Recommendations for Improving Serum Creatinine Measurement: A Report from the Laboratory Working Group of the National Kidney Disease Education Program

GARY L. MYERS,1* W. GREG MILLER,2 JOSEF CORESPH,3 JAMES FLEMING,4 NEIL GREENBERG,5 TOM GREENE,6 THOMAS HOSTETTER,7 ANDREW S. LEVEY,8 MAURO PANTEGHINI,9 MICHAEL WELCH,10 and JOHN H. ECKFELDT11 for the National Kidney Disease Education Program Laboratory Working Group

Background: Reliable serum creatinine measurements in glomerular filtration rate (GFR) estimation are critical to ongoing global public health efforts to increase the diagnosis and treatment of chronic kidney disease (CKD). We present an overview of the commonly used methods for the determination of serum creatinine, method limitations, and method performance in conjunction with the development of analytical performance criteria. Available resources for standardization of serum creatinine measurement are discussed, and recommendations for measurement improvement are given.

Methods: The National Kidney Disease Education Program (NKDEP) Laboratory Working Group reviewed problems related to serum creatinine measurement for estimating GFR and prepared recommendations to standardize and improve creatinine measurement.

Results: The NKDEP Laboratory Working Group, in collaboration with international professional organizations, has developed a plan that enables standardization and improved accuracy (trueness) of serum creatinine measurements in clinical laboratories worldwide that includes the use of the estimating equation for GFR based on serum creatinine concentration that was developed from the Modification of Diet in Renal Disease (MDRD) study.

Conclusions: The current variability in serum creatinine measurements renders all estimating equations for GFR, including the MDRD Study equation, less accurate in the normal and slightly increased range of serum creatinine concentrations [<133 μmol/L (1.5 mg/dL)], which is the relevant range for detecting CKD [<60 mL·min⁻¹·(1.73 m²)⁻¹]. Many automated routine methods for serum creatinine measurement meet or exceed the required precision; therefore, reduction of analytical bias in creatinine assays is needed. Standardization of calibration does not correct for analytical interferences (nonspecificity bias). The bias and nonspecificity problems associated with some of the routine methods must be addressed.

Chronic kidney disease (CKD)12 is a major public health problem in the United States. The incidence and prevalence of end-stage renal disease, kidney failure treated by

12 Nonstandard abbreviations: CKD, chronic kidney disease; GFR, glomerular filtration rate; K/DOQI, Kidney Disease Outcomes Quality Initiative; NKDEP, National Kidney Disease Education Program; MDRD, Modification of Diet and Renal Disease; IDMS, isotope dilution mass spectrometry; GC, gas chromatography; LC, liquid chromatography; PT, proficiency testing; EQAS, external quality assurance scheme; CAP, College of American Pathologists; IVD, in vitro diagnostics; ISO, International Organization for Standardization; JCTLM, Joint Committee on Traceability in Laboratory Medicine; and CLSI, Clinical and Laboratory Standards Institute.
dialysis, and transplantation have more than quadrupled over the last 2 decades (1). The estimated number of people with earlier stages of CKD is ~19 million, including ~8 million people with a reduced glomerular filtration rate (GFR) < 60 mL min⁻¹ (1.73 m²)⁻¹ and another ~11 million with a GFR > 60 mL min⁻¹ (1.73 m²)⁻¹ but an abnormally high albumin excretion (urine albumin-to-creatinine ratio > 30 mg/g) (2).

The National Kidney Foundation Kidney Disease Outcomes Quality Initiative (K/DOQI) and the National Kidney Disease Education Program (NKDEP) within the National Institute of Diabetes and Digestive and Kidney Diseases recently defined CKD as either kidney damage or a GFR

Diseases recently defined CKD as either kidney damage or a GFR

comes Quality Initiative (K/DOQI) and the National

Kidney Disease Education Program (NKDEP) within the

National Institute of Diabetes and Digestive and Kidney

Diseases recently defined CKD as either kidney damage

or a GFR

Diseases recently defined CKD as either kidney damage

National Institute of Diabetes and Digestive and Kidney

Parkinson’s disease (3). GFR is traditionally considered the best overall index of kidney function (4). The threshold of GFR < 60 mL min⁻¹ (1.73 m²)⁻¹ was selected as the definition of CKD because at this value approximately one half of an adult’s normal kidney function is lost, leading to several possible complications (3).

Understanding by laboratorians worldwide of the importance of reliable serum creatinine measurements in GFR estimation and of factors that may affect creatinine measurement is critical to ongoing global public health efforts to increase the diagnosis and treatment of patients with CKD. The NKDEP Laboratory Working Group, in collaboration with international professional organizations, has developed a plan that enables standardization and improved accuracy (trueeness) of serum creatinine measurements in clinical laboratories worldwide.

Materials and Methods for Estimating GFR

GFR cannot be measured by direct means, but it can be assessed by measuring the urinary clearance of endogenous filtration markers such as inulin, iohexol, or iothalamate (5–7). However, because of difficulty in use, expense, radiation exposure, and radionuclide regulatory requirements, these methods have limited use in the routine laboratory and are typically confined to the research setting.

GFR is often estimated clinically from serum concentrations of endogenous creatinine (8) or cystatin C (9, 10). Serum cystatin C has not yet been adequately evaluated as an index of GFR (11), however, and serum creatinine alone should not be used to assess the GFR or to detect the presence of CKD because it is affected by the GFR and by factors independent of GFR, including age, sex, race, body size, diet, certain drugs, and laboratory analytical methods (12, 13). More accurate and precise estimations of GFR can be obtained with equations that empirically combine all of the average effects from factors that affect serum creatinine other than GFR (14). The currently recommended estimating equation was developed from the Modification of Diet in Renal Disease (MDRD) Study (15) and is based on GFR values measured by iothalamate clearance in 1628 adults and subsequently validated in another 1775 adults in the African American Study of Kidney Disease (AASK) (16). The “four-variable” MDRD Study equation (Eq. 1) uses age, sex, race (African-American vs non–African-American), and serum creatinine (sCr) (17):

\[
\text{GFR} \text{[mL} \cdot \text{min}^{-1} \cdot (1.73 \text{m}^2)^{-1}] = 186 \times (\text{sCr})^{-1.154} \\
\times \text{(age)}^{-0.203} \times (0.742 \text{if female}) \times \text{(1.210 if African-American)}
\]

The MDRD Study equation does not require a body weight variable because it normalizes GFR for a standard body surface area of 1.73 m². The MDRD Study equation has been demonstrated to be useful for CKD patients and performs similarly in diabetic vs nondiabetic individuals (18), but its use is unclear for healthy individuals and is not recommended for hospitalized patients (19).

Because of the current variability in calibration of serum creatinine assays, assays not calibrated in agreement with the kinetic alkaline picrate assay used in the MDRD Study introduce a source of error into GFR estimates. This calibration error is relatively greater and contributes to larger uncertainty in GFR estimates at lower creatinine values near the upper limit of the reference interval (20). The progressively larger effect on estimated GFR of different calibration biases of creatinine methods is shown in Fig. 1 (21), and the progressively larger effect of measurement imprecision at lower creatinine values is shown in Fig. 2. Thus, calibration bias and measurement imprecision for serum creatinine have a much larger impact on the uncertainty in estimated GFR when serum creatinine is close to the reference value,
which is the relevant range for detecting early CKD [GFR <60 mL·min⁻¹·(1.73 m²)⁻¹]. This limitation applies to all estimating equations based on serum creatinine, not just the MDRD Study equation (22). For this reason, the NKDEP currently recommends that GFR estimates above 60 mL·min⁻¹·(1.73 m²)⁻¹ be reported simply as “>60 mL·min⁻¹·(1.73 m²)⁻¹” rather than as a discrete numeric value (3, 18). Variability in creatinine calibration and measurement imprecision also contributes to substantial uncertainty in estimating GFR in children, who usually have lower serum creatinine concentrations than do adults. For estimating GFR in children, the Schwartz and the Counahan–Barratt equations are recommended (23–26). Both provide GFR estimates based on a constant multiplied by the child’s height divided by the measured serum creatinine concentration.

**Sources of Variability in Estimating GFR**

Sources of variability in GFR estimates include-underlying biological variability in GFR, biological variability in serum creatinine, and errors in the measurement of serum creatinine and in the estimating equation.

GFR may vary in response to meals, exercise, posture, changes in blood pressure, and other conditions. GFR is also affected by pregnancy, glucose control in diabetes, extracellular fluid volume, antihypertensive medications, and acute and chronic kidney disease (27). Error may occur in measurement of serum and urine filtration markers or of urine flow rate, or in techniques for urine collection. Variability among clearance periods during GFR measurement may also lead to error (28). Median intra-individual CVs reported for measured GFR ranged from 6.3% to 7.5% (6, 29). These GFR measurements were made in controlled studies; consequently, the intra-individual variability was likely lower than would be observed in a typical clinical setting.

In 2 published studies, the mean intraindividual CVs for serum creatinine were 4.1% and 4.3%, respectively (30, 31). Neither of these studies included patients with CKD.

**Analytical Performance Specifications for GFR Estimates**

Percentile distribution of the differences between estimated and measured GFR is a useful measure to assess the accuracy of GFR estimates. The K/DOQI reported that for an independent sample of 1070 participants evaluated in the GFR range <90 mL·min⁻¹·(1.73 m²)⁻¹, ~90% of GFR estimates calculated by use of the MDRD Study equation were within 30% of the measured GFR (3). This overall measure of clinical performance included error components from several sources: measurement of serum creatinine, including specimen nonspecificity effects and the effects on determinants of serum creatinine other than GFR, including generation, secretion, and elimination; and from measurement of GFR as iothalamate clearance, including physiologic differences in renal function and various comorbid conditions.

Considering the various types of error, an estimated GFR within 30% of a measured GFR was considered acceptable by K/DOQI for clinical interpretation to identify individuals with CKD as defined by GFR <60 mL·min⁻¹·(1.73 m²)⁻¹ for 3 months or more and to follow subsequent progression of the disease. For example, at a GFR of 60 mL·min⁻¹·(1.73 m²)⁻¹, the range of GFR values would be 42–78 mL·min⁻¹·(1.73 m²)⁻¹.

**Analytical Performance Specifications for Serum Creatinine Measurement**

Serum creatinine measurements must have a small enough total error that the impact on the total uncertainty of estimated GFR remains within clinically acceptable limits. The critical serum creatinine concentration corresponding to a GFR of 60 mL·min⁻¹·(1.73 m²)⁻¹ varies with the age, sex, and race of the patient (3). Typical values for serum creatinine at this critical GFR are 88.4 μmol/L (1.00 mg/dL) for a 60-year-old non–African-American female, 99 μmol/L (1.18 mg/dL) for a 60-year-old African-American female, 114 μmol/L (1.30 mg/dL) for a 60-year-old non–African-American male, and 135 μmol/L (1.53 mg/dL) for a 60-year-old African-American male. Thus, creatinine values within or very close to many published reference intervals are consistent with substantial reduction in GFR in some patients. For the same demographic groups at an estimated GFR of 30 mL·min⁻¹·(1.73 m²)⁻¹, the serum creatinine values are 162, 190, 209, and 247 μmol/L (1.83, 2.15, 2.37, and 2.79 mg/dL), respectively. Because of the dramatic increase in the impact of creatinine bias and imprecision on the error of an estimated GFR as the serum creatinine value gets smaller (Figs. 1 and 2), the laboratory measurement performance goal is currently targeted at a creatinine

![Fig. 2. Effect of creatinine measurement imprecision on estimated GFR.](Image)

Solid lines represent the upper and lower limits of the 95% confidence interval for estimated GFR for a 60-year-old non–African-American female for whom the estimated GFR is 60 mL·min⁻¹·(1.73 m²)⁻¹ at a creatinine of 88.4 μmol/L (1.00 mg/dL), using a value of 5.3 μmol/L (0.06 mg/dL) as the measurement SD. This SD was the median SD observed for 50 different method groups assaying a fresh-frozen serum specimen with a creatinine value of 80 μmol/L (0.90 mg/dL) in the 2003 CAP survey (126). The dashed lines represent the upper and lower limits of the 95% confidence interval for estimated GFR based on the largest peer-group SD, 12 μmol/L (0.13 mg/dL), observed in the survey.
A concentration of 88.4 μmol/L (1.00 mg/dL), which is consistent with a GFR of 60 mL·min⁻¹·(1.73 m²)⁻¹ for some adult demographic groups and is at the lower range of measurement, where the impact of bias and imprecision will be greater than at higher values.

The 2 primary components of measurement error in serum creatinine are systematic bias, a consistent error typically resulting from calibration differences between measurement procedures, and random measurement error, including within-laboratory effects, between-laboratory random variability in day-to-day calibration, and specimen-specific effects. In a simulation study, gaussian-distributed random errors and increasing systematic biases were added to the baseline serum creatinine measurements of 491 patients in the MDRD Study validation subset who had serum creatinine measurements between 88.4 and 132 μmol/L (1.00 and 1.50 mg/dL). The increase in root mean square error in the estimated GFR, compared with an iothalamate-measured GFR, was calculated for each increment in added bias and imprecision.

The upper bounds for combinations of systematic bias and imprecision in a serum creatinine measurement that would increase the root mean square error in estimating GFR by no more than 10% [an arbitrary modest increase consistent with previous recommendations for the impact of measurement error on clinical utility of laboratory results (32)] are shown in Fig. 3. The serum creatinine measurements obtained in the MDRD Study were assumed to have zero bias; thus, the bias increments should be interpreted as a difference from a zero bias condition. The SD increments were added to the underlying SD in the MDRD Study [2.65 μmol/L (0.03 mg/dL) for creatinine in the 88.4–133 μmol/L (1.00–1.50 mg/dL) range]. The line in Fig. 3 represents combinations of added bias and SD at which the root mean square error was ≤12.22 mL·min⁻¹·(1.73 m²)⁻¹ (a 10% increase). Under the conditions of the simulation analysis, the method performance parameters in Fig. 3 should generalize to other settings. A more detailed description of the simulation analysis can be found in the online Data Supplement that accompanies this report at http://www.clinchem.org/content/vol52/issue1/.

The required laboratory measurement performance for serum creatinine can also be estimated based solely on the underlying biological variability (33). Any approach must consider both imprecision and bias in making an estimate of analytical performance required to meet a clinical interpretation goal. A desirable imprecision goal has been proposed as one-half the intrasubject biological variability because this will not increase the total error more than 12% (32). A more recent recommendation for analytical performance goals based on intra- and interindividual biological variability has included both imprecision and bias and has empirically proposed tiered recommendations in categories consistent with minimum, desirable, and optimal method performance to support clinical interpretation of a result (33). The desirable imprecision is consistent with the previous maximum 12% increase in total error, and the other categories for imprecision and bias are arbitrary extensions to allow categorization of method performance to identify those that may need improvement. The analytical goals for serum creatinine measurement using this approach are summarized in Table 1. The minimum total error goal is estimated at 11.4% and the desirable total error goal at 7.6%.

**Clinical Laboratory–Based Analytical Systems for Measuring Serum Creatinine to Assess GFR**

The methods most widely used to measure serum creatinine are alkaline picrate methods, enzymatic or partially enzymatic assays, and HPLC methods. Isotope-dilution mass spectrometry (IDMS) high-order reference methods have been developed for assignment of reference materials but are available in only a few highly specialized laboratories worldwide.

Because no systematic differences between serum and plasma measurements have been reported, we consider serum and plasma results as equivalent (34). Serum creatinine has been found to remain stable during long-term storage and after repeated thawing and refreezing (35) and for up to 24 h in clotted whole blood at room temperature (36).

**ALKALINE PICRATE METHODS**

The method of Jaffe (37) is commonly used to measure serum creatinine in routine laboratories. The presence of interfering substances, particularly proteins, in serum can lead to the overestimation of serum creatinine by as much as 15%–25% by various Jaffe methodologic applications. Many endogenous and exogenous interfering substances contribute to the lack of analytical specificity in the Jaffe method (38–76). Interferences from glucose (65–67) and acetoacetate (58) are particularly important because diabetic persons are a high-risk population to develop CKD. Several modifications, including optimization of kinetic
enzymatic creatinine methods widely adopted for routine clinical laboratory use are method. The only alternative methods that have been because they have not demonstrated improved perfor-

assays, have been made to improve method specificity and minimize susceptibility to interfering substances (38,77–81).

In a study to evaluate the effect of a compensated Jaffe method on estimated GFR, the MDRD Study equation overestimated GFR by ~50% in individuals with serum creatinine concentrations <155 μmol/L (1.75 mg/dL) (82). In another study, serum creatinine was measured by a new modified kinetic Jaffe reaction and a modified version of the established enzymatic creatinine p-amino-

Table 1. Analytical performance goals for creatinine measurements based on biological variability.

<table>
<thead>
<tr>
<th>CVI*</th>
<th>CVg*</th>
<th>CVa goal*</th>
<th>Bias goal</th>
<th>TE goald</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3%</td>
<td>12.9%</td>
<td>Minimum (0.75 CVI), 3.2%</td>
<td>Minimum [0.375 (CVI² + CVg²)¹/₂], 5.1%</td>
<td>Minimum acceptable performance, 11.4%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Desirable (0.50 CVI), 2.2%</td>
<td>Desirable [0.25 (CVI² + CVg²)¹/₂], 3.4%</td>
<td>Desirable performance, 7.6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimum (0.25 CVI), 1.1%</td>
<td>Optimum [0.125 (CVI² + CVg²)¹/₂], 1.7%</td>
<td>Optimum performance, 3.8%</td>
</tr>
</tbody>
</table>

* CVI, mean intraindividual biological variation (136).
* CVg, mean interindividual biological variation (136).
* CVa, analytical imprecision. When the minimum acceptable imprecision goal is achieved, the contribution of analytical variability to total variability is at most 25%. When the desirable imprecision goal is achieved, the contribution of analytical variability to total variability is <12%. When the optimum imprecision goal is achieved, the contribution of analytical variability to total variability is a maximum of 3% (33).
* TE, total error. Goal was calculated as: bias goal + (1.96 × CVa goal).

ENZYMATIC METHODS

Inorganic chemical–based methods (84,85) that have been developed as alternatives to the alkaline picrate methods have not been widely implemented clinically because they have not demonstrated improved performance compared with the various adaptations of the Jaffe method. The only alternative methods that have been widely adopted for routine clinical laboratory use are enzymatic creatinine methods (86–89). Although the enzymatic methods have been reported to have generally fewer interferences than the Jaffe methods, there have been reports of various substances that do interfere (90–98).

HPLC PROCEDURES

Early reports suggested that HPLC was a fairly sensitive and analytically specific method for measuring serum creatinine (99,100). More than 50 methods for the analysis of creatinine have been described, including cation-exchange, normal-phase, reversed-phase, and reversed-phase ion-pair chromatography (101). Sample deproteinization improves the specificity of creatinine measurement by HPLC by removing many protein-bound endogenous and exogenous compounds without altering the quantification of creatinine (101–111). Several authors have described direct column injection techniques without the deproteinization step (99,112,113).

Interference studies have demonstrated that HPLC methods have greater analytical specificity than conventional methods (101,103,106,107,109–111,114–117). Sample deproteinization combined with the selectivity of HPLC mobile-phase conditions make it unlikely that many substances will interfere; thus, HPLC appears to provide an excellent designated comparison method for in-house use by manufacturers.

MS-BASED PROCEDURES

GC-IDMS is considered the method of choice for establishing the true concentration of creatinine in serum because of its excellent specificity and relative SD <0.3% (118–120). In this procedure, creatinine must be derivatized before GC analysis because of its polarity. In addition, a cation-exchange clean-up step before GC analysis is also necessary because creatine (a compound similar to creatinine) is derivatized into the same chemical species as creatinine. More recently, a method coupling HPLC with IDMS for the direct quantification of creatinine was reported (121). This procedure offers simplicity and speed of analysis with the potential for much quicker turnaround of highly accurate serum creatinine results because only a simple protein precipitation without derivatization is required. A blind international interlaboratory comparison study demonstrated that the liquid chromatography (LC)-IDMS method is comparable to the GC-IDMS method, with an observed bias <0.2% and an expanded uncertainty <0.3% (k = 2) (121).

Performance of Current Routine Methods for Creatinine Measurement

Several proficiency testing (PT) providers offer external quality assessment schemes (EQAS) for assessing accuracy of serum creatinine measurements. Unfortunately, the materials typically used for most interlaboratory PT programs do not give the same numeric relationship between 2 methods as that observed for native clinical samples (i.e., noncommutable) for the majority of routine
methods, including creatinine, used by clinical laboratories (122–125). A recent study by the College of American Pathologists (CAP) found that conventional PT specimens were not commutable with a fresh-frozen serum specimen for 69% of creatinine methods (126). This limitation prevents use of PT results from conventional specimens to evaluate accuracy for an individual laboratory or trueness for a method peer group compared with a reference measurement procedure such as GC-IDMS.

PT and EQAS programs that used carefully collected frozen off-the-clot serum pools, presumably commutable with native clinical sera, and used GC-IDMS as the reference measurement procedure have reported results for evaluation of creatinine (122, 125–128). Collectively, these observations from PT/EQAS programs suggest that a large number of routine methods for serum creatinine are biased high [range, −5.3 to 27 µmol/L (~0.06 to 0.31 mg/dL) at a concentration of ~80 µmol/L (0.90 mg/dL)] and that a standardization program traceable to a high-order reference measurement procedure would allow manufacturers to achieve substantially improved trueness in creatinine results with routine methods.

PT data from a method peer group also provide useful information on the interlaboratory SD, representing total imprecision and including contributions from calibration uniformity within a method group and from within-laboratory imprecision, for measuring the dispersion of routine method results, which affects the total error for creatinine measurement. The 2 largest studies that used a commutable serum sample reported similar method group SDs [0.088–12 µmol/L (0.001–0.131 mg/dL) with a median SD of 5.1 µmol/L (0.058 mg/dL)] and a standardization program traceable to a high-order reference measurement procedure would allow manufacturers to achieve substantially improved trueness in creatinine results with routine methods.

PT data from a method peer group also provide useful information on the interlaboratory SD, representing total imprecision and including contributions from calibration uniformity within a method group and from within-laboratory imprecision, for measuring the dispersion of routine method results, which affects the total error for creatinine measurement. The 2 largest studies that used a commutable serum sample reported similar method group SDs [0.088–12 µmol/L (0.001–0.131 mg/dL) with a median SD of 5.1 µmol/L (0.058 mg/dL)] and a standardization program traceable to a high-order reference measurement procedure would allow manufacturers to achieve substantially improved trueness in creatinine results with routine methods.

Fig. 4. Performance of routine methods compared with the total error limit for serum creatinine.

The data points represent the bias vs a GC-IDMS reference measurement procedure and the interlaboratory SD for 50 routine method peer groups for assay of a fresh-frozen serum pool with creatinine 80 µmol/L (0.90 mg/dL) in the 2003 CAP survey (126) superimposed on the total error budget for creatinine measurements from Fig. 3. Correction of bias would achieve method performance within the limits consistent with a maximum contribution of 10% to the error of estimated GFR for 41 of 50 peer groups.

Within-laboratory imprecision can be determined from internal daily quality-control data. On the basis of data submitted to an interlaboratory quality-control monitoring program (G Cooper, Bio-Rad Laboratories Inc, Quality Systems Division, Irvine, CA., personal communication), intralaboratory CVs correspond to SDs ranging from 1.8 to 7.5 µmol/L (0.02 to 0.085 mg/dL) at the critical creatinine concentration of 88.4 µmol/L (1.00 mg/dL). These data suggest that intra- and interlaboratory error are of similar magnitude and must be controlled to allow method performance within the desired total error goals.

**Resources for Standardization of Serum Creatinine Measurement**

To universally implement GFR estimations based on serum creatinine measurements, a unified effort to standardize routine serum creatinine measurements is required. Establishing measurement traceability is an important tool to achieve the needed comparability in serum creatinine measurement results regardless of the method used and/or the laboratory where the analyses are performed. This effort must involve international cooperation among the in vitro diagnostic (IVD) manufacturers, clinical laboratories, professional organizations, government agencies, and EQAS/PT providers.

Achieving traceability of clinical laboratory serum creatinine measurements through IVD manufacturers would be considerably more efficient than doing so through each of the tens of thousands of clinical laboratories internationally that perform serum creatinine measurements [e.g., in the United States, 97% of participants in the 2003 CAP survey used instruments from 5 manufacturers (126)]. To achieve improved accuracy of creatinine results requires that the values assigned by manufacturers to calibrators and control materials are traceable to high-order reference measurement procedures and reference materials. The International Organization for Standardization (ISO) has developed a written standard that details a pathway for establishing traceability of clinical laboratory measurement results (129).

The only exception to this intervention at the level of the IVD manufacturer is for laboratories that elect to use a nonhomogeneous system, in which a reagent from one manufacturer is used with an instrument or a calibrator from another manufacturer. In these situations, the laboratory must assume primary responsibility and be able to document the accuracy of the reported serum creatinine values.

To address the global need for standardization and traceability of clinical laboratory measurements, the Joint Committee on Traceability in Laboratory Medicine (JCTLM) was created. The JCTLM recently completed the
task of critically reviewing reference materials and reference measurement procedures that were submitted for consideration using criteria set forth in ISO Standards 15194 (130) and 15193 (131), which define the characteristics of higher order reference materials and reference measurement procedures, respectively. A first list of approved reference materials and reference measurement procedures is now available through the Bureau International des Poids et Mesures website (http://www.bipm.org/en/committees/jc/jctlm). These materials and procedures are tools that the IVD industry can use to demonstrate an unbroken chain of traceability back to the SI unit.

REFERENCE MATERIALS
The NIST Standard Reference Material (SRM) 914a, crystalline creatinine, is intended for use in calibration of reference measurement procedures. Calibrator solutions of SRM 914a, prepared by dissolving crystalline creatinine in aqueous buffer, are intended primarily for use in high-order reference measurement procedures (e.g., GC-IDMS and LC-IDMS) and are not generally suitable for direct assay by routine clinical analyzers. NIST (SRM 909b-1 and -2) and the Institute for Reference Materials and Measurements (IRMM; BCR 573, 574, and 575) offer multilevel lyophilized human serum–based certified reference materials with GC-IDMS-assigned values. These materials are intended as trueness control products for high-order reference measurement procedures. Although these are human serum–based materials, the matrix has been altered by converting plasma to serum and by lyophilization, potentially altering the recovery of creatinine in these fluids by routine methods. Because the commutability of these materials with native clinical sera has not been established for routine methods, caution should be exercised when using them as reference materials for calibration purposes or as trueness controls. Reference material that is noncommutable with native clinical serum samples can cause significant error in method calibration.

COMMUTABLE REFERENCE MATERIALS
The NKDEP, CAP, and NIST have collaborated to prepare a human serum-creatinine reference material with acceptable commutability with native clinical specimens in routine methods. This material is a fresh-frozen human serum pool prepared according to Clinical and Laboratory Standards Institute (CLSI) guideline C-37A (132). Two concentrations, ~71 and 354 μmol/L (0.8 and 4.0 mg/dL), have been prepared by supplementation with crystalline creatinine to achieve a material with increased concentration. These materials will be value-assigned by NIST with the GC-IDMS and LC-IDMS reference measurement procedures. The materials will be designated NIST SRM 967, are expected to be commutable with native human sera, and will be validated for commutability with a variety of routine methods. NIST plans to submit this new reference material to JCTLM for inclusion in the approved list.

COMMUTABLE PT/EQA MATERIALS
EQAS and PT providers should make available commutable materials for regularly recurring assessment of serum creatinine measurement performance in routine clinical laboratories. CAP, in collaboration with NKDEP and NIST, has introduced a Calibration Verification/Linearity Survey (LN-24) for serum creatinine [see page 87 in the 2005 Surveys and Educational Anatomic Pathology Programs catalog (http://www.cap.org/apps/docs/proficiency_testing/Surveys_catalog_05.pdf)]. This new proficiency testing survey was initiated in 2004 and covers the range needed to detect early stages of CKD, 44–354 μmol/L (0.50–4.00 mg/dL). Target values are assigned by NIST, using IDMS. This external PT program uses frozen off-the-clot human serum pools prepared according to CLSI C-37A (132). NIST, in collaboration with CAP and NKDEP, will evaluate the commutability of LN-24. EQA and PT surveys will provide IVD manufacturers and individual clinical laboratories an excellent way to validate the traceability of their clinical measurement procedures.

REFERENCE MEASUREMENT PROCEDURES
Three GC-IDMS methods, nominated by the University of Ghent (Belgium), the German Society of Clinical Chemistry and Laboratory Medicine (DGKC), and NIST, have been approved by the JCTLM as reference measurement procedures for serum creatinine (118–120). All 3 of these methods require a separation step to remove creatine, which gives the same derivatization product as creatinine, before derivatization of the creatinine or the GC step and are therefore very time-consuming procedures with limited sample throughput. An alternative approach to standardizing results and establishing traceability to a reference measurement procedure, particularly when the commutability of reference materials is not known, is for IVD manufacturers to split samples with a laboratory performing a reference measurement procedure. An LC-IDMS method has been developed for serum creatinine and nominated in the JCTLM cycle II for consideration as a reference measurement procedure (121). LC-IDMS methods have much simpler and faster sample preparation than do GC-IDMS methods and would be much more amenable to split-sample comparisons on a more timely basis.

Implementation of Estimated GFR Calculated from Serum Creatinine by Use of the Current MDRD Study Equation
The NKDEP recommends immediate implementation of GFR estimation with the MDRD Study equation, before creatinine standardization is implemented. A routine method in a central laboratory was used to develop and validate the MDRD Study equation, but not unexpectedly this method had a small positive bias compared with the
GC-IDMS reference measurement procedure. The difference between the calibration bias for this method and that of most current routine methods is relatively small. Consequently, the impact of bias on the uncertainty in estimated GFR, although not meeting the desirable goals, is still thought to provide substantial clinical benefit in identifying patients with early stages of CKD.

In implementing estimated GFR calculated with the MDRD Study equation, laboratories must address several issues. The equation is applicable to adults 18 years and older. The impact of creatinine measurement error increases dramatically at higher GFRs (lower creatinine); therefore, for an estimated GFR >60 mL·min⁻¹·(1.73 m²)⁻¹, reporting of numeric values is not recommended. Reporting of two values for estimated GFR is recommended: one if the patient is African-American and a second if the patient is not African-American, because the equation does not address mixed ethnicity. The computer implementation must not compute a value if there is any missing information. Finally, the implementation must correct the estimated GFR if there is a correction made to the serum creatinine, to the birth date (age), or to the sex.

When creatinine method standardization and traceability to IDMS are introduced, a clear understanding of the clinical impact is necessary. A systematic program to inform laboratories of the clinical ramifications of a change in calibration of serum creatinine needs to be developed in collaboration with method manufacturers. Because most current routine serum creatinine measurement procedures have a small positive bias, recalibration will decrease the numeric value of the serum creatinine concentration. Because the creatinine method used to develop and validate the MDRD Study equation was not traceable to IDMS, the appropriate changes in the MDRD equation coefficients must be coordinated with the change in creatinine calibration traceability. An improved GFR-estimating equation based on serum creatinine values traceable to IDMS reference measurement procedures will be presented in the fall of 2005 (133).

When introducing revised serum creatinine calibration to be traceable to IDMS, laboratories will need to communicate the following to healthcare providers: the serum creatinine reference interval will change to lower values, calculations of estimated GFR used by pharmacies or other groups to adjust drug dosages will be affected by the decreased creatinine values, measured and calculated creatinine clearance values will increase, and the corresponding reference interval will be different. Note that estimated GFR obtained with the MDRD Study equation is more accurate than creatinine clearance (3), and use of creatinine clearance should be discouraged in patients with normal body size. The magnitude and impact of this calibration change must be carefully established for each creatinine routine method so that after standardization, if an adjustment in decision-making criteria is required, a quantitative link with the current criteria will be available.

Recommendations from Australia (134) and the United Kingdom (135) have recently been published.

**Limitations of the MDRD Study Equation**

There are limitations to use of the MDRD Study equation:

- The performance of the MDRD Study equation in healthy individuals or in patients with CKD and a GFR ≥90 mL·min⁻¹·(1.73 m²)⁻¹ is unclear.
- The MDRD Study equation has not been tested in children, the elderly >75 years, pregnant women, patients with serious comorbid conditions, or persons with extremes of body size, muscle mass, or nutritional status. Application to these patient groups may lead to errors in GFR estimation.
- Variability in serum creatinine measurement renders all estimating equations for GFR, including the MDRD Study equation, substantially less accurate in the reference (normal) and slightly increased range of serum creatinine concentrations [<133 μmol/L (1.50 mg/dL)], which is the relevant range for detecting CKD [GFR <60 mL·min⁻¹·(1.73 m²)⁻¹] in adults.
- The majority of routine methods meet the required precision performance; however, current bias performance of routine methods for serum creatinine is suboptimal for meeting the total error goal for estimating GFR using the MDRD Study equation.
- Analytical nonspecificity biases found in individual patient samples can significantly affect the accuracy of GFR estimates computed from serum creatinine.
- Standardization of calibration does not correct for analytical interferences (nonspecificity bias). Nonspecificity issues must be addressed by IVD manufacturers.
- Implementing traceability of serum creatinine assays to GC- or LC-IDMS will lead to changes in the clinical decision-making criteria currently used for serum creatinine and creatinine clearance and will compromise any clinical decisions based on the estimated GFR unless the estimated GFR is calculated by use of the new MDRD equation based on creatinine values traceable to an IDMS reference method (to be presented in the fall of 2005).
- Adjustment of routine method calibration to be traceable to GC- or LC-IDMS will impact clinical interpretation of creatinine results and will require the following: communication of reference interval changes; communication to pharmacies of the change in creatinine values that will impact drug dose adjustment; communication of change in creatinine clearance values and the corresponding reference intervals; and adjustment of coefficients in the equation used to calculate estimated GFR.

**Recommendations of the NKDEP Laboratory Working Group**

For both IVD Manufacturers and clinical laboratories, the NKDEP Laboratory Working Group is making the following recommendations:
• Implement estimated GFR now, using the MDRD Study equation for routine methods that have not been recalibrated to be traceable to IDMS until a revised MDRD Study equation and routine methods traceable to IDMS are ready for use.

• IVD manufacturers should recalibrate serum creatinine methods to be traceable to IDMS and should coordinate the introduction of recalibrated serum creatinine methods with the introduction of a revised GFR-estimating equation appropriate for use with zero-biased routine methods. If coordination cannot be accommodated, IVD manufacturers should collaborate with the NKDEP and other professional organizations to communicate to customers the clinical issues associated with recalibrating serum creatinine (see the fourth recommendation for NKDEP in collaboration with other professional organizations).

• Report estimated GFR values above 60 mL·min⁻¹·(1.73 m²)⁻¹ as “≥60 mL·min⁻¹·(1.73 m²)⁻¹” and not as an exact number. For values <60 mL·min⁻¹·(1.73 m²)⁻¹, the report should give the numeric estimate rounded to the nearest whole number, such as “35 mL·min⁻¹·(1.73 m²)⁻¹”.

• Report serum creatinine values as mg/dL to 2 decimal places (e.g., 0.92 mg/dL instead of 0.9 mg/dL). Serum creatinine values reported as µmol/L should be reported as the nearest whole number (e.g., 109 µmol/L instead of 109.3 µmol/L).

• After recalibration to IDMS, a realistic total error goal for creatinine measurement is a maximum 10% increase in the relative error of the estimated GFR. Routine methods could achieve this total error goal if analytical imprecision (including between-laboratory calibration variability) is <8% and analytical bias (compared with an IDMS reference measurement procedure) is <5% at all serum creatinine concentrations ≥88.4 µmol/L (1.00 mg/dL).

FOR NKDEP IN COLLABORATION WITH OTHER PROFESSIONAL ORGANIZATIONS

• Determine the differences in clinical decision criteria that may result from establishing traceability of serum creatinine methods to IDMS.

• Develop a replacement for the MDRD Study equation to estimate GFR using serum creatinine measurements that are traceable to IDMS, taking care to ensure that any modification is quantitatively linked to decision outcomes made with the current MDRD Study equation.

• Work with IVD manufacturers and clinical laboratories to coordinate introduction of method traceability to IDMS with introduction of a revised GFR-estimating equation.

• Develop guidelines to help manufacturers communicate to clinical laboratories the best approach to coordinate the introduction of serum creatinine traceability to IDMS with the simultaneous introduction of a new GFR-estimating equation and the resulting changes in clinical interpretation of serum creatinine. The following will need to be addressed: change in the serum creatinine reference intervals; changes in creatinine clearance values and reference intervals; communication to pharmacies of the change in creatinine values that will impact drug dose adjustment; adjustment of coefficients in the equation used to calculate the estimated GFR.

• Coordinate introduction of traceability of serum creatinine to IDMS with PT/EQAS providers to ensure that appropriate participant grading adjustments are made during the transition between calibration schemes.

• Establish a small group of reference laboratories, in the United States and other countries, that can perform high-throughput reference measurement procedures for serum creatinine and provide <1 month turnaround time at reasonable cost such that IVD manufacturers, PT providers, and other interested parties can obtain creatinine results that are accurate and traceable to a high-order reference method. The IFCC has convened a Working Group on GFR that will establish a reference laboratory network for serum creatinine.

• If, in the future, an alternative filtration marker replaces serum creatinine or is used in addition to serum creatinine to estimate GFR, a similar program for assay standardization should be developed. A working group on the standardization of cystatin C is already active in IFCC.

• Professional organizations should implement pro-
grams to educate their memberships regarding the need for reliable serum creatinine measurements and the proper use of the MDRD Study equation to assess CKD risk. The IFCC Working Group on GFR, in collaboration with NKDEP, will coordinate the global introduction of standardized creatinine together with new GFR-estimating equations as well as education of laboratory professionals regarding the importance of assessing CKD risk.

FOR NATIONAL METROLOGY INSTITUTES, REFERENCE LABORATORIES, AND ORGANIZATIONAL MEMBERS OF JCTLM

- Provide tools to assist IVD manufacturers to reduce analytical bias because many automated routine methods can meet or exceed the imprecision goal (less than \( \pm 8\% \)) necessary to meet the maximum 10\% impact on estimated GFR.
- Develop readily available reference materials for serum creatinine with demonstrated commutability to individual patient sera with a wide variety of routine methods and submit them to JCTLM for review and acceptance. NIST SRM 967 [two concentrations, \( \sim 71 \mu mol/L (0.80 mg/dL) \) and 354 \( \mu mol/L (4.00 mg/dL) \)] is expected to fulfill this need when available.
- Make available a high-order reference measurement procedure (e.g., LC-IDMS) with high throughput and validated to have little or no bias relative to GC-IDMS. Such a high-order, high-throughput reference measurement procedure can assist IVD manufacturers in validating the trueness of their methods and can assist in validating commutability of reference materials. Additional reference laboratories will be needed to meet the anticipated demand for analytical services to establish and validate traceability to the reference method.

FOR PT AND EQAS PROVIDERS

- Introduce a regularly recurring program that uses commutable serum materials with target values traceable to IDMS reference measurement procedures, so that IVD manufacturers can, on an ongoing basis, use routine method results to assess the performance of clinical laboratories and the success of their accuracy transfer processes.

Members of the NKDEP Laboratory Working Group

The members of the NKDEP Laboratory Working Group are John Eckfeldt, MD, PhD, University of Minnesota, National Kidney Disease Education Program, Laboratory Working Group Chair; Thomas Hostetter, MD, Director, National Kidney Disease Education Program and Director, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health; Elisa Gladstone, MPH, National Kidney Disease Education Program Associate Director, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health; Joris Delanghe, MD, PhD, University Hospital Ghent (European Communities Confederation of Clinical Chemistry and Laboratory Medicine); James Fleming, PhD, FACB, Laboratory Corporation of America; Neil Greenberg, PhD, Ortho Clinical Diagnostics (Advamed); Ethan Hausman, MD, FAAP, FCAP (US Food and Drug Administration); Glen Hortin, MD, PhD, National Institutes of Health; Harvey Kaufman, MD, Quest Diagnostics; Anthony Killeen, MD, PhD, University of Minnesota (College of American Pathologists); David A. Lacher, MD, MEd, Centers for Disease Control and Prevention; Timothy Larson, MD, Mayo Clinic Renal Laboratory (AACC); Andrew S. Levey, MD, Tufts-New England Medical Center; Frederick Van Lente, PhD, Cleveland Clinic (CKD Epidemiology Collaboration); Greg Miller, PhD, Virginia Commonwealth University (Clinical and Laboratory Standards Institute, formerly NCCLS); Gary Myers, PhD (Centers for Disease Control and Prevention); Mauro Panterghini, MD, Università degli Studi di Milano (International Federation of Clinical Chemistry and Laboratory Medicine); David W. Seccombe, MD, PhD, Canadian External Quality Assessment Laboratory; Michael Welch, PhD, National Institute of Standards and Technology. Names that appear with an organization in parentheses have been designated that organization’s official working group representative.

We thank Elisa Gladstone for her outstanding assistance with the preparation of this manuscript and for her excellent logistic support of the Laboratory Working Group.

References


90. Mitchell RT, Marshall LH, Lefkowitz LB, Stratton CW. Falsely


