account several factors: they collect data from interviewing the patients or clinical investigations, they analyse results of several laboratory tests and image techniques (as e.g. ultrasonography, computer topography and X-ray contrast tests). The patient's status resulting from the analysis of these data and the clinical experience of the physician leads to determination of indications and contr-indications for the particular mode of treatment. The physician's clinical experience is mainly based on his current knowledge and experience with a given disease as well as the therapeutic efficiency of the chosen mode of its treatment.

The progress in medical sciences in last years has significantly increased the amount of available measurement data. These data may have, however, different practical importance. So, they are analysed in order to find and select the most important and valuable data elements for the medical interpretation. Typical tasks in the analysis of the medical data, in particular concerning the problems of diagnosing and/or treatment of a given disease, are the following:

- to identify the most significant attributes for the patients' classification (resulting from the diagnostic and therapeutic point of view),
- to discover the dependencies between values of the significant attributes and the patients' classification.

It must be noticed that the medical data systems are characterized by many properties that make these data specific to analyse, e.g.

- the number of attributes is too large comparing to the number of patients,
- there are too many independent and 'noisy' attributes,
- most of the attributes have a qualitative character,

- classes of the patients' classification are non-balanced taking into account the number of individuals; one class is often a strong majority class,
- observations describing patients may have individual character and be difficult to generalize.

2 The ESWL problem - description of the data set

Data concerning patients with urolithiasis treated by the extracorporeal shock wave lithotripsy (ESWL) were collected at the one clinic of University of Medical Sciences in Poznań. In order to qualify patients for the ESWL treatment different data are taken into account, i.e. anamnesis (i.e. information coming from investigating patients by the physician), laboratory and imaging tests. Although the current experience includes over 1000 patients per year, we could analyse part of it only, i.e. data about patients with completely defined pre-operation attributes and with known and verified long term results of the treatment.

The patients are described by 33 pre-operation attributes currently considered in urological practice. These are the following attributes: 1 - age, 2 - sex, 3 - duration of disease, 4 - type of urolithiasis, 5 - lithuresis, 6 - operations in the past, 7 - nephrectomy, 8 - PCNL, 9 - number of the ESWL treatment previously done, 10 - evacuation of calculi by zeiss cathether, 11 - lumbar region pains, 12 - dyspeptic symptoms, 13 - basic dysuric symptoms, 14 - other dysuric symptoms, 15 - body temperature, 16 - general uriscopy, 17 - urine reaction, 18 - erythrocyturia, 19 - leucocyturia, 20 - bacteriuria, 21 - crystaluria, 22 - proteinuria, 23 - urea, 24 - creatinine, 25 - bacteriological test, 26 - kidney location, 27 - kidney size, 28 - kidney defect, 29 - status of urinary system, 30 - secretion of urinary contrast, 31 - location of the concrement, 32 - calixcalculus, 33 - stone size. Let us notice that nearly all of these attributes have a qualitative character. Their domains usually consist of a limited number of values which are qualitative and linguistic terms. In addition, the domains of many attributes cannot be ordered.

The post-operation conditions of the patients were described by two attributes having the following clinical meaning:

1. A patient's physical condition after the lithotripsy, i.e.:
 - without complications,
 - with complications.
2. Long term results of the treatment:
 - recovery (good results),
 - no recovery,
 - lack of effects.

The both post-operation attributes define two classifications of patients, denoted as \mathcal{Y}_1 and \mathcal{Y}_2 respectively. Values of these classifications will be further called *decision classes*. These classifications are typical standards used to evaluate the medical treatment.