Clinica Chimica Acta 396 (2008) 49–55

Contents lists available at ScienceDirect

Clinica Chimica Acta

journal homepage: www.elsevier.com/locate/clinchim

Establishing age/sex related serum creatinine reference intervals from hospital laboratory data based on different statistical methods

Hans Pottel [a,*], Nicolas Vrydags [b], Boris Mahieu [b], Emmanuel Vandewynckele [b], Kathleen Croes [b], Frank Martens [b]

[a] Interdisciplinary Research Center, Katholieke Universiteit Leuven Campus Kortrijk, Kortrijk, Belgium
[b] Department of Clinical Chemistry, AZ Groeninge Hospital, Kortrijk, Belgium

ARTICLE INFO

Article history:
Received 16 June 2008
Received in revised form 19 June 2008
Accepted 19 June 2008
Available online 23 June 2008

Keywords:
Serum creatinine
Age/sex reference intervals
Non-parametric and parametric methods

ABSTRACT

Background: This is a retrospective study on a large hospital database to establish age- and sex-related mean values and reference ranges for serum creatinine (Scr), obtained with an IDMS-traceable, enzymatic method, in a Caucasian population.
Methods: The database was filtered for unique entries to reduce the presence of correlated and pathological data. Three different statistical methods, a non-parametric method, the Bhattacharya procedure and a non-linear fit of the cumulative Gaussian distribution were used to estimate the serum creatinine–age dependency for men and women, from birth till 100 years of age.
Results: Scr increases with age, equal for boys and girls, up to 14 years and with a much steeper slope for boys than for girls between 14 and 20 years. We show that the Scr–age pattern is constant between 20 and 70 years with a mean of 0.90 mg/dL [0.63–1.16 mg/dL] for men and 0.70 mg/dL [0.48–0.93 mg/dL] for women. Above 70, Scr starts to slowly increase again.
Conclusions: Indirect methods confirm the available reference intervals from healthy-volunteer studies and add information on age-periods not covered by these studies. As such, indirect methods can be used complementary to healthy-volunteer studies.

1. Introduction

Although the use of estimated glomerular filtration rate (eGFR) from statistical models, such as the Modification of Diet in Renal Disease (MDRD) Study equation [1,2], is getting more and more common to assess renal function, reference values for serum creatinine (Scr) may still be necessary to compare patients' results. The reason for this is that statistical models are based on assumptions and have their own limitations. For instance, the MDRD Study equation, which is recommended for diagnosing Chronic Kidney Disease (CKD) by US, UK, New Zealand, Australia and guidelines of other countries, is only valid in adults between 20 and 65–75 years old, with decreased renal function (eGFR <60 ml/min/1.73 m^2) [3–5]. This means that children, the elderly and healthy individuals are not included in the model. Scr reference values may be helpful in these age/sex subgroups to spot abnormalities.

In the past few years, many publications have addressed the topic of the diversity of Scr assays [6,7], each with their specific problems. Most clinical labs are using an alkaline picrate assay (so-called Jaffe method). One major issue was that the alkaline picrate method has shown a considerable lack of specificity. Attempts to correct for this nonspecificity resulted in the use of many different assays among which a compensated Jaffe method, where a correction for a constant bias as compared to the Isotope Dilution Mass Spectrometry (IDMS) reference method has been introduced. Although many attempts to solve this specificity problem, the current tendency amongst clinical labs is the increasing use of enzymatic methods to determine Scr. These methods have the advantage of being IDMS-traceable, that is, mathematical correction factors are no longer necessary. Although not current common practice, the scientific community is clear in its recommendation to use enzymatic assays for the determination of Scr [6]. As a consequence, the number of publications reporting reference values for Scr, determined by enzymatic assays, is very limited up till now, and can only date back to 2002, the time when sufficient commercial enzymatic assays became available. This also demonstrates one of the drawbacks of so-called healthy-volunteers' studies, which are recommended by the IFCC to determine reference values [8,9]: if the biochemical method improves or changes, the study has to be repeated. With the continuous evolution of measurement procedures, this process has to be repeated too often. In such studies, the "a priori" selection of reference individuals for obtaining reference

* Corresponding author. Interdisciplinary Research Center, KU Leuven Campus Kortrijk, E. Sabbelaan 53, B-8500 Kortrijk, Belgium. Tel.: +32 56 246022; fax: +32 56 246994.
E-mail address: hans.pottel@kuleuven-kortrijk.be (H. Pottel).

0009-8981/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.cca.2008.06.017

values is required. This selection is difficult, costly and time consuming: physical examinations, laboratory tests and a medical interview based on a questionnaire should guarantee the health status of the selected individuals. There is no doubt that this reduces the risk of including abnormal values in the data, but it can never be 100% guaranteed. Clinical laboratories, seeking accreditation for compliance with ISO 15189:2003 need to demonstrate that the physiological reference intervals are appropriate for the patient population served and for their measurement systems as well. Giving the cost, time and difficulty of healthy-volunteer studies, few labs do so. Most of them use the reference values from scientific or commercial literature. Literature for reference values for Scr measured by an enzymatic method and obtained according to the IFCC guidelines is quite limited, as shown by Ceriotti et al. [10]. In their paper they referred to only one paper with pediatric data [11] and 5 reports with adult data, of which three of them were on the same group of subjects and had to be considered together. From the remaining three reports on adults, only two were obtained in a European population [12,13], the other in an Australian white population [14]. The results were comparable, giving reference limits of 0.67–1.19 mg/dL for men and 0.51–1.02 mg/dL for women (smallest and largest values of three reference papers are presented here). These data were not able to cover the complete age-range, so clear gaps are observed: newborns (although average values for neonates between 0 and 21 days are given), the adolescent period (15–20 years) and the elderly (above 65–75 years old). For these age groups, there are no recent IDMS-traceable Scr reference values available.

Indirect methods have been described to produce reference intervals [15–17]. These methods use mathematical and statistical procedures, sometimes in combination with exclusion criteria, on retrospective data of patients obtained daily in the laboratory. These methods 'try' to differentiate between pathological and non-pathological results. Although, we do not see these methods as a complete alternative for IFCC recommendations to produce reference intervals, they can be complementary, filling the gaps where healthy-volunteer study results are lacking. On the other hand, healthy-volunteer results may be used to confirm assumptions on which indirect methods are based; as such, these healthy-volunteer results can be used to validate the indirect method. Indirect methods have the advantage that they are cheap, fast and can be repeated at any time, acting as self-validating. When the measurement method evolves, indirect methods can be used to generate new reference intervals, as soon as enough data from the new measurement method is available. Moreover, data mining should not only be seen as a means to obtain reference intervals. Median or means of the data are much more robust than the tailing reference values, especially if the majority of the age/sex subgroup distribution is coming from healthy people. Then, trends may be observed which may help to understand the evolution of Scr with age or sex, in newborns, young adolescents or the elderly.

In this study we present age/sex reference intervals for Scr, obtained from an enzymatic assay, in a Caucasian population, spanning the complete lifetime.

2. Materials and methods

This is a retrospective study using three different indirect methods to analyse the Scr-age/sex dependency over the age-range from birth till approximately 100 years of age. The aim was to obtain information on trends and reference intervals for Scr, obtained by an IDMS-traceable enzymatic assay, complementary to previously published healthy-volunteer study results [10].

2.1. Enzymatic assay

The enzymatic assay for serum creatinine as used from July 2003 in the clinical lab of the AZ Groeninge Hospital, Kortrijk, Belgium, was performed with the Roche kit, Creatinine Plus version 2, cat no 03263991, system ID 0766127 [18]. All tests were performed on the Integra 800 Roche. Within run CV was 1.3% and total CV was 1.8%. The enzymatic method was considered traceable to the IDMS-reference method, with an obtained method comparison line of Scr(IDMS) = 1.010 × Scr(Enz) + 0.0037; R^2 = 0.999; the slope 1.010 with 95%CI [0.992; 1.028] not significantly different from 1, the intercept 0.0037 with 95%CI [−0.158; + 0.168] not significantly different from zero. This linear Scr (IDMS)–Scr(Enz) relation was established on samples of 20 65-year old patients with serum creatinine approximately equally distributed between 1.63 mg/dL and 15.36 mg/dL. The IDMS-traceability was established by applying the procedure described by Stöckl [19] and Thienpont [20] at the department of analytical chemistry, faculty of Pharmaceutical Sciences, University of Ghent.

2.2. Statistical methods

Our approaches are all based on the fact that a hospital population can be regarded as a main population consisting of patients with non-pathological laboratory test results and a small(er) disturbing population with pathological results. These two populations can be separated mathematically if the distribution of the main population is known (parametric method) and if the sample size is sufficiently large, or based on exclusion criteria, or using both techniques.

We used three different statistical approaches to estimate the mean, median, 2.5th and 97.5th percentile for serum creatinine age/sex subgroups:

a) A non-parametric method was applied on the database where all serial measurements were discarded. Discarding serial measurements was considered as an attempt to eliminate possible pathological data. Software used was SAS 9.1.3.
b) The Bhattacharya procedure, described elsewhere [15]. The frequency distribution is divided into a number of classes with the same width (e.g. 0.3 mg/dL). We modified the procedure by applying a five-point Savitzky–Golay smoothing procedure before applying a weighted least square on the linear part of the Bhattacharya plot, as suggested by Oosterhuis et al. [21]. One of the limitations of the Bhattacharya procedure is the rather large number of data needed to obtain relevant statistics, a condition that was met in most of the age/sex subgroups we defined. The Bhattacharya procedure was also used in combination with omitting all serial measurements within the age/sex subgroup. Software used was MS Excel 2003 and the Solver add-in.
c) The cumulative Gaussian distribution was fitted using a non-linear least squares procedure. Here we transformed the serum creatinine values (x) into $z=(x-\mu)/\sigma$ and fitted $y=(\text{top}/2)\times(1+\text{erf}(z/\sqrt{2}))$, where erf is the error function and y is the cumulative fraction. The same classes as for the Bhattacharya procedure were used. We minimized the sum of squared deviations between observed and predicted to obtain the fit parameters. When the distribution was skewed to the right (a phenomenon observed mainly in the elderly), we truncated the sum of squares to a serum creatinine value of 1.0 (for females) and 1.2 (for males), values which are about 0.3 mg/dL above the mean of the non-pathological distribution. This avoided fitting the (right) tail of the distribution. Software used was MS Excel 2003 and Excel's Solver for non-linear curve fitting, minimizing the sum of squares by changing the parameters μ, σ and top with constraint top ≤ 1.

2.3. Preliminary assumptions

2.3.1. The majority of the Scr results are from healthy individuals

Scr serves as a surrogate marker of kidney function, and its measurement has been accepted for the examination of kidney function in health and disease. The majority of Scr-measurements are routinely ordered as part of a basic investigation panel in clinical biochemistry [22]. Consequently, most Scr concentrations are to be expected within the so-called normal range. For adults (20–65 years), the reference values from literature [10] can be used to check the percentage of abnormal values in the database. The total number of male adults in the pooled databases (including serial measurements) was 49,493 of which 34,429 (70%) were between 0.67 and 1.19 mg/dL; for females the total number was 53,906 of which 44,624 (83%) were between 0.51 and 1.02 mg/dL. The total percentage of adults between these reference values was 76%, rising to 92% when serial measurements are omitted. The total percentage of children between the reference

Table 1
Serum creatinine (mg/dL[a]) for neonates as a function of days after birth

Group	Days	Mean	Median	Pct2.5	Pct97.5	Number	Signif
A (cord sera)	0	0.75	0.75	0.56	1.05	79	D–L
B	1	0.81	0.85	0.59	0.97	29	D–L
C	2	0.72	0.75	0.53	0.93	12	E–L
D	3	0.60	0.61	0.45	0.78	14	I–L
E	4	0.55	0.53	0.41	0.75	12	I–L
F	5	0.49	0.50	0.32	0.62	18	J–L
G	6	0.48	0.46	0.36	0.69	16	J–L
H	7	0.45	0.40	0.33	0.62	7	L
I	>7 and <=14	0.40	0.39	0.26	0.61	37	K–L
J	>14 and <=21	0.33	0.32	0.23	0.47	25	
K	>21 and <=28	0.31	0.30	0.22	0.45	33	
L	>28 and <=36	0.29	0.29	0.18	0.38	26	

Signif. means statistically significant differences between group means, corrected for multiple testing.
[a] To express serum creatinine in mol/L, multiply the values by 88.4.

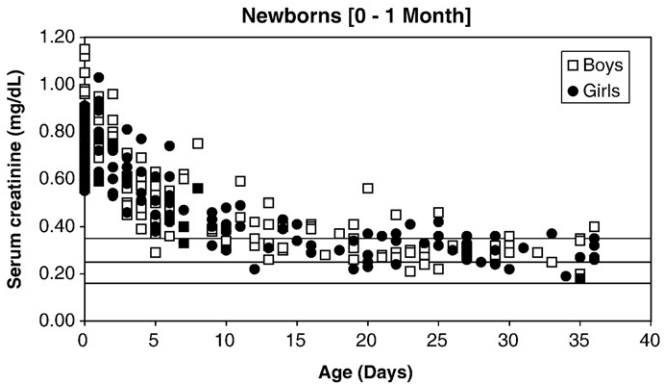

Fig. 1. Serum creatinine (mg/dL) as a function of days after birth for neonates. The horizontal lines are lower, mean and upper values for 1-year old children (see Table 2).

values of Ceriotti et al. [10] was 88%, rising to 90% after removal of serial measurements. These high percentages of values between the normal reference limits published by Ceriotti et al. [10], allow us validating the assumption that the majority of chronologically obtained Scr values are coming from healthy individuals.

2.3.2. Serial measurements were disregarded from the database

We based this decision on two arguments:

1) Serial measurements are strongly correlated, especially if they were taken shortly after each other in time,
2) Closely spaced serial measurements are mostly taken for patients who are suspected of having an underlying pathology related to kidney malfunction.

By disregarding serial data our statistical analysis focused on a cross-sectional setting, also avoiding the necessity to properly account for serial correlation between multiple observations per patient. In fact, we did not even retain one value for patients with serial measurements, because of the higher probability of abnormal data. We know that people with regular health checks may have 2 or more measurements as well, not necessarily with abnormal Scr values, so we inevitably took out data from healthy people. As we did not have any information on the health status of the subjects, we took this decision to reduce the risk of including abnormal data. The database was seriously reduced by applying this criterion, with only 1/6 of the data remaining.

2.4. Age/sex subgroups

Age and sex subgroups were defined as follows: newborns up to 1 month, older than 1 month to 1 year, [1–2y[,...,]19–20y[,]20–25y[25–30y[,..., [95–100y[for male and female separately. The age groups for children were restricted to 1 year because serum creatinine is rapidly changing during childhood. In the age group of the newborns, the exact birth date was known, so a detailed analysis per day was possible. The population studied was mainly Caucasian with a minority of North-African immigrants (comparable muscle mass).

2.5. Non-parametric or parametric approach

In the non-parametric approach, median, lower and upper reference limits (LRL and URL) (as 2.5th and 97.5th percentiles), were calculated after taking out all serial measurements and extreme outliers (which were removed after visual inspection assisted by a statistical outlier test (Grubb's test [23])) in each age/sex subgroup. Normality of the data was checked via visual inspection of the histogram and the location of mean and median. Due to rounding Scr values to two digits, tied (equal) values were responsible for rejecting normality by traditional normality tests (Kolmogorov–Smirnov, Darling–Anderson). However, visual checks confirmed the bell-shaped Gaussian distribution combined with a small subpopulation of larger values. This subpopulation was extremely small in children and young adults, allowing the use of both parametric and non-parametric methods. The right tail was increasing in older adults and becoming quite pronounced in the elderly, resulting in deviating results for the parametric methods (Bhattacharya and cumulative Gaussian fit) as compared to the non-parametric method, especially for the upper limit.

It is known that the Bhattacharya procedure gives the best results when applied to normal distributions that are contaminated with abnormal data (up to 40%) at one side [21]. We found that a small subpopulation in the elderly had very low Scr values. When omitting serial measurements, this subpopulation disappeared from the data, resulting in higher estimates for the lower limit and better correspondence between lower limits determined by the three methods.

2.6. Method validation

We included another important data mining step, by validating our findings on databases from different time-periods. We used

a) The data from July 2006 till December 2007: DB1
b) The data from July 2005 till December 2006: DB2
c) The data from July 2003 till December 2004: DB3.

DB1 and DB2 partially overlap (data from 1/2 year), but this was done to maintain the same time-coverage. DB3 is completely independent from DB1 and DB2. The Hospital changed its storage method in 2005, so it was impossible to obtain data in the same format for the time-period in-between Jan 2005 and July 2005. The data in DB3 had a different format, but the same information was obtained for all three databases.

The pooled database contained 317,045 entries, covering mid-2003 till end 2004 and mid-2005 till end 2007. After removal of all serial measurements (including the first measurement), we obtained a database of 51,304 observations, on which we performed the Bhattacharya and cumulative Gaussian fit procedure. The non-parametric method was applied after we additionally removed 862 outlying values (that is, less than 1.7%) to obtain a more reliable non-parametric estimate of the 97.5th percentile. When the first measurement of a series was maintained in the database, the total number of observations was 87,138.

3. Results

3.1. Method validation

Median/mean, lower and upper limits were calculated for each database (DB1, DB2, DB3), based on the same parametric and non-

Table 2
Median or mean, lower and upper reference limits for serum creatinine (mg/dL[a]) for children of 1-year age-periods, calculated from three statistical methods: non-parametric method (median, 2.5th and 97.5th percentile), Bhattacharya procedure and cumulative Gaussian non-linear fitting (mean, 2.5th and 97.5th probability limits)

Children		Median/mean			Lower limit			Upper limit		
Age	*n*	Non-par	Bhatt	Cumul	Non-par	Bhatt	Cumul	Non-par	Bhatt	Cumul
Newborns	428	0.54	–	–	0.23	–	–	0.98	–	–
1 month to <1 year	1637	0.25	0.25	0.25	0.17	0.15	0.16	0.36	0.36	0.34
1 to <2 years	1061	0.27	0.27	0.26	0.19	0.18	0.18	0.39	0.37	0.35
2 to <3 years	743	0.30	0.30	0.29	0.21	0.19	0.18	0.42	0.42	0.40
3 to <4 years	629	0.33	0.33	0.32	0.23	0.23	0.20	0.46	0.44	0.43
4 to <5 years	419	0.36	0.36	0.35	0.26	0.23	0.23	0.50	0.49	0.47
5 to <6 years	329	0.38	0.38	0.37	0.27	0.25	0.24	0.53	0.51	0.50
6 to <7 years	279	0.43	0.42	0.42	0.29	0.30	0.28	0.58	0.54	0.55
7 to <8 years	252	0.45	0.44	0.44	0.33	0.32	0.29	0.60	0.56	0.58
8 to <9 years	216	0.47	0.45	0.46	0.34	0.28	0.31	0.62	0.63	0.61
9 to <10 years	220	0.50	0.49	0.48	0.36	0.34	0.31	0.69	0.65	0.65
10 to <11 years	239	0.52	0.52	0.51	0.37	0.37	0.35	0.71	0.66	0.67
11 to <12 years	205	0.54	0.53	0.52	0.41	0.39	0.35	0.71	0.68	0.70
12 to <13 years	201	0.57	0.57	0.56	0.43	0.38	0.38	0.74	0.76	0.75
13 to <14 years	201	0.61	0.61	0.60	0.44	0.42	0.39	0.83	0.80	0.81

The sample size *n* is after removal of serial measurements.
[a] To express serum creatinine in µmol/L, multiply the values by 88.4.

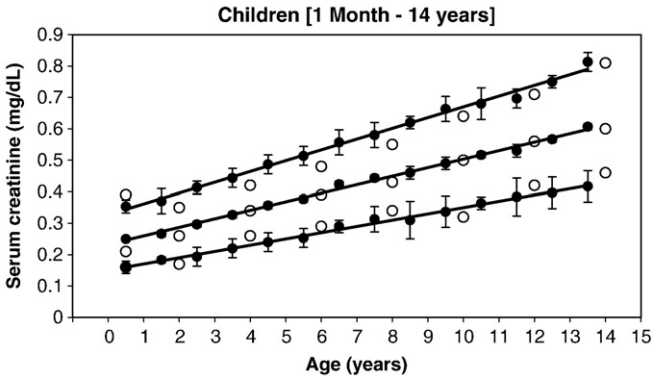

Fig. 2. Linear dependency of Scr (mg/dL) with age for children up to 14 years. The open circles are the reported reference values by Ceriotti et al. [10] and median values from Schlebusch et al. [11]. The error bars are the range of results obtained from the three statistical methods.

parametric procedures. Repeated analyses on separate databases covering different time-periods resulted in comparable results: the slopes of the time comparisons for median, pct2.5 deviated less than 2% from '1' (intercepts were equal to zero) and for pct97.5 the deviation was less than 5% when the elderly (≥85 years) were not taken into account (too small sample size for fair comparison). We therefore decided to pool the data of all three databases and performed the analysis with the three methods on the pooled database.

3.2. Age/sex dependency of serum creatinine

In the next sections we present our findings for newborns, children, adults and the elderly and compare them (where possible) with previously reported literature values (when available for the enzymatic assay) or with previously reported trends (dependency of age and sex).

Table 3
Median or mean, lower and upper reference limits for serum creatinine (mg/dL[a]) for adult males of 1-year age-periods (between 15 and 20 years) and 5-year age-periods (above 20 years) (mid age value is given) calculated according to three different statistical methods: non-parametric method, Bhattacharya procedure and non-linear Gaussian fit of the cumulative distribution

Males		Median/mean			Lower limit			Upper limit		
Age	*n*	Non-par	Bhatt	Cumul	Non-par	Bhatt	Cumul	Non-par	Bhatt	Cumul
14.5	104	0.68	0.68	0.67	0.50	0.47	0.46	0.91	0.89	0.87
15.5	94	0.78	0.78	0.77	0.53	0.51	0.51	1.01	1.05	1.02
16.5	149	0.82	0.82	0.80	0.58	0.54	0.54	1.07	1.09	1.06
17.5	166	0.85	0.85	0.84	0.64	0.62	0.61	1.10	1.09	1.06
18.5	184	0.86	0.86	0.85	0.63	0.62	0.62	1.14	1.11	1.09
19.5	216	0.88	0.88	0.87	0.66	0.64	0.64	1.11	1.13	1.10
22.5	1039	0.90	0.90	0.88	0.65	0.64	0.63	1.16	1.15	1.13
27.5	1099	0.90	0.90	0.89	0.67	0.66	0.65	1.19	1.15	1.13
32.5	1308	0.91	0.90	0.88	0.66	0.64	0.63	1.21	1.15	1.13
37.5	1470	0.91	0.91	0.89	0.63	0.64	0.63	1.23	1.17	1.15
42.5	1776	0.90	0.90	0.88	0.62	0.63	0.61	1.19	1.16	1.15
47.5	1694	0.90	0.89	0.87	0.61	0.60	0.58	1.21	1.18	1.16
52.5	1643	0.91	0.90	0.89	0.61	0.59	0.58	1.24	1.20	1.19
57.5	1581	0.90	0.89	0.88	0.63	0.58	0.58	1.25	1.21	1.18
62.5	1355	0.92	0.94	0.93	0.63	0.64	0.62	1.31	1.25	1.24
67.5	1347	0.94	0.92	0.92	0.63	0.60	0.58	1.42[b]	1.24	1.25
72.5	1372	0.97	0.94	0.94	0.63	0.60	0.59	1.55[b]	1.28	1.30
77.5	1070	1.00	1.00	0.97	0.65	0.59	0.61	1.68[b]	1.41	1.34
82.5	721	1.01	0.96	0.98	0.64	0.61	0.58	1.76[b]	1.31	1.37
87.5	262	1.11	1.03	1.02	0.67	0.62	0.62	1.94[b]	1.43	1.43
92.5	72	1.12	0.99	1.04	0.71	0.73	0.65	2.12[b]	1.25	1.44
97.5	17	1.32	–	–	0.91	–	–	–	–	–

n is the number of observations in each age/sex subgroup.
[a] To express serum creatinine in μmol/L, multiply the values by 88.4.
[b] Non-parametric 97.5th percentile is not used to calculate the mean upper limit in Fig. 3.

Table 4
Median or mean, lower and upper reference limits for serum creatinine (mg/dL[a]) for adult females of 1-year age-periods (between 15 and 20 years) and 5-year age-periods (above 20 years) (mid age value is given) calculated according to three different statistical methods: non-parametric method, Bhattacharya procedure and non-linear Gaussian fit of the cumulative distribution

Females		Median/mean			Lower limit			Upper limit		
Age	*n*	Non-par	Bhatt	Cumul	Non-par	Bhatt	Cumul	Non-par	Bhatt	Cumul
14.5	111	0.62	0.63	0.61	0.46	0.42	0.42	0.78	0.84	0.81
15.5	180	0.68	0.67	0.66	0.52	0.48	0.47	0.92	0.86	0.85
16.5	182	0.70	0.70	0.69	0.50	0.49	0.48	0.95	0.91	0.90
17.5	198	0.71	0.71	0.70	0.53	0.50	0.50	0.94	0.91	0.90
18.5	220	0.71	0.71	0.69	0.45	0.46	0.46	0.93	0.95	0.93
19.5	255	0.70	0.70	0.69	0.47	0.46	0.46	0.93	0.94	0.92
22.5	1474	0.68	0.68	0.67	0.45	0.45	0.44	0.94	0.91	0.90
27.5	2013	0.66	0.65	0.65	0.45	0.41	0.40	0.91	0.90	0.89
32.5	1942	0.67	0.67	0.66	0.46	0.43	0.42	0.92	0.91	0.89
37.5	1949	0.69	0.69	0.67	0.49	0.47	0.46	0.95	0.90	0.89
42.5	2048	0.70	0.70	0.68	0.49	0.49	0.47	0.96	0.91	0.90
47.5	2101	0.71	0.70	0.69	0.49	0.48	0.47	0.98	0.93	0.91
52.5	1868	0.72	0.71	0.70	0.50	0.48	0.47	0.98	0.94	0.92
57.5	1709	0.72	0.71	0.70	0.50	0.49	0.47	1.01[b]	0.93	0.92
62.5	1480	0.73	0.72	0.71	0.50	0.48	0.48	1.03[b]	0.96	0.94
67.5	1507	0.74	0.73	0.72	0.49	0.46	0.46	1.10[b]	1.00	0.98
72.5	1621	0.76	0.74	0.72	0.51	0.47	0.48	1.16[b]	1.01	0.97
77.5	1542	0.79	0.76	0.74	0.49	0.43	0.45	1.34[b]	1.10	1.04
82.5	1259	0.81	0.78	0.78	0.51	0.42	0.43	1.48[b]	1.14	1.14
87.5	566	0.84	0.80	0.80	0.47	0.44	0.43	1.67[b]	1.16	1.16
92.5	311	0.88	0.81	0.83	0.53	0.41	0.43	1.61[b]	1.21	1.24
97.5	66	0.96	0.87	0.86	0.51	0.47	0.43	1.86[b]	1.27	1.30

n is the number of observations in each age/sex subgroup.
[a] To express serum creatinine in μmol/L, multiply the values by 88.4.
[b] Non-parametric 97.5th percentile is not used to calculate the mean upper limit in Fig. 4.

3.2.1. Newborns

Our database contained information of the birthday for all observations from mid-2006 on. This allowed us to make a detailed day-to-day analysis of newborns, of which we had 428 observations in our pooled database (after removal of serial values), but only 308 with exact birthday information. We were not able to make a difference between term and preterm neonates as our database did not contain this information.

Table 1 presents Scr (mg/dL) summary statistics for neonates in subgroups defined by the number of days after birth. Fig. 1 shows the

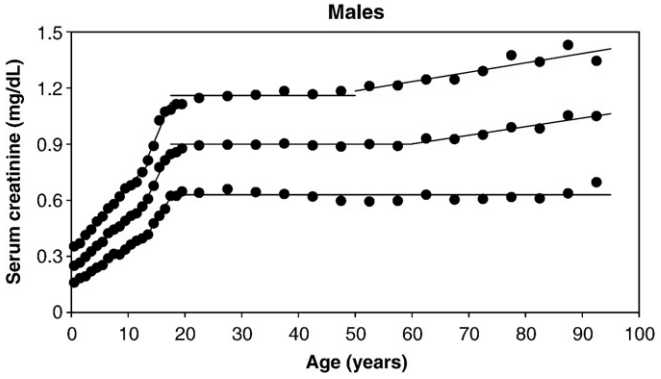

Fig. 3. Age dependency of Scr (mg/dL) for males for lower limit, mean and upper limit.

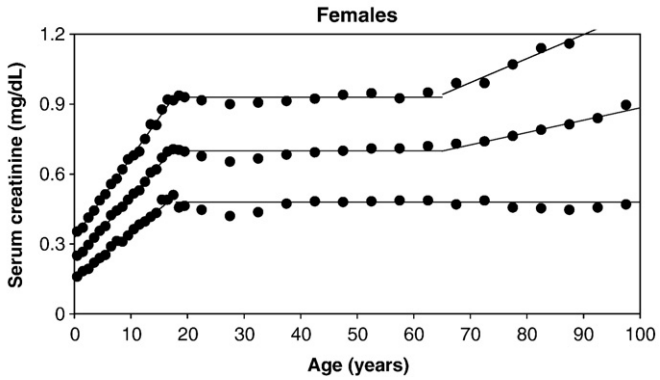

Fig. 4. Age dependency of Scr (mg/dL) for females for lower limit, mean and upper limit.

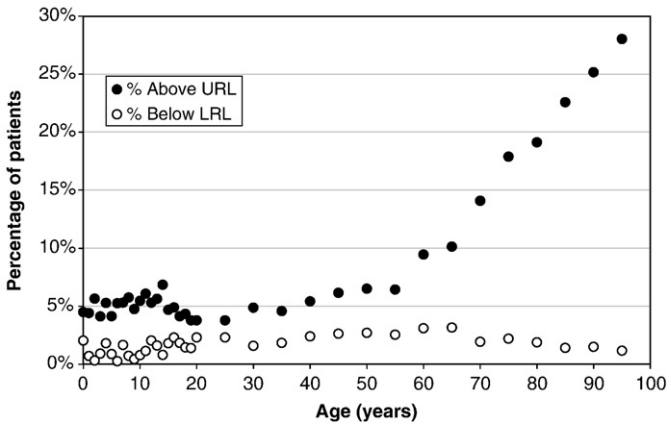

Fig. 6. the percentage of patients below the lower reference limit and above the upper reference limit for Scr as a function of age in the pooled database.

graph with the individual data for boys and girls as a function of days after birth. Comparison between the groups defined in Table 1 using one-way ANOVA, followed by Tukey's multiple comparison test, showed many significant differences, indicating that Scr was significantly decreasing in the first 3 weeks after birth.

3.2.2. Children

Bhattacharya procedure and cumulative Gaussian distribution non-linear fitting on all data (excluding serial measurements) resulted in the same results as the non-parametric method (deviations were mostly smaller than 0.02 mg/dL) (see Table 2). The 97.5th percentile showed the largest deviations between methods, mainly because omitting serial measurements did not remove all large values, resulting in slightly higher values for the non-parametric method. As the differences are small, we used the mean of the three methods and the range of the results as error bars for the graph in Fig. 2. As the distribution for newborns (0–1 month) was clearly not Gaussian and could not be considered as a combination of a Gaussian distribution and a small disturbing population, we used the non-parametric method only to give an overall estimate of median, lower and upper limit for that subgroup, a value which can be compared to the values reported by Ceriotti et al. (0.31–0.98 mg/dL).

3.2.3. Adults

The same analysis was performed on age/sex subgroups of male and female adolescents, adults and elderly. The results are presented in Tables 3 and 4 and Figs. 3 and 4.

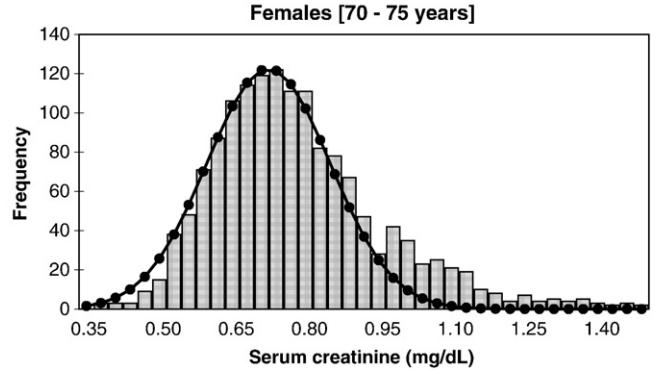

Fig. 5. Frequency distribution for Scr (mg/dL) of 70– 75-year old females. The fitted result from the cumulative Gaussian non-linear method is shown as the bold solid line.

3.2.4. Elderly

The results for the elderly are part of Tables 3 and 4 and Figs. 3 and 4.

The correspondence between the results obtained from the three methods is worse in the elderly subgroups. The reported parametric results are based on the main (normal) population within the elderly subgroups; the non-parametric result is strongly influenced by the right-skewness of the distribution. For ages above 90 years, the sample size for males became too small to obtain reliable results.

An example of the right-skewness of the frequency distribution for an older population is presented in Fig. 5. The change in distribution resulted in an increased number of subjects having Scr values outside the reference intervals (see Fig. 6).

4. Discussion

The variation of Scr with age and sex makes the task of defining scientifically sound reference intervals very demanding. Healthy-volunteer studies are mostly based on small studies, making it difficult to obtain reliable estimates for reference limits when these limits vary with age and sex. Indirect methods may be helpful and complementary to these healthy-volunteer studies.

Reference intervals, no matter if they are determined by a direct or indirect method, are based on the tails of the distribution. It is an arbitrary but common convention to use the 2.5th and 97.5th percentile (parametric or non-parametric) as reference intervals, as they include 95% of the study subjects. Clearly, these values are susceptible to errors, depending on the sample size and outlying observations. For instance, in a healthy-volunteer study of 100 subjects (which is already a large sample size for these kind of studies), the second smallest and second largest values define the non-parametric reference interval. Outliers may have a tremendous effect on these values and should be omitted, no matter if direct or indirect methods are used. Omitting outliers is clearly a decision based on the data, not on the health status of the individual.

Another important remark is that a sample size of 100 subjects covering a broad age-range does not allow the accurate determination of age-specific reference intervals if the biochemical compound varies with age. Sometimes it is even impossible to discover age tendencies in such a small group of subjects.

4.1. Newborns

The analysis shows that at birth, the median Scr value is 0.75 mg/dL, with a LRL=0.56 mg/dL and an URL=1.05 mg/dL, reflecting the maternal creatinine level (see the results for adult females). These

values of lower and upper limit correspond very well with the values for cord sera reported by Ceriotti et al. [10] (0.52 mg/dL and 0.97 mg/dL). The values reported for the lower and upper limit by Ceriotti et al. for preterm neonates (n = 58) (0.32–0.98 mg/dL) and term neonates (n = 69) (0.31–0.92 mg/dL) are averaged over 21 and 14 days after birth resp. Our analysis shows the fast decay of Scr during the first days of life, indicating that Scr values are of less clinical importance in the first 3–4 weeks after birth. From Fig. 1, it can be seen that there is no difference between boys and girls and that after 1 month, nearly all Scr values fall between the lower and upper limit of 1-year old children (the horizontal lines in Fig. 1).

The finding that Scr is increased in newborns, especially in their early days of life, has been reported previously [24–26]. It has been reported that the rise in creatinine is mostly observed in the first 48 h and drops down thereafter [24]. Although it was mentioned that, after the third day, the mean Scr value of the cohort in Manzar et al. [26] dropped from a level of 0.64 mg/dL to 0.44 mg/dL, the median values we found in the newborns are much higher at birth (0.75 mg/dL) and are going down even further, to median values of 0.25 mg/dL (the median value of the 1-year old children). However, it was unclear from the paper of Manzar et al. which measurement method was used to determine Scr. The trend in Fig. 1 suggests that even after the third day of life, Scr values drop further. Miall et al. [24] reported that Scr dramatically rises in the first 48 h of life in preterm infants. We also found an increased and statistically significant difference in mean values between day 0 and day 1 (2-sided t-test, p = 0.031, not corrected for multiple testing), which confirms the findings of Miall et al.

4.2. Children

Table 2 and Fig. 2 summarize the results for children between 0 and 14 years. Serum creatinine increases with age during childhood, equal for boys and girls, up to the age of 14 years. As there is a logical linearly increasing trend with age, the Scr-age dependency could further be modelled by linear regression equations for median/mean, lower and upper limit. This also corrects for small irregularities in the pattern when calculations are based on smaller age/sex subgroups.

$$\begin{aligned}\mathrm{Scr(LRL)} &= 0.0199 \times \mathrm{Age} + 0.1504 \left(R^2 = 0.994\right)\\ \mathrm{Scr(Median)} &= 0.0270 \times \mathrm{Age} + 0.2329 \left(R^2 = 0.996\right)\\ \mathrm{Scr(URL)} &= 0.0343 \times \mathrm{Age} + 0.3272 \left(R^2 = 0.994\right)\end{aligned}.$$

Interesting also is to compare our results with the only published reference limits for children obtained from healthy volunteers and for IDMS-traceable Scr-measurements. The reference limits from Ceriotti et al. are plotted on Fig. 2, for sake of comparison. There is nearly perfect agreement for the lower limit and median (obtained from Schlebusch et al. [11]), but the upper limit obtained from our analysis is slightly higher than the values reported by Ceriotti et al.

An important remark is that the reference limits from Ceriotti et al. were for 2-year broad age-ranges, except for the 2-month–1 -year age-period. Although the correspondence with the limits reported by Ceriotti et al. for children is very good, there are some differences: a) our upper limit is mostly higher (for all three methods used), although not much; b) we report different limits for 1-year broad age-ranges, indicating that Scr is changing fast and c) the limits we report are showing a smoother pattern as a function of age than the bumpier pattern of the Ceriotti limits. This might indicate that the lower and upper limits are sensitive to variability in the data, probably as a consequence of their small sample size.

4.3. Adults

During childhood, there is no difference between males and females. At the age of 14–15 years, the Scr-age dependency starts to deviate between males and females, with a steeper Scr-increase for boys between 15 and 20 years, than for girls. This corresponds to the average growth curves of boys and girls (Flandres) which start deviating at the age of 14 years: girls reach their final height around 14–15 years, whereas boys continue to grow up to the age of 19–20 years, a difference that explains the steep increase in Scr differences between sexes.

Between 20 and 70 years old, the mean Scr for males is stable and equal to 0.90 mg/dL. After the age of 70, mean serum creatinine slightly starts increasing again. The linear dependency with age for mean is: Scr (mean) = 0.0046 × Age + 0.6254 (R^2 = 0.975). The lower limit is constant over a very broad age-range of 17–95 years and equal to 0.63 mg/dL. The upper limit is constant over the age-range of 17–55 years and equal to 1.16 mg/dL and is slightly increasing from 55 years on following the relation Scr(URL) = 0.0050 × Age + 0.9343 (R^2 = 0.945).

Scr increases for girls just like for boys during childhood, however, the linear dependency with age is valid up to 20 years for girls. Between 20 and 70 years, mean Scr is approximately stable and equal to 0.70 mg/dL. After the age of 70 years, mean Scr starts to increase slightly again, following a linear relationship with age: Scr(mean) = 0.0053 × Age + 0.3542 (R^2 = 0.989). The lower limit remains about 0.48 mg/dL over the complete 20– 100-year age-range. The upper limit is constant between 20 and 70 years and equal to 0.93 mg/dL. After 70 years of age, the distribution becomes right-skewed and the upper limit is slightly going up, according to Scr(URL) = 0.0103 × Age + 0.2713.

Note that the difference between males and females increases up to a value of 0.25 mg/dL at the age of 30 years (mostly because females have a small decrease around 25–30 years, a fact that we found in the three databases DB1, DB2 and DB3), when it starts to decrease and becomes constant at an average value of 0.20 mg/dL.

4.4. Elderly

The reference values we report for Scr should be carefully used in assessing renal function, as it has been mentioned previously that Scr is an inadequate screening test for renal failure in elderly patients [27,28]. Scr-level may estimate GFR imprecisely in the elderly because aging leads to a loss of lean muscle mass, and thus a concurrent decline in creatinine production. Parallel to this process is a real decline in renal function with age, resulting in an increase of Scr. As a result, Scr-levels are frequently within normal limits despite the presence of underlying renal impairment.

Reference values for Scr in elderly healthy individuals are not available for the enzymatic assay. However, trends in the data have been reported [29] for Scr measured by alkaline picrate assays and confirm the trend obtained in our analysis. Mean Scr is increasing and distributions are significantly skewed. The Framingham Heart Study [30] also showed that elevated Scr-levels are common in the community and are strongly associated with older age.

As age in itself is not a disease, the change from a Gaussian to a right-skewed frequency distribution (Fig. 5) might be considered normal and much higher upper limits would then be calculated, based on the non-parametric approach (see Tables 3 and 4, data marked with “b”). The parametric approaches described here are not affected by the right-tail of the distribution. As there is no exact generally accepted definition of pathologic versus degenerative changes in the ageing kidney, the most appropriate criteria for a healthy elderly group are difficult to identify. In the Framingham Heart Study, cutoff points for Scr were 30–40% above mean Scr values, corresponding with our parametric 97.5th percentile.

Therefore, we used the obtained parametric reference intervals to calculate the percentage of subjects outside the lower and upper limits in each age/sex subgroup (see Fig. 6) from the database (including the first observation of a series). The percentage of subjects with Scr lower than the lower limit is roughly unaffected by age and is between 0% and 3% (a value expected, as the lower limit is defined as the 2.5th percentile). Below the age of 60, approximately 5% of the observations is above the upper limit, with a linearly increasing percentage from

approximately 10% to 30% between 60 and 100 years of age (approximately 0.5% increase per year), probably reflecting the increasing prevalence with age of renal impairment.

5. Conclusions

Reference intervals for Scr remain relevant despite the current emphasis on the use of the estimated glomerular filtration rate for assessing renal function. This is due to the limitations of the statistical models to estimate GFR. The determination of reference values is not straightforward when age/sex dependency has to be considered.

Advantages of indirect (data mining) methods applied on large hospital databases over direct methods (study in healthy individuals) are that a) the numbers are much larger, covering the complete age-range; b) much more detailed analysis is possible; c) it is possible to fill the gaps uncovered by healthy-volunteer studies by making use of healthy-volunteer study results to validate the assumptions.

There are also clear disadvantages to using indirect methods. Mathematical methods can be used to differentiate between pathological and non-pathological data, but these methods are based on assumptions on the underlying frequency distribution. These methods may be combined with inclusion/exclusion criteria to better define the population under study, but, again, these criteria are based on assumptions. As assumptions are food for discussion, the application of indirect methods will always be criticized.

Therefore, a combination of direct and indirect methods may result in lifetime spanned reference intervals for both sexes. Combination of both methods allows verifying the correctness of data mining methods with the more limited results of direct methods.

By using this approach we presented:

a) Data for newborns as a function of days after birth, showing a fast decay in Scr during the first month;
b) Data for children between 14 and 20 years which are rarely obtained from healthy individual studies as Scr is changing very rapidly in this age group and consequently large numbers are required for estimating reliable limits;
c) Data for older people (>75 years) which are also rarely published for many reasons, among which the difficulty to declare such a person healthy and to find a reliable sample size in these age groups.

By using literature reference intervals to check our data mining method, we were able to determine Scr reference intervals for the enzymatic assay, spanning the complete lifetime for both men and women in a Caucasian population.

Acknowledgment

The authors would particularly like to thank Dr. S. Vanderschueren for providing laboratory data from the LIS, AZ Groeninge, Kortrijk.

References

[1] Levey AS, Bosch JP, Breyer Lewis J, Greene T, Rogers N, Roth D. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Ann Intern Med 1999;130:461–70.
[2] Levey AS, Coresh J, Greene T, et al. Expressing the modification of diet in renal disease study equation for estimating glomerular filtration rate with standardized serum creatinine values. Clin Chem 2007;53:766–72.
[3] National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis 2002;39: S1–S266.
[4] National Kidney Disease Education Programme. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification and stratification. Part 5. Evaluation of laboratory measurements for clinical assessment of kidney disease. Guideline 4; May 22, 2008. Estimation of GFR. http://www.kidney.org//professionals/kdoqi/guidelines_ckd/p5_lab_g4.htm [accessed.
[5] Joint Specialty Committee for Renal Disease Royal College of Physicians of London and the Renal Association. Chronic Kidney Disease in Adults: UK Guidelines for Identification, Management and Referral. Royal College of Physicians of London; 2005. Available at: http://www.renal.org/CKDguide/full/UKCKDfull.pdf [accessed May 22, 2008].
[6] Panteghini M. Enzymatic assays for creatinine: time for action. Scand J Clin Lab Invest 2008;68:84–8.
[7] Myers GL, Miller WG, Coresh J, et al. Recommendations for improving serum creatinine measurement: a report from the Laboratory Working Group of the National Kidney Disease Education Program. Clin Chem 2006;52:5–18.
[8] Solberg HE, International Federation of Clinical Chemistry (IFCC), Scientific Committee, Clinical Section, Expert Panel on Theory of Reference Values and International Committee of Standardization in Haematology (ICSH), Standing Committee on Reference Values. Approved recommendation (1986) on the theory of reference values. Part 1. The concept of reference values. J Clin Chem Clin Biochem 1987; 25: 337–342.
[9] CLSI document C28-A2. How to define and determine reference intervals in the clinical laboratory. Approved guideline. 2nd edition. Wayne, PA: Clinical and Laboratory Standards Institute; 2000.
[10] Ceriotti F, Boyd JC, Klein G, et al. Reference intervals for serum creatinine concentrations: assessment of available data for global application. Clin Chem 2008;54:559–66.
[11] Schlebusch H, Liappis N, Kalina E, Klein G. High sensitivity CRP and creatinine: reference intervals from infancy to childhood. J Lab Med 2002;26:341–6.
[12] Junge W, Wilke B, Halabi A, Klein G. Determination of reference intervals for serum creatinine, creatinine excretion and creatinine clearance with an enzymatic method and a modified Jaffe method. Clin Chim Acta 2004;344:137–48.
[13] Rustad P, Felding P, Franzson I, et al. The Nordic Reference Interval Project 2000: recommended reference intervals for 25 common biochemical properties. Scand J Clin Lab Invest 2004;64:271–84.
[14] Mazzachi BC, Peake MJ, Ehrhardt V. Reference range and method comparison studies for enzymatic and Jaffe creatinine assays in plasma and serum and early morning urine. Clin Lab 2000;46:53–5.
[15] Bhattacharya CG. A simple method of resolution of a distribution into Gaussian components. Biometrics 1967;23:115–35.
[16] Martin HF, Hologgitas JV, Driscoll J, Fanger H, Gudzinowicz BJ. Reference values based on populations accessible to hospitals. In: Gräsbeck R, Alström T, editors. Reference values in laboratory medicine. Wiley: Chichester; 1981. p. 233–62.
[17] Kairisto V, Hänninen KP, Leino A, et al. Generation of reference values for cardiac enzymes from hospital admission laboratory data. Eur J Clin Chem Clin Biochem 1994;32:789–96.
[18] Guder WG, Hoffmann GE, Hubbuch A, Poppe WA, Siedel J, Price CP. Multicentre evaluation of an enzymatic method for creatinine determination using a sensitive colour reagent. J Clin Chem Clin Biochem 1986;24:889–902.
[19] Stöckl D, Reinauer H. Candidate reference methods for determining target values for cholesterol, creatinine, uric acid, and glucose in external quality assessment and internal accuracy control. I Method setup. Clin Chem 1993;39:993–1000.
[20] Thienpont LM, De Leenheer AP, Stöckl D, Reinauer H. Candidate reference methods for determining target values for cholesterol, creatinine, uric acid, and glucose in external quality assessment and internal accuracy control. II Method transfer. Clin Chem 1993;39:1001–6.
[21] Oosterhuis WP, Modderman TA, Pronk C. Reference values: Bhattacharya or the method proposed by the IFCC? Ann Clin Biochem 1990;27:359–65.
[22] Levey AS, Stevens LA, Hostetter Th. Automatic reporting of estimated glomerular filtration rate — just what the doctor ordered. Clin Chem 2006;52:2188–93.
[23] Grubbs F. Procedures for detecting outlying observations in samples. Technometrics 1969;11:1–2.
[24] Miall LS, Henderson MJ, Turner AJ, et al. Plasma creatinine rises dramatically in the first 48 hours of life in preterm infants. Pediatrics 1999;104:e76.
[25] Gordjani N, Burghard R, Leititis JU, Brandis M. Serum creatinine and creatinine clearance in healthy neonates and prematures during the first 10 days of life. Eur J Pediatr 1988;148:143–5.
[26] Manzar S, Al-Umran K, Al-Awary BH, Al-Faraidy A. Changes in plasma creatinine in first 72 hours of life. Arch Dis Child Fetal Neonatal Ed 2001;85:F146–7.
[27] Swedko PJ, Clark HD, Paramsothy K, Akbari A. Serum creatinine is an inadequate screening test for renal failure in elderly patients. Arch Intern Med 2003;163:356–60.
[28] Perrone RD, Madias NE, Levey AS. Serum creatinine as an index of renal function: new insights into old concepts. Clin Chem 1992;38:1933–53.
[29] Papaioannou A, Ray JG, Ferko NC, Clarke JA, Campbell G, Adachi JD. Estimation of creatinine clearance in elderly persons in long-term care facilities. Am J Med 2001;111:569–73.
[30] Culleton BF, Larson MG, Evans JC, et al. Prevalence and correlates of elevated serum creatinine levels: the Framingham Heart Study. Arc Intern Med 1999;159:1785–90.