3-O-Methyldopamine (3-O-Methoxytyramine) Interferes with the Internal Standard 3,4-Dihydroxybenzylamine in a Plasma Catecholamine HPLC Method

To the Editor:

Plasma catecholamines are routinely assayed for the detection of pheochromocytoma and in clinical trials evaluating hemodynamic function in intensive care patients (1). HPLC combined with electrochemical detection is widely applied for the quantification of plasma catecholamines (2). We use a commercially available method for plasma catecholamine measurement (Chromsystems), including a pretreatment step for plasma samples that uses the relative nonspecific absorption of analytes onto Al₂O₃. Isocratic separation is then performed on a reversed-phase C18 column. Assay conditions include a 0.8 mL/min flow rate and a mobile phase containing salts, methanol, and an ion-pairing reagent. Quantification is performed with an L 3500 electrochemical detector (Recipe) at a potential of +500 mV.

Fig. 1. HPLC chromatograms of plasma catecholamine analysis.

The length of the run as depicted is 25 min. The retention times of norepinephrine, epinephrine, DHBA, and dopamine were 8 min, 9 min, 13–14 min, and 19–21 min, respectively. The retention time of DHBA is indicated by an arrow. Inserts (black frame) show the internal standards (IS) of the calibrator of the respective runs for comparison. (A), chromatographic separation of plasma catecholamines of a patient receiving dopamine in therapeutic dosage; the spuriously increased internal standard is indicated (arrow). (B), chromatographic separation of an aqueous solution of dopamine 10 mg/L (Carinopharm). In addition to the expected dopamine peak, an increase of the internal standard peak is also visible (arrow).
3,4-Dihydroxybenzylamine (DHBA) is used as the internal standard.

We observed an interference with the internal standard in patients taking part in a clinical study in our intensive care unit. Evaluation of the case histories revealed that these patients have received dopamine intravenously before blood sampling. An example of this interference in a patient receiving dopamine is shown in Fig. 1A. The sample from this patient showed an approximately 1.5-fold spurious increase in the response of the internal standard compared to other samples in the same run. Because calculations of catecholamine concentrations were performed relative to the internal standard, falsely low results (approximately 67% of the actual concentration) would be reported if the falsely increased internal standard response was not manually corrected.

Testing dopamine (Carinopharm) in vitro, we found the same interference (Fig. 1B). Because dopamine has a retention time of approximately 20 min with this method, a dopamine impurity might lead to the spuriously increased internal standard readings. Therefore, we asked the manufacturer about potential contaminants and learned that minor amounts of 3-O-methyldopamine, 4-O-methyldopamine, and homoveratrylamine are detected via thin-layer chromatography within the QC procedure. Therefore, we analyzed the retention time of these substances. We found that 3-O-methyldopamine (Fisher Scientific) interfered with the HPLC analysis, leading to an increase of the internal standard only with substantial amounts, e.g., 100 mg/L. The other tested substances showed no interference with the internal standard (data not shown). To analyze whether the disturbance is specific for the applied dopamine, we tested dopamine standards from 2 other sources, Sigma-Aldrich and Fresenius-Kabi, and also found an artificial increase of the internal standard (data not shown). In vivo, 3-O-methyldopamine is a degradation product of dopamine. Dopamine is converted to 3-O-methyldopamine via catechol-O-methyltransferase (3). The concentration of 3-O-methyldopamine in plasma is usually <1 µg/L, but may increase markedly in pheochromocytoma patients. We tested whether 3-O-methyldopamine also interfered with catecholamine determination in urine. To retain catecholamines we applied a urine pretreatment process that immobilizes phenylboronic acid via formation of cyclic esters. This specific interaction is possible only for cis-diol compounds. Indeed, this pretreatment eliminated the disturbance caused by 3-O-methyldopamine (data not shown).

Because we knew that the chromatographic separation of 3-O-methyldopamine, the peak with the retention time of 13 min, should contain 2 different substances, DHBA and 3-O-methyldopamine, we repeated the analysis of 3-O-methyldopamine with increased resolution of our method and a new column. The columns not able to resolve the peak, however, were maintained according to the manufacturer’s instructions. Alternatively, more specific detection methods, such as mass spectrometry, might be applied, but such methods are not as widely used for catecholamine analysis. Interference with the internal standard is a major pitfall in HPLC analysis. Numerous reports are published that describe such interferences, which are often found coincidentally (4, 5), and they probably go unrecognized even more frequently.

Our findings show that the possibility for spurious increases in the internal standard must be considered because they may lead to falsely low results. DHBA, a widely used internal standard for the determination of catecholamines, would not be easy to replace, but minor modifications of the method might be beneficial and must be developed. As this report exemplifies, manufacturers and analysts are permanently challenged to improve their methods.

Grant/funding support: None declared. Financial disclosures: None declared.

Acknowledgments: We thank Steffi Hasanovic for excellent technical assistance.

References

Albrecht Pfafflin
Erwin Schleicher
Karsten Müssig

Department of Internal Medicine IV, Clinical Chemistry University of Tuebingen
Tuebingen, Germany

* Address correspondence to this author: Department of Internal Medicine IV, Clinical Chemistry (Central Laboratory), University of Tuebingen, Hoppe-Seyler-Str. 3, D-72076 Tuebingen, Germany. Fax: 49-7071-29-4696; e-mail Karsten.Muessig@med.uni-tuebingen.de.

DOI: 10.1373/clinchem.2007.094276

The Vietnamese Khin Population Harbors Particular N-Acetyltransferase 2 Allele Frequencies

To the Editor: Vietnam ranks 13th among the 23 countries that must deal with 80% of global deaths from tuberculosis (1). The National Tuberculosis Control Program of Vietnam implemented the WHO DOTS (directly observed therapy, short course) strategy in 1989. The standard treatment regi-
men for previously untreated patients consists of 2 months of streptomycin, isoniazid (INH), rifampin, and pyrazinamide, followed by 6 months of INH and ethambutol (the 2SHZ/6HE regimen).

INH, a central component of this chemotherapy, is extensively metabolized by the polymorphic N-acetyltransferase 2 (NAT2) enzyme. N-acetyltransferase 2 (arylamine N-acetyltransferase) (NAT2) gene polymorphisms are associated with individual phenotypes generally classified as INH rapid (homozygous or heterozygous) and slow acetylators. Among the described NAT2 alleles, NAT2*5, *6, *7, *12, *14, *17, and *19 have been associated with the slow acetylator phenotype, and NAT2*13 and *4 are associated with fast acetylation. Under treatment with standard dosages, patients who are slow metabolizers may be at risk for INH side effects, which include peripheral neuropathy, hepatic toxicity, and psychotic symptoms. On the other hand, insufficient drug concentrations may be more frequent among fast metabolizers. The determination of the NAT2 gene allele pattern of a population can supply valuable baseline information for determining drug dosage and expected efficacy for specific groups.

Although Vietnam has exceptional DOTS coverage, with extensive use of INH, information on population-related NAT2 pharmacogenetics has been unavailable. Accordingly, we sequenced the full NAT2 gene in 72 unrelated adult Vietnamese volunteers [51 male and 21 female, average age: 29 (18–45) years], sampled at the National Institute of Malaria Research, Parasitology and Entomology, Hanoi, Vietnam. With the exception of 2 Thai individuals, all volunteers were self-identified as belonging to the Khin ethnic group, the prevalent population in Vietnam. All study participants gave written informed consent. All protocols were approved by the Ministry of Health Hanoi, Vietnam; the Swedish Medical Products Agency, Uppsala, Sweden; and the Ethics Committee of Göteborg University.

We used an ABI 6100 Nucleic Acid PrepStation (Applied Biosystems) to extract genomic DNA from peripheral blood samples. The NAT2 gene was analyzed through full sequencing with PCR amplification in ABI 7420 thermocyclers. The primers used were as follows: 5′-TCACAC GAGGAAATCAAATG-3′ (NAT2fw) and 5′-GTGCTATCAGATATCCTCT CTACC-3′ (NAT2rev).

Our results show that 90% (65 of 72) of the investigated individuals harbored at least 1 NAT2 fast allele (Table 1), suggesting a particularly low frequency of slow acetylators, even in an Asian population. This high prevalence of fast acetylator–associated alleles is essentially attributable to the exceptionally high frequency of the NAT2*13. Although this allele has been observed at frequencies ≤5% in other populations, in our study we detected it at a prevalence of 31%, significantly different from previous studies on a sample of 264 Thai individuals (*13 <5.1%; P <0.0001) (2) (Table 1). Notably, the NAT2*7 allele, consistently found in Asian populations, was not observed in our study. This finding also contrasts with the 20% prevalence in the Northeast Thai population, and the 6.3% prevalence among 64 individuals from the Cambodian Khmer population (6.3%) (2) (Table 1).

### Table 1. Genotype frequencies and deduced phenotype among Vietnamese individuals predominantly from the Khin ethnic group.

<table>
<thead>
<tr>
<th>Phenotype</th>
<th>NAT2 genotype</th>
<th>n</th>
<th>Frequency (95% CI)</th>
<th>alleleb n</th>
<th>Vietnamc</th>
<th>Thailande</th>
<th>Cambodiaf</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rapid</td>
<td>*4/*4</td>
<td>14</td>
<td>0.194 (0.111–0.305)</td>
<td>*4 63</td>
<td>0.444 (0.362–0.525)</td>
<td>0.381</td>
<td>0.484</td>
</tr>
<tr>
<td>(homozygous)</td>
<td>*4/*13</td>
<td>28</td>
<td>0.389 (0.276–0.511)</td>
<td>*13 44</td>
<td>0.310 (0.234–0.386)</td>
<td>&lt;0.051</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>*13/*13</td>
<td>5</td>
<td>0.069 (0.023–0.155)</td>
<td>*12 4</td>
<td>0.028 (0.008–0.071)</td>
<td>&lt;0.051</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>*6B/*13</td>
<td>21</td>
<td>0.051 (0.013–0.136)</td>
<td>*5 4</td>
<td>0.028 (0.008–0.071)</td>
<td>0.156</td>
<td>0.038</td>
</tr>
<tr>
<td></td>
<td>*6B/<em>13</em>6A</td>
<td>21</td>
<td>0.051 (0.013–0.136)</td>
<td>*6 25</td>
<td>0.176 (0.113–0.239)</td>
<td>0.325</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td>*6B/<em>13</em>6A</td>
<td>7</td>
<td>0.056 (0.015–0.136)</td>
<td>*7 0</td>
<td>0</td>
<td>0.204</td>
<td>0.063</td>
</tr>
<tr>
<td>(heterozygous)</td>
<td>*13/*12B or *12A/*13B</td>
<td>1</td>
<td>0.014 (0.0003–0.075)</td>
<td>*19 2</td>
<td>0.014 (0.002–0.050)</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Slow</td>
<td>*6A/*6A</td>
<td>7</td>
<td>0.097 (0.040–0.190)</td>
<td>*6A 7</td>
<td>0.097 (0.040–0.190)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* All alleles were found to be in Hardy-Weinberg equilibrium.
* The ambiguous *4/*12B or *12A/*13B haplotype found in 1 individual was not included in the allele frequency calculations.
* 95% CI.
* Ref. (2).
* Distinction was not possible between *4/*6A and *6B/*13. Because *6B was not otherwise found, the former was assumed.
* ND, not detected, although some alleles might be included in the NAT4 group, due to allele differentiation limitations of the performed essays.
* Nondistinguishable haplotypes: *5g/*12A from *5B/*12B (1 individual), as well as *5B/*12A from *5C/*12C (3 individuals)—the overall group was classified as *5/*12, because significant differences between the suballeles concerning the metabolism of INH have not been documented.
* Distinction was not possible between these 2 haplotypes, but predicted to represent a fast acetylator.
During analysis we found cases in which it was not possible to distinguish between haplotypes (Table 1) because the presence of multiple heterozygous single-nucleotide polymorphisms prevented us from determining which alleles were in cis (same chromosome) or trans (different chromosomes). For example, when we have both the 282 C>T and 803 A>G in the heterozygote form, 2 possible haplotypes appear: NAT2*4/*12B (one chromosome 282C/803A + the other 282T/803G) and NAT2*12A/*13 (282C/803G + 282T/803A).

The observed 9.7% frequency of slow acetylators among the Vietnamese Khin predicts a low incidence of adverse effects caused by overexposure to INH, and potential drug-drug interactions associated with this drug (3). For this minority group of slow acetylators, a decrease in drug dosage would likely reduce the risk of adverse events without compromising the efficacy of the drug, because a dose of INH as low as 3 mg/kg has recently been shown to be sufficient for a successful therapy (4). On the other hand, a dose reduction below 6 mg/kg in fast acetylators can compromise INH efficacy. The observed high frequency of fast alleles may hypothetically affect the success of the standard INH dosing in the overall Vietnamese population, particularly in terms of early bactericidal activity (5).

Limited pharmacogenetic data are available to account for the diversity of the Southeast Asian populations. Future studies should be performed to identify particular population characteristics that might influence pivotal pharmacotherapy in these geographic regions.

Grant/funding support: This study was supported by the Swedish International Development Cooperation Agency (SIDA/SAREC). I.C. is the recipient of a scholarship from Fundação para a Ciência e Tecnologia, Portugal (Ref. SFRH/BD/2002/8887). Financial disclosures: None declared.

References

C-Reactive Protein (CRP) in Neonates: Comparing VITROS Slide and High-Sensitivity CRP Methods

To the Editor:
C-reactive protein (CRP) is measured on VITROS® Chemistry Systems (Ortho-Clinical Diagnostics) using a “MicroSlide” method. At our hospitals plasma CRP concentrations <10 mg/L are used as an indicator that it is safe to discontinue antibiotic therapy in neonates with known or suspected sepsis (1). After implementing the neonatal sepsis protocol, we found that the VITROS CRP slide gave results up to 33 mg/L higher in neonatal specimens (Fig. 1, open triangles) than the Behring Nephelometer II (BNII, Dade Behring) high-sensitivity CRP (hsCRP) assay, a well-characterized method for CRP measurements (2).

According to the manufacturer’s instructions, hemolysis falsely increases VITROS CRP slide results, but the increases observed were not limited to hemolyzed specimens. It has been reported (3) that collecting specimens in serum separator tubes increases CRP values measured by the VITROS CRP slide; however, the differences we saw were much greater than those attributed to serum separator tubes. Because the falsely increased results in neonates were at or near our decision point of 10 mg/L, we concluded that the VITROS CRP slide method was not suitable for use with our neonatal
sepsis protocol. We began referring neonatal specimens to laboratories able to run CRP on the BNII, the IMMULITE® immunoassay system (Diagnostic Products), or the COBAS Integra® 400 (Roche Diagnostics). We had previously shown (data not included) that the IMMULITE and Integra CRP methods correlated very well with the BNII.

At this point we purchased the VITROS® 5,1 FS Chemistry System, which in addition to analyzing CRP by the MicroSlide method, also determines hsCRP by an immuno-turbidimetric MicroTip method. According to the manufacturer, specimens with hsCRP greater than the reportable range of 15 mg/L should not be diluted but retested. Because the MicroSlide method was not suitable for neonates, and we were seeking a faster alternative to sending specimens to another laboratory, we evaluated the VITROS hsCRP MicroTip method, including dilutions, for use with neonates (age <6 months). We tested 249 serum and heparinized plasma specimens left over after clinical testing by one of the other analyzers. This sample set included 143 specimens by the BNII (range 0.2–351 mg/L), 95 by the IMMULITE (range 0.1–139 mg/L), and 11 by the Integra (range 1.0–49 mg/L). Specimens with CRP concentrations >15 mg/L were diluted with saline by the instrument and retested. The institutional review boards of the University of Utah and Intermountain Healthcare gave approval for submission of this data.

Our results are shown by the solid squares in Fig. 1. The VITROS hsCRP method showed markedly better agreement with the BNII, IMMULITE, and Integra than the MicroSlide CRP method, especially at our decision point of 10 mg/L. The simple linear regression line between VITROS hsCRP MicroTip and the 3 comparative methods, taken together, had a slope of 1.16 (5%–95% CI 1.14–1.17), an intercept of –0.9 mg/L, and R = 0.996. Regression analysis data for individual instruments were as follows: BNII slope 1.16, intercept –0.6 mg/L, R = 0.997; IMMULITE slope 1.12, intercept –0.9 mg/L, R 0.987; and Integra slope 1.14, intercept 0.6 mg/L, and R = 0.992.

Using specimens from both neonates and older patients, we found that the positive bias was introduced by diluting the specimen, manually or by the instrument, with deionized water, saline, or the 0 calibrator. The bias was 16% using a 10-fold dilution and 21% using a 30-fold dilution. Undiluted specimens showed no bias. We speculate that the dilution bias may be the reason that diluting the specimen is not recommended by the manufacturer. Although the hsCRP method is positively biased when the specimen is diluted, the bias is reproducible and relatively small, and does not preclude using the test to manage neonates with sepsis.

Grant/funding support: None declared. Financial disclosures: None declared.

References


Phillip R. Bach1*
Brian W. Davis2
David Loughmiller3
Janet Oertli4
Jacquie Taylor5

1 Primary Children’s Medical Center
Department of Pathology
Salt Lake City, UT

2 Utah Valley Regional Medical Center, Provo, UT

3 Dixie Regional Medical Center
St. George, UT

Fig. 1. Bland-Altman plot showing the difference between CRP by VITROS and comparative methods.

Δ, VITROS CRP Slide vs Behring Nephelometer II. ■, VITROS hsCRP MicroTip method vs Behring Nephelometer II, DPC IMMULITE, or Roche Integra. Dotted vertical line is the sepsis protocol decision point of 10 mg/L.
To the Editor:

Whether serum or plasma is the best specimen for determination of matrix metalloproteinases (MMPs) is a matter of debate, as are the influences of sample collection and processing on MMP concentrations (1–3). MMP-2 concentrations do not differ significantly in plasma and serum, whereas MMP-9 concentrations are significantly higher in serum than in plasma, and in plasma differently affected by anticoagulants (2,3). Furthermore, both unexpected protein peaks and increased MMP-9 concentrations occur in serum samples collected in tubes coated with clot accelerators such as kaolin or silica gel (2,4).

To find an explanation of the differences in MMP-9 concentrations among citrate plasma, serum, and serum with clot accelerators (Sca), we analyzed the effects of silicate on MMP in whole blood (WB), plasma, serum, and buffy coat (BC), as well as in culture media of U-937 myelomonocytic leukemia cells (SF U937).

We obtained peripheral blood (PB) and BC samples from 30 healthy volunteers age 23–55 years (median 36 years). Samples were collected into plastic tubes (Vacutainer® from Becton Dickinson). After centrifugation at 500 g for 15 min at 4 °C, the supernatants of cell culture, BCs, plasma, and serum with and without silicate were analyzed by use of Western blot (75–777 and GE-213 monoclonal antibodies against MMP-2 and -9, respectively; Calbiochem) and gelatin zymography (7.5% polyacrylamide gels containing 2 g/L gelatin 90 Bloom type A from porcine skin; Sigma) (2). U-937 cells were cultured in serum-free conditions (to avoid endogenous bovine serum gelatinases) and treated for 3 h with silicate (54 mg/L). Calibrators were prepared from capillary WB (2). The accuracy/precision of gelatinolytic activities were evaluated by zymogram densitometry with Image Pro-Plus software (Cybernetics) (4). Differences were compared using the Mann–Whitney U-test; P values <0.05 were considered statistically significant. All study participants gave informed consent, and the work was carried out in accordance with the ethics standards of the Helsinki Declaration of 1975, as revised in 1983.

Western blots identified pro-MMP-2 (Gelatinase A, EC 3.4.24.24, 72 kDa) and pro- and complexed forms of MMP-9 (Gelatinases B, EC 3.4.24.35, of 92, 130, and 225 kDa) in whole PB (Fig. 1, lanes 2 and 3). Citrate plasma results showed that pro-MMP-2 expression did not differ significantly between serum and plasma, nor did it change with silicate treatment in any of the paired sample types. MMP-9 forms were found in significantly higher amounts in serum (mean 5-fold higher; P <0.001) than in citrate samples (Fig. 1, lanes 6 vs 4 serum). However, addition of silicate to previously separated citrate plasma and serum did not noticeably change the zymographic profile of MMP-9 with respect to untreated samples (Fig. 1, lane 4 vs 5, and lanes 6 vs 7).

Addition of silicate to citrate plasma tubes before PB collection increased MMP-9 (Fig. 1, lanes 8 and 9) a mean of 4-fold (P <0.001). Addition of silicate into empty plastic tubes for serum collection before PB addition also significantly increased MMP-9 concentrations (P <0.001) (data not shown). Addition of silicate to citrate and serum tubes before PB addition resulted in similar trends of MMP increase vs silicate concentration: MMP-9 activity (µg/L) = 28.0 × silicate (mg/L) – 9.4, r² = 0.94, and MMP-9 activity (µg/L) = 30.6 × silicate (mg/L) + 91.6, r² = 0.88, respectively.

When samples were collected into empty serum tubes to which buffered silicate (silicate dissolved in PBS containing 137 mmol/L NaCl, 10 mmol/L phosphate, 2.7 mmol/L KCl, pH 7.4) was added before PB collection, all MMP-9 forms were increased, and the zymographic profile was similar to that of Sca (Fig. 1, lanes 10 and 11). Samples collected in the presence of nonbuffered silica (nonsoluble silica particles sprayed into plastic tubes or with silica-gel) (http://catalog.bd.com/ecat/msds/d01/vs60313.pdf) showed MMP-9 release 1.5-fold higher than for buffered silicate (data not shown). Serum collected in plastic tubes with clot accelerators showed the highest MMP-9 activities (Fig. 1, lane 11). The addition of silicate to BCs isolated from citrate PB significantly enhanced MMP-9 release in buffered solution (Fig. 1, lanes 12 and 13). Similarly, silicate addition to U-937 cells cultured in serum-free media significantly increased MMP secretion (Fig. 1, lanes 14 and 15). Thus silicates increase in vitro release of MMP-9 forms from leukocytes. Our observations are consistent with the findings that during silicosis both macrophages and lymphocytes secrete enhanced amounts of MMP-9 forms (5).

To optimize the diagnostic accuracy of PB MMPs as biomarkers, we strongly recommend avoiding the use of serum samples, particularly in serum with clot activators containing silica/silicate. We believe the increased MMP-9 observed in these specimens reflects both the interfering effects of the coagulation/fibrinolysis processes (4) and the induction by silicates of MMP-9 release from leukocytes.
Grant/funding support: None declared. Financial disclosures: None declared.

References


Ferdinando Mannello1*
Gaetana A. Tonti1
Jose E. Tanus-Santos2
Raquel F. Gerlach3

1 Institute of Histology and Laboratory Analysis, Faculty of Sciences and Technologies University of Urbino Carlo Bo Urbino, Italy

2 Department of Pharmacology Faculty of Medicine University of Sao Paulo Ribeirao Preto, Brazil

3 Department of Morphology Estomatology and Physiology Dental School University of Sao Paulo Ribeirao Preto, Brazil

* Address correspondence to this author at: Institute of Histology and Laboratory Analysis, Faculty of Sciences and Technologies, University Carlo Bo, Via O. Ubaldini 7, 61029 Urbino (PU), Italy. Fax 39-0722-322370; e-mail f.mannello@uniurb.it.

DOI: 10.1373/clinchem.2007.090548
Endorsement of the STARD Statement by Biomedical Journals: Survey of Instructions for Authors

To the Editor:
The Standards for the Reporting of Diagnostic accuracy studies (STARD) statement was first published in 2003 (1, 2). In February 2006, the International Committee of Medical Journal Editors decided to encourage the use of the STARD statement by including it in their Uniform Requirements for Manuscripts Submitted to Biomedical journals (www.icmje.com).

The STARD statement contains a checklist of 25 recommended items for the reporting of diagnostic accuracy studies and encourages the use of a flow diagram to represent the design of the study and the flow of patients through the study (1, 2).

Early evaluation of the STARD statement showed that the quality of diagnostic research reports had improved slightly over time, without a more pronounced effect in adopting journals (i.e., journals mentioning STARD in their Instructions for Authors) (3). We determined to what extent journals have adopted the STARD statement and incorporated the recommendations in their Instructions for Authors.

We identified the top 50 journals that frequently publish diagnostic accuracy studies, including general and internal medical journals (26%) and specialty journals (74%) (4). In February 2007, 1 reviewer (N.S.) searched for Instructions for Authors on each journal’s website and extracted any text mentioning STARD. Editors of 8 biomedical journals who were included in the STARD evaluation study (3) were asked to participate in a discussion about the quality of diagnostic research reports, factors affecting the use of the STARD statement within the editorial process, and how to improve the quality of diagnostic accuracy studies. A survey of editors of nonadopting journals was also carried out to evaluate their interest in adopting the STARD statement.

The STARD statement was mentioned in the Instructions for Authors in 19 of the 50 journals (38%). STARD was mentioned more often in general and internal medicine journals (6 of 13, 46%) than in specialty journals (13 of 37, 35%). The 19 adopting journals referred to the STARD initiative paper (11 of 19, 58%) and/or the STARD statement on the CONSORT website (www.consort-statement.org) (8 of 19, 42%). Three journals referred to the STARD background document. Only 2 journals referred directly to the STARD website (www.stard-statement.org).

In the adopting journals, we observed a broad variation in the language explaining how to use the STARD statement. Five journals clearly described their expectation regarding the use of the STARD statement, asking authors to submit the STARD checklist and/or to include a flow diagram in the manuscript. Most adopting journals, however, only referred to the STARD statement, without describing their expectations regarding use of the statement (Table 1).

Editors of nonadopting journals indicated that they do not need these guidelines because their reviewers and editors are well attuned to the items listed on the STARD checklist, and they publish relatively few diagnostic accuracy studies. Only 1 nonadopting journal became an adopter after our survey (Table 1). Editors of adopting journals indicated that examples of flow diagrams and proper references to statistical papers would be helpful for authors.

The STARD statement was developed to improve the quality of diagnostic research reports and to facilitate adequate appraisal of the internal and external validity of the study being reported. Our results show that not all journals that publish diagnostic accuracy studies refer to the STARD statement in their Instructions for Authors, and that those that do refer to STARD use varying language in their instructions. This variability may be a reason for the incompleteness of reporting still observed in major journals (3).

Editors have suggested that it takes time before authors and reviewers actually use guidelines such as STARD. Dissemination takes time, but we question whether implementation is a gradual and automatic process. We are aware that editors must make decisions and compromises in what they ask from authors and reviewers. Some editors believe that they ask too much from authors if they require a STARD checklist and flow diagram, whereas others feel that such a requirement may be the solution to improving the completeness of reporting of diagnostic accuracy studies.

In addition to including the STARD statement in the Instructions for Authors, journals could improve the quality of diagnostic research reports by more clearly communicating their expectations regarding use of the STARD statement and providing reader- and printer-friendly versions of these instructions on the journal website. The STARD checklist can be used by reviewers as a tool during the evaluation process and by authors as a means to assess the completeness of the reporting in their manuscripts. We believe that continuous attention of all parties involved is needed to improve the quality of diagnostic research reports.

Grant/funding support: This project was funded by grants from The Netherlands Organisation for Health Research and Development, International Federation of Clinical Chemistry and Laboratory Medicine, and the NHS National Knowledge Service (as part of EQUATOR project).

Financial disclosures: None declared.

Acknowledgments: We thank the editors of the British Medical Journal, the Lancet, The Journal of the American Medical Association, New England Journal of Medicine, Annals of Internal Medicine, Clinical Chemistry, Gut, and Radiology and their collaborators for allowing us to have meetings to discuss measures to improve the quality of diagnostic research papers.
Table 1. The citations of STARD statement and the strength of direction regarding the use of the STARD statement by the 19 adopting journals.

<table>
<thead>
<tr>
<th>Adopting journals</th>
<th>Citations of the STARD statement</th>
<th>Date of last update of the Instructions for Authors*</th>
<th>Strength of direction regarding the use of the STARD statement in the Instructions for Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annals of Emergency Medicine</td>
<td>No No Yes Yes No Yes NR</td>
<td></td>
<td>Should reflect</td>
</tr>
<tr>
<td>Annals of Internal Medicine(^d)(^e)</td>
<td>No No No Yes Yes No NR</td>
<td></td>
<td>Consult</td>
</tr>
<tr>
<td>Annals of the Rheumatic Diseases</td>
<td>No No Yes Yes No No NR</td>
<td></td>
<td>Should conform</td>
</tr>
<tr>
<td>Archives of Disease in Childhood</td>
<td>No No Yes No No No NR</td>
<td></td>
<td>Should conform</td>
</tr>
<tr>
<td>British Journal of Ophthalmology</td>
<td>No No Yes No No No NR</td>
<td></td>
<td>Should conform</td>
</tr>
<tr>
<td>British Journal of General Practice</td>
<td>No Yes No No No No NR</td>
<td></td>
<td>Should follow</td>
</tr>
<tr>
<td>British Medical Journal(^d)</td>
<td>No Yes Yes Yes No No NR</td>
<td></td>
<td>Please follow/report</td>
</tr>
<tr>
<td>Cancer</td>
<td>No No Yes No No No NR</td>
<td>July 2007</td>
<td>Encourage to submit</td>
</tr>
<tr>
<td>Circulation</td>
<td>No No No Yes No No NR</td>
<td>See</td>
<td>Encourage to submit</td>
</tr>
<tr>
<td>Clinical Chemistry(^d)(^e)(^r)</td>
<td>No No Yes Yes Yes Yes April 2006</td>
<td>Complete</td>
<td>Strongly recommended</td>
</tr>
<tr>
<td>Clinical Radiology(^d)(^f)</td>
<td>No Yes No No No No NR</td>
<td></td>
<td>Should refer to</td>
</tr>
<tr>
<td>Gut</td>
<td>No No Yes No No No NR</td>
<td>January 2007</td>
<td>Encourage to submit</td>
</tr>
<tr>
<td>The Journal of the American Medical Association(^f)</td>
<td>No Yes No No No No NR</td>
<td>January 2007</td>
<td>Encourage to submit</td>
</tr>
<tr>
<td>Journal of Clinical Microbiology</td>
<td>No Yes No No No No Jan 2007</td>
<td></td>
<td>Should refer to</td>
</tr>
<tr>
<td>Lancet(^f)</td>
<td>Yes No No No No NR</td>
<td></td>
<td>Must be reported</td>
</tr>
<tr>
<td>Neurology(^f)</td>
<td>No Yes No No No No December 2006</td>
<td>Requires</td>
<td>Requires</td>
</tr>
<tr>
<td>Obstetrics and Gynaecology</td>
<td>No No Yes No No Yes July 2007</td>
<td></td>
<td>As resource</td>
</tr>
<tr>
<td>Radiology(^d)</td>
<td>No No Yes No No No NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>American Journal of Roentgenology(^d)(^d)</td>
<td>Yes Yes No No No No July 2007</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^b\)www.consort-statement.org.
\(^c\) Date of last update of Instructions for Authors on journals’ website (search was carried out on 11 July 2007); NR, not reported.
\(^d\)Published the STARD initiative paper.
\(^e\)Published the background document.
\(^f\)Published an editorial or commentary about the STARD statement.
\(^g\)Became an adopter after the survey in May 2007.
To the Editor:

Measurement of immunoglobulin $\kappa$ and $\lambda$ free light chains (FLC) has improved the diagnostic evaluation and monitoring of monoclonal gamopathies (1), particularly light chain multiple myeloma (2), nonsecretory multiple myeloma, plasmacytoma, systemic amyloid light chain (AL) amyloidosis, heavy chain myeloma (3), and monoclonal gammapathy of undetermined significance (4). FLC measurement methods, however, are affected by analytical problems inherent to nephelometric techniques.

A 76-year-old woman was admitted for evaluation of bone pain and hypercalcemia. The pains were located in the left inguinal fold. Examination showed no lymphadenopathy or hepatosplenomegaly. X-rays of the skeleton showed 2 voluminous gaps in the pubic branches. Laboratory examinations showed hypercalcemia of 2.91 mmol/L and renal insufficiency. Complete blood cell count showed a macrocytic anemia with a hemoglobin concentration of 97 g/L, rouleaux formation, and circulating plasmacytes. Bone marrow contained 85% dystrophic plasmacytes. Serum total protein was 68 g/L. Protein electrophoresis showed hypogammaglobulinemia of 2.2 g/L [reference interval (RI) 7–11 g/L] with a moderate increase of the $\alpha$-2 fraction (12.2 g/L, RI, 6–9 g/L). Immunofixation electrophoresis of blood and urine showed a major band of monoclonal $\kappa$ FLC migrating in the $\alpha$-2 region. Serum FLC measurement showed a major increase in the $\kappa$ FLC (Table 1). We arrived at a diagnosis of light chain $\kappa$ multiple myeloma.

In our laboratory, serum $\kappa$ and $\lambda$ FLC measurements are carried out with a nephelometric technique on the BNprospec automat (Dade Behring) using the Freelite reagents (The Binding Site) according to the manufacturer’s recommendations. A 1st FLC measurement is recommended on a sample diluted at 1:100. If the result is lower than the low end of the assay range, the machine automatically starts the assay again by diluting at 1:20 and then at 1:5. Likewise, if the concentration in FLC is higher than the high value of the range, the analyzer automatically starts serial dilutions and continues to a dilution of 1:32 000.

Initially, the $\kappa$ chains were <0.294 mg/L [RI, 3.3–19.4 mg/L (5)] and $\lambda$ chains were <0.405 mg/L [RI, 5.7–26.3 mg/L (5)]. In view of this unusual result (double negativity), a 1:32 000 dilution was made manually and the $\kappa$-chain concentrations were then >60 100 mg/L. Measurement at a dilution of 1:2000 was then carried out for the $\lambda$ chains, and the result was <0.41 mg/L (Table 1). Serum immunofixation confirmed the result, with a strong $\kappa$ monoclonal band. Thirty-five days later the $\kappa$ chains were 112 mg/L and $\lambda$ chains <0.405 mg/L. Given the patient’s history, however, a measurement at a dilution of 1:32 000 was made manually for $\kappa$ FLC, and the concentration obtained was 12 600 mg/L (Table 1).

In this patient, we repeatedly observed that very high FLC concentrations yielded paradoxically low results. This situation is consistent with antigen excess. The “limitations” section of the product insert from the company does alert the user to the possibility of this phenomenon, but to our knowledge this report is the 1st to demonstrate that the problem of antigen excess in this assay is not merely theoretical but can actually occur clinically. The main risk is that a value within the RI, or one that is substantially underestimated, will be reported instead of the true value. To eradicate antigen excess, any suspect sample must be tested again at a higher dilution. Serum FLC measurements should always be interpreted together with other laboratory test results and clinical findings. This protocol requires a laboratory capable of performing serum protein electrophoresis, immunofixation, serum FLC measurement, and measurements of IgG, $\kappa$, A, and M. Dissociation of these techniques among various laboratories can cause errors in interpretation.

In conclusion, we have illustrated that antigen excess can cause falsely low serum FLC results with nephelometric techniques. Clinically suspicious results should be repeated after sample dilution.

**References**


Grant/funding support: None declared.
Financial disclosures: None declared.

References

Table 1. Serum FLC measurements according to the dilutions.

<table>
<thead>
<tr>
<th>Dilution</th>
<th>$\kappa$, mg/L</th>
<th>$\lambda$, mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/100</td>
<td>&lt;5.88</td>
<td>&lt;8.09</td>
</tr>
<tr>
<td>1/20</td>
<td>&lt;1.17</td>
<td>&lt;1.62</td>
</tr>
<tr>
<td>1/5</td>
<td>&lt;0.294</td>
<td>&lt;0.405</td>
</tr>
<tr>
<td>1/32 000$^a$</td>
<td>&gt;60 100</td>
<td></td>
</tr>
<tr>
<td>1/2000$^a$</td>
<td></td>
<td>&lt;0.410</td>
</tr>
</tbody>
</table>

Day 0

<table>
<thead>
<tr>
<th>Dilution</th>
<th>$\kappa$, mg/L</th>
<th>$\lambda$, mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/100</td>
<td>112</td>
<td>&lt;8.09</td>
</tr>
<tr>
<td>1/20</td>
<td>&lt;1.62</td>
<td>&lt;0.405</td>
</tr>
<tr>
<td>1/5</td>
<td>&lt;0.405</td>
<td></td>
</tr>
<tr>
<td>1/32 000$^a$</td>
<td>12 600</td>
<td>&lt;0.410</td>
</tr>
</tbody>
</table>

Day 35

<table>
<thead>
<tr>
<th>Dilution</th>
<th>$\kappa$, mg/L</th>
<th>$\lambda$, mg/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/100</td>
<td>1/32 000</td>
<td></td>
</tr>
<tr>
<td>1/2000$^a$</td>
<td></td>
<td>&lt;0.410</td>
</tr>
</tbody>
</table>

$^a$ Serum FLC measurements reported to the medical record after manual dilution.

3 Centre Hospitalier Universitaire Clermont-Ferrand
Department of Rheumatology
G Montpied Hospital
Clermont-Ferrand, France

These authors contributed equally to this work.

* Address correspondence to this author at: Department of Immunology, Hotel-Dieu, Blvd. Leon Malfreyt, F-63058 Clermont-Ferrand, France. Fax 35-4-73-750-637; e-mail bevrard@chu-clermont-ferrand.fr.

DOI: 10.1373/clinchem.2007.093377

The Effect of Freezing, Thawing, and Short- and Long-Term Storage on Serum Thyrotropin, Thyroid Hormones, and Thyroid Autoantibodies: Implications for Analyzing Samples Stored in Serum Banks

To the Editor:

Data concerning the effects of freezing, thawing, and storage of samples for up to 23 years on TSH, fT4, fT3, TPO-Ab, and TG-Ab concentrations in human serum (Architect i2000, Abbott Diagnostics). Whole blood samples for the Finnish Maternity Cohort (FMC) were collected without preservative from pregnant women, most of whom were in the 1st trimester of pregnancy. Each sample was stored in a single serum aliquot in a polypropylene cryovial (Nunc GmbH & Co.) at −25 °C. The ethics boards of the Finnish National Public Health Institute and Oulu University Hospital approved this study.

The effect of freezing and thawing was evaluated by comparing frozen serum samples (n = 50) with fresh samples (n = 50) from FMC. Samples stored for 6 months vs 2–23 years at −25 °C (50 different samples collected from FMC at every time point, total n = 645) were analyzed to evaluate the effect of storage time. Short-term storage of up to 6 days at 4 °C was studied in fresh sera (n = 8) from nonpregnant women. Details of the methods are available on request.

After 0, 1, 3, and 6 days at 4 °C, no statistically significant effect was seen on any tests studied. There were no differences in TSH, fT4, TPO-Ab, or TG-Ab concentrations when 50 frozen and thawed serum samples were compared with 50 fresh serum samples. fT3 concentrations were significantly higher (Student t-test, $P < 0.001$) in frozen samples but remained within reference intervals.

We analyzed 645 samples from pregnant women with a mean gestational age of 11.9 weeks (range 6.7–18.0) to evaluate the effect of storage time. Storage time had no effect on TSH or fT3 concentrations (Fig. 1). Concentrations of fT4 remained within reference intervals at all time points, although they differed significantly from baseline after 10, 12, and 20–23 years of storage (Fig. 1). Because fT4 concentrations were comparable between 0 and 14–18 years, the difference at 10 and 12 years might be the result of random variation. Storage time explained only 5.5% of the total variation in...
fT4 concentrations (linear regression analysis, \( P < 0.001 \)), which was smaller than the interassay variation of the method.

The concentrations of TPO-Ab increased steadily with extended storage time, and a significant difference compared with baseline was seen when storage time was >2 years (Dunnett test, \( P = 0.047 \)). Thereafter, TPO-Ab concentrations steadily increased with increasing storage time (Dunnett test, \( P < 0.001 \)) until 14 years, at which point a more marked increase was seen (Dunnett test, \( P < 0.001 \); Fig. 1). Storage time explained 19.7% of the total variation in TPO-Ab concentrations (linear regression analysis, \( P < 0.001 \)). Although median TPO-Ab concentrations stayed under the upper limit of the reference interval (5.61 kIU/L), considerable change was seen after 14 years of storage (Fig. 1).

The concentrations of TG-Ab increased with extended storage time, differing significantly from the baseline after storage for >6 years (Dunnett test, \( P < 0.05 \)), with the exception of the 10th year of storage. Storage time explained 19.4% of the total variation seen in TG-Ab concentrations (linear regression analysis, \( P < 0.001 \)). Median TG-Ab concentrations exceeded the upper limits of the reference interval (4.11 kIU/L) after 14 years of storage (Fig. 1).

Higher thyroid autoantibody concentrations were likely to be the result of storage time, because preanalytic conditions (sampling, tubes, transportation, sample processing, freezing, thawing, or storage conditions) did not differ in our study. The causes of changes in thyroid autoantibody instability during long-term storage are unknown and probably of complex character.

We conclude that TSH, fT4, and fT3 can reliably be analyzed in samples stored for 23 years at \(-25^\circ C\), and that TPO-Ab and TG-Ab are also stable for 14 years of storage. Cross-sectional comparison of samples is possible at every time point, but comparison between samples with different storage times, at least when exceeding 14 years, should be carried out cautiously.

Grant/funding support: This study was supported in part by grants from Alma and K.A. Snellman Foundation, Oulu, Finland; the Jalmari and Rauha Ahokas Foundation, Finland; and the Lilly Foundation, Finland.

Financial disclosures: None declared.

Acknowledgments: We thank Jouni Salminen and Frank Quinn from Abbott Laboratories for providing laboratory kits.

References

Tuija Männistö1,2*, Heljä-Marja Surcel3, Aini Bloigu3, Aimo Ruokonen4, Anna-Liisa Hartikainen1, Marjo-Riitta Järvelin2,5, Anneli Pouta3, Marja Vääräsmäki1, Eila Suvanto-Luukkonen1

Departments of
1 Obstetrics and Gynecology and Public Health Science and General Practice University of Oulu Oulu, Finland
2 National Public Health Institute Oulu, Finland
3 Department of Clinical Chemistry University of Oulu Oulu, Finland
4 Department of Epidemiology and Public Health Imperial College London London, United Kingdom

* Address correspondence to this author at: Department of Obstetrics and Gynecology, University of Oulu, PO Box 5000, 90014 Oulu, Finland. Fax 358-8-3154310; e-mail Tuija.Mannisto@oulu.fi.

DOI: 10.1373/clinchem.2007.091371
The Combination of Cystatin C and Serum Creatinine Improves the Monitoring of Kidney Function in Patients with Diabetes and Chronic Kidney Disease

To the Editor:

Pucci et al. (1) recently reported their study of the use of cystatin C (cysC) to detect decreases in renal function in patients with diabetes. CysC had an advantage over other methods only for detecting very early impairment of renal function [glomerular filtration rate (GFR) < 90 and 75 mL·min⁻¹·(1.73 m²)⁻¹], whereas detection of the 70 patients with GFR < 60 studied by Pucci et al. (1) was not improved by use of cysC compared with conventional estimated GFR estimations (e-GFR) predicted by Cockcroft and Gault (CG) and Modification of Diet in Renal Disease (MDRD) equations. Better GFR estimations are required for those patients considered to have chronic kidney disease (CKD) according to the new American Diabetes Association recommendations, particularly to assess the progression of their CKD, which may not be accurately estimated by the CG and MDRD equations (2). Rule et al. (3) recently proposed a composite GFR estimation based on both serum creatinine (sCr) and cysC, highly correlated with the results of 204 iothalamate clearance determinations in patients with diseases of the native (non-transplanted) kidney. We wondered whether this estimation could be useful in the assessment and monitoring of kidney function in patients with diabetes and CKD.

Seventy-six patients [43 males, 33 females, mean (SD) age 65 (12) years, body mass index 27.6 (4.9) kg/m²] gave informed consent to participate in the study, which was performed in accordance with the declaration of Helsinki. The inclusion criteria were diabetes [19 type 1 and 57 type 2, hemoglobin A1C 8.1 (1.4%)] and CKD [MDRD<60 mL·min⁻¹·(1.73 m²)⁻¹, sCr 182 (84) μmol/L]. GFR was measured by ⁵¹Cr-EDTA clearance [i-GFR, 33.2 (14.8) mL·min⁻¹·(1.73 m²)⁻¹], and compared to the recommended and new composite estimations by use of correlations, paired t tests, Bland-Altman procedures, and absolute percentage difference with i-GFR. sCr was quantified on an Olympus AU 640 analyzer by use of the Jaffé method with bichromatic measurements according to the manufacturer’s specifications. The MDRD estimation was calculated by the abbreviated 4-variable version of the equation. The composite estimation was based on both sCr and cysC (3):

\[\sqrt{(66.8 \times \text{cysC}^{-1.10} \times (273 \times \text{sCr}^{-1.22}) \times \text{age}^{-0.299} \times 0.738 \text{ (if female)}}].\]

The relationships between i-GFR and 1/sCr, 1/cysC, and both were tested by linear regression. The diagnostic accuracy of each estimation for the diagnosis of severe renal failure (i-GFR<30) was tested by ROC curve analysis. Sixteen patients had a 2nd i-GFR [31.0 (18.0) mL·min⁻¹·(1.73 m²)⁻¹ vs 36.1 (14.0) at baseline, not significant (NS)], and an e-GFR determination 2 years later to test whether e-GFR changes correlated to i-GFR changes [−5.1 (17.5) mL·min⁻¹·(1.73 m²)⁻¹].

The initial mean (SD) cysC was 1.97 (0.83) mg/L. i-GFR correlated to 1/sCr (r = 0.65, P < 0.001) and 1/cysC (r = 0.64, P < 0.001). The statistical significance was maintained when both were included in a multiple linear regression model (r = 0.69, P < 0.001; P = 0.002 for 1/sCr and P = 0.009 for 1/cysC). Areas under the ROC curves were 0.874 for cysC and 0.870 for sCr (NS). The results of the GFR estimations are presented in Table 1.

As already reported in patients with diabetes (4), the performances of the CG equation (correlation with i-GFR, accuracy, area under the ROC curve, Bland-Altman procedure) were poor. In accordance with the findings of Pucci et al. (1), the MDRD remains the best estimation of GFR in renally insufficient diabetic patients: the best estimates for correlation coefficient, absolute percentage difference, and ROC curve for the diagnosis of severe renal failure were all obtained with this equation. However, 1/cysC still predicted i-GFR after taking account of 1/sCr. The performances of the composite estimations of GFR, which

<table>
<thead>
<tr>
<th>Table 1. The performances of the CG, MDRD, and composite e-GFR in the 76 patients with diabetes and CKD.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CG</strong></td>
</tr>
<tr>
<td>e-GFR [mL·min⁻¹·(1.73 m²)⁻¹]</td>
</tr>
<tr>
<td>r/i-GFR (correlation)</td>
</tr>
<tr>
<td>P vs i-GFR (paired t test)</td>
</tr>
<tr>
<td>Absolute percentage difference, %</td>
</tr>
<tr>
<td>Bland-Altman</td>
</tr>
<tr>
<td>2 (SD) [mL·min⁻¹·(1.73 m²)⁻¹]</td>
</tr>
<tr>
<td>r</td>
</tr>
<tr>
<td>P</td>
</tr>
<tr>
<td>AUC ROC/GFR&lt;30</td>
</tr>
<tr>
<td>GFR change at 2 years [mL·min⁻¹·(1.73 m²)⁻¹]</td>
</tr>
<tr>
<td>r vs i-GFR change</td>
</tr>
<tr>
<td>P</td>
</tr>
</tbody>
</table>

a, P < 0.05 vs all other values.
included both markers, were close to the MDRD equation. These similar performances contrast with the better correlation coefficient for the composite estimation ($r^2 = 0.891$) than the MDRD ($r^2 = 0.825$) in the report by Rule et al. (3), but this outcome is not unexpected because the lower i-GFR for our patients [33.2 (14.8) mL/min (1.73 m$^2$)] vs 57 (29) for the patients studied by Rule et al.] decreased the impact of the well-known underestimation of high GFR by the MDRD. Most interestingly, the correlation between the composite estimation change and the i-GFR change almost reached significance, whereas the MDRD did not. Even for advanced CKD patients, the inclusion of cysC in the composite estimation may have diagnostic value, for monitoring GFR decline, a finding that is in line with recent reports for the use of $100/cysC$ (5).

Grant/funding support: None declared. Financial disclosures: None declared.

References