UCSF
UC San Francisco Previously Published Works

Title

Laboratory reproducibility of biochemical markers of bone turnover in clinical practice.

Permalink

Journal

ISSN

Authors

Schafer, AL
Vittinghoff, E
Ramachandran, R
et al.

Publication Date

DOI

Copyright Information

Peer reviewed

ORIGINAL ARTICLE

Laboratory reproducibility of biochemical markers of bone turnover in clinical practice

A. L. Schafer · E. Vittinghoff · R. Ramachandran ·
N. Mahmoudi · D. C. Bauer

Abstract

Summary To determine the laboratory reproducibility of
urine *N*-telopeptide and serum bone-specific alkaline
phosphatase measurements, we sent identical specimens to
six US commercial labs over an 8-month period. Longitu-
dinal and within-run laboratory reproducibility varied
substantially. Efforts to improve the reproducibility of these
tests are needed.

Introduction We assessed the laboratory reproducibility of
urine *N*-telopeptide (NTX) and serum bone-specific alka-
line phosphatase (BAP).

Methods Serum and urine were collected from five
postmenopausal women, pooled, divided into identical
aliquots, and frozen. To evaluate longitudinal reproduc-
ibility, identical specimens were sent to six US commer-
cial labs on five dates over an 8-month period. To

evaluate within-run reproducibility, on the fifth date, each
lab was sent five identical specimens. Labs were unaware
of the investigation.

Results Longitudinal coefficients of variation (CVs) ranged
from 5.4% to 37.6% for NTX and from 3.1% to 23.6% for
BAP. Within-run CVs ranged from 1.5% to 17.2% for
NTX. Compared to the Osteomark NTX assay, the Vitros
ECi NTX assay had significantly higher longitudinal
reproducibility (mean CV 7.2% vs. 30.3%, $p < 0.0005$) and
within-run reproducibility (mean CV 3.5% vs. 12.7%, $p <
0.0005$).

Conclusions Reproducibility of urine NTX and serum BAP
varies substantially across US labs.

Keywords Bone-specific alkaline phosphatase · Bone
turnover markers · Laboratory assay · Reproducibility ·
Urine *N*-telopeptide

A. L. Schafer · D. C. Bauer
Department of Medicine, University of California,
San Francisco, CA, USA

E. Vittinghoff · D. C. Bauer
Department of Epidemiology and Biostatistics,
University of California,
San Francisco, CA, USA

R. Ramachandran
Department of Pathology, University of California,
San Francisco, CA, USA

N. Mahmoudi
Department of Medicine, Kaiser Permanente,
Oakland, California, USA

A. L. Schafer (✉)
2200 Post Street, Room C-409,
San Francisco, CA 94115, USA
e-mail: anne.schafer@ucsf.edu

Introduction

Recent investigation has shown that biochemical markers of
bone turnover, both markers of bone resorption and markers
of bone formation, can confirm a biochemical response to
treatment of osteoporosis with antiresorptive agents [1], and
early changes in these markers can predict long-term
changes in bone mineral density [2]. Further, changes in
markers are associated with fracture risk [3–5].

Although these findings have secured a place for the use
of bone turnover markers in research trials, markers still are
not used frequently in clinical practice. Use in the diagnosis
and treatment of individual patients has largely been limited
by cost, by the data supporting marker significance, and by
variability, both pre-analytical and analytical. Pre-analytical

variability includes biological variability, which comprises that from circadian rhythms, diet, age, and gender [6], as well as that due to sample handling and storage. Analytical variability, in contrast, is that which originates from the laboratory measurements themselves. While laboratory assays are studied rigorously in standardized settings, data are lacking about the reproducibility of bone turnover marker measurements in actual clinical practice. The data that do exist raise concerns: a European investigation involving interlaboratory variation found that results for most biochemical markers of bone turnover differed markedly among laboratories [7]. In the USA, laboratory standards are determined by the Clinical Laboratory Improvement Amendments and assessed by proficiency-testing providers such as the College of American Pathologists, but the results of cross-laboratory proficiency testing are not routinely available to clinicians.

The evaluation of laboratory reproducibility in clinical practice is especially important as laboratory assays evolve. For some markers, manual enzyme-linked immunosorbant assays (ELISAs) are being replaced by assays using the same monoclonal antibodies but run on automated platforms. Different laboratories may use distinct assays on clinical specimens.

This study aimed to determine the laboratory reproducibility of two biochemical markers of bone turnover: urine cross-linked N-telopeptide of type I collagen (NTX), a marker of bone resorption, and serum bone-specific alkaline phosphatase (BAP), a marker of bone formation.

Methods

Postmenopausal women older than 55 years of age were recruited with advertising flyers posted around a large academic medical center and in community businesses. Volunteers were excluded if they were using current pharmacologic therapy for osteoporosis, with relevant therapy defined as estrogen, calcitonin, a selective estrogen receptor modulator, a bisphosphonate, or teriparatide; calcium and vitamin D supplements were permitted. All volunteers provided verbal informed consent with the assistance of an information sheet, given the minimal risks involved in participation. The institutional review board of the University of California, San Francisco approved the study protocol prior to initiation of the study.

A pool of serum and a pool of urine were created from specimens from five volunteers, in order to create samples sufficiently large for the investigation and also in order to minimize the interfering effects of medications or other factors specific to a single volunteer. To create the pool of serum, fasting morning blood from the participating women was collected in eight gold-top serum separator tubes,

allowed to clot at room temperature for 30 min, and then placed on ice, centrifuged, and separated. The pooled serum was then stirred for 10 min in an ice water bath, divided into 1.2 mL aliquots, and flash-frozen. To create the pool of urine, fasting second-morning urine from the participating women was collected, placed on ice, pooled, stirred for 10 min in an ice water bath, divided into 4 mL aliquots, and flash-frozen. The serum and urine aliquots were then frozen at −80°C.

Six US laboratories were selected for investigation, each a recognized, high-volume commercial laboratory that offers urine NTX and serum BAP testing: ARUP Laboratories (Salt Lake City, UT, USA), Esoterix Laboratory Services (Calabasas Hills, CA, USA), Laboratory Corporation of America (LabCorp; Burlington, NC, USA), Mayo Medical Laboratories (Rochester, MN, USA), Quest Diagnostics (Nichols Institute, San Juan Capistrano, CA, USA), and Specialty Laboratories (Valencia, CA, USA). To prevent bias, the laboratories were unaware of the investigation; source-masked identifiers were used for all specimens, and the specimens were sent by the authors' institutional clinical laboratory as routine clinical specimens ordered by clinicians would be sent. The laboratories were paid in full via the standard contractual arrangements in place with the authors' clinical laboratory. Each laboratory was sent a serum and a urine specimen on five dates over an 8-month period, in order to assess longitudinal (between-run) variability of the marker measurements. The dates were 6 to 7 weeks apart, with the exception of those sent to Specialty, for which the interval between the first and second dates was 14 weeks. For all laboratories, on the fifth date, five serum and five urine specimens were sent to each laboratory in order to assess within-run variability of the marker measurements.

Each of the six laboratories used one of two assays for urine NTX measurements and one of two assays for serum BAP measurements. For urine NTX, two laboratories (LabCorp and Specialty) used the Osteomark assay (Inverness Medical Innovations, Waltham, MA, USA), an ELISA using a monoclonal antibody directed against a urinary pool of collagen cross-links originally derived from a patient with Paget's disease. Four laboratories (ARUP, Esoterix, Mayo, and Quest) used the Vitros enhanced chemiluminescence (ECi) assay (Ortho-Clinical Diagnostics, Rochester, NY, USA), a fully automated platform using the same antigen. For serum BAP, one laboratory (Specialty) used the Metra BAP enzyme immunoassay (Quidel, San Diego, CA, USA), while five laboratories (ARUP, Quest, Esoterix, Mayo, and LabCorp) used Access Ostase (Beckman Coulter, Fullerton, CA, USA), another enzyme immunoassay. Of note, Metra BAP was formerly called Alkphase-B. Access Ostase was formerly Hybritech Tandem-MP Ostase, which itself was developed from the monoclonal antibody

Table 1 Longitudinal reproducibility of urine NTX

Units for reference ranges, means and SDs: nM BCE/mM Cr

[a] Reference ranges, provided by each laboratory, are for postmenopausal women for ARUP and Esoterix, premenopausal women for Mayo and Quest, and not specified for LabCorp and Specialty

Lab	Assay	Reference range[a]	Mean ± SD	CV, % (95% CI)
ARUP	Vitros ECi	26–124	35.8±1.9	5.4 (3.2–15.5)
Esoterix	Vitros ECi	25–110	35.8±2.9	8.0 (4.5–30.4)
LabCorp	Osteomark	5–65	74.2±19.3	25.9 (15.2–87.9)
Mayo	Vitros ECi	19–63	35.0±3.0	8.6 (5.1–25.0)
Quest	Vitros ECi	4–64	34.0±2.2	6.6 (3.9–19.1)
Specialty	Osteomark	14–74	42.8±16.0	37.6 (21.6–168.0)
Vitros ECi (all)			35.1±2.5	7.2 (5.5–10.6)
Osteomark (all)			58.5±17.7	30.3 (20.4–60.5)

used for the Hybritech Tandem-R Ostase immunoradiometric assay.

The laboratories communicated the results by fax to the authors' institutional clinical laboratory, as is done for routine clinical specimens. Urine NTX values were reported by all labs in whole numbers; BAP values were reported by four of the labs to one tenth of a microgram per liter or unit per liter but by Esoterix and Mayo as whole numbers. Following standard practice, labs corrected urine NTX values for dilution by urinary creatinine analysis and reported results as NTX/creatinine ratios (to be referred to simply as NTX in this paper).

Means, SDs, and coefficients of variation (CVs, defined as mean/SD) with 95% confidence intervals (CIs) were calculated [8]. A CV for within-run reproducibility for BAP could not be computed for Esoterix because the reported values were rounded to the nearest microgram per liter and did not vary. Two sensitivity analyses were performed: first, a uniform random variate on the interval [−0.5, 0.5] was added to the BAP values reported by that lab and by Mayo, which also rounded to the nearest microgram per liter. Then, the perturbed results were rounded to the nearest 0.1 μg/L, as reported by the other labs. Second, CVs were computed after rounding reported values from all six labs to the nearest microgram per liter (or, for Metra, the nearest U/L). Assay-specific CVs were computed for NTX and BAP measurements as the ratio of the average within-lab SDs, obtained from a linear regression of the measurement on laboratory, stratified by assay type, to the overall average of the measurements for that assay; CVs were compared across assays using the methods of Feltz and Miller [9].

Results

The participating postmenopausal women were Caucasian and ranged in age from 57 to 74 years (mean ± SD age 65± 6.3 years).

Longitudinal reproducibility was evaluated by sending one specimen to each lab on each of five dates. For urine

NTX (Table 1, Fig. 1), CVs varied from 5.4% to 37.6%: CVs were 5.4% (95% CI 3.2–15.5) for ARUP, 8.0% (CI 4.5–30.4) for Esoterix, 25.9% (CI 15.2–87.9) for LabCorp, 8.6% (CI 5.1–25.0) for Mayo, 6.6% (CI 3.9–19.1) for Quest, and 37.6% (CI 21.6–168.0) for Specialty. Longitudinal reproducibility was significantly lower for labs using the Osteomark assay (CV 30.3%, CI 20.4–60.5) than for those using the Vitros ECi assay (CV 7.2%, CI 5.5–10.6; $p<0.0005$ for comparison between assays).

For BAP (Table 2, Fig. 2), longitudinal CVs ranged from 3.1% (CI 1.9–9.1) for Esoterix to 23.6% (CI 13.9–77.2) for LabCorp. Analyses using perturbed data, done because some labs' results were in whole numbers and some to one tenth of a microgram per liter or unit per liter, gave similar results. For example, the longitudinal CV for Esoterix, which reported its results as whole numbers, became 4.5% (CI 2.7–13.0) when the values were perturbed by random variables before computations were performed, and the CV
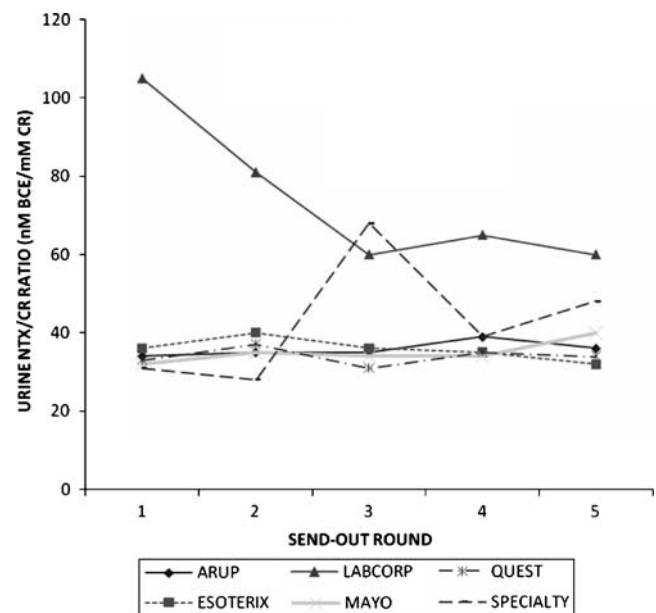

Fig. 1 Urine NTX measurements for the six laboratories. Send-out rounds were of identical specimens and were 6 to 7 weeks apart, with the exception of those sent to Specialty, for which the interval between the first and second dates was 14 weeks

Table 2 Longitudinal reproducibility of serum BAP

Lab	Assay	Reference range[a]	Mean ± SD	CV, % (95% CI)
ARUP	Ostase	7.0–22.4	13.8±1.3	9.3 (5.6–27.3)
Esoterix	Ostase	≤22.4	14.2±0.4	3.1 (1.9–9.1)
LabCorp	Ostase	0.0–21.3	11.4±2.7	23.6 (13.9–77.2)
Mayo	Ostase	≤22	14.4±0.9	6.2 (3.7–18.0)
Quest	Ostase	5.6–29.0	14.4±1.5	10.4 (6.2–30.7)
Specialty	Metra BAP	14.2–42.7	24.0±1.4	5.6 (3.4–16.3)
Ostase (all)			13.6±1.6	11.4 (8.9–16.0)
Metra BAP			24.0±1.4	5.6 (3.4–16.3)

Units for reference ranges, means and SDs: µg/L, except U/L for Metra

[a] Reference ranges, provided by each laboratory, are for postmenopausal women for ARUP, Mayo, and Esoterix, and not specified for Quest, LabCorp, and Specialty

for LabCorp, which reported its results to a tenth of a microgram per liter, became 24.3% (CI 14.3–80.2) when the values were rounded to whole numbers before computations were performed.

Within-run reproducibility was evaluated as each lab was sent five identical specimens on one date. For urine NTX (Table 3), CVs ranged from 1.5% (CI 0.9–4.3) for ARUP to 17.2% (CI 10.2–52.9) for Specialty. A comparison of assays revealed a statistically significant difference, with within-run CVs 12.7% (CI 8.7–23.5) for the Osteomark assay and 3.5% (CI 2.6–5.1) for the Vitros ECi assay ($p<$ 0.0005 for comparison between assays).
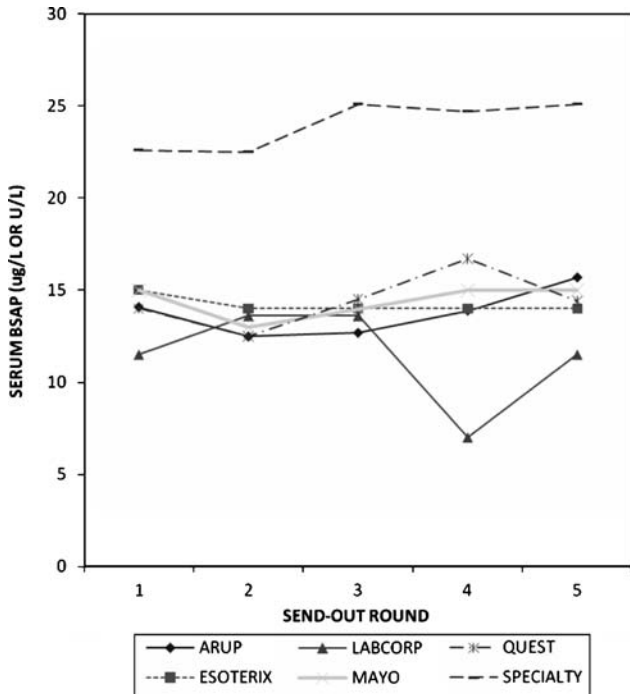

Fig. 2 Serum BAP measurements for the six laboratories. Measurements of BAP by the Metra assay, used by Specialty Labs, are in units per liter, while measurements by the Ostase assay, used by the other five laboratories, are in micrograms per liter. Send-out rounds were of identical specimens and were 6 to 7 weeks apart

For BAP (Table 4), Esoterix produced five identical measurements, and within-run CVs for the other labs ranged from 2.2% (CI 1.3–6.3) for Quest to 15.5% (CI 9.2–47.1) for LabCorp. Analyses using perturbed data, done because some labs' results were in whole numbers and some to one tenth of a microgram per liter or unit per liter, gave similar results. For example, the longitudinal CV for Quest, which reported its results to a tenth of a microgram per liter, became 3.8% (CI 2.3–11.0) when the values were rounded to whole numbers before computations were performed, and the CV for LabCorp, which also reported its results to a tenth of a microgram per liter, became 15.1% (CI 9.0–45.5). The CV for Mayo, which reported its results as whole numbers, was 8.3% (CI 5.0–24.2) using the values reported and became 9.3% (CI 5.3–27.3) when the values were perturbed by random variables before computations were performed. Of the five identical serum specimens sent on one date to LabCorp, one was not processed, with the reason cited "quantity not sufficient."

In addition to means, SDs, and CVs for the NTX/creatinine ratio (referred to simply as NTX in this paper), computations were also done for NTX itself (uncorrected) and for urine creatinine alone. CVs obtained for NTX itself (uncorrected) appeared similar to those for the ratio (data not shown).

Discussion

Despite their use in research trials, biochemical markers of bone turnover still are not used frequently in clinical practice, in part due to concerns about analytical variability. In this masked study of identical specimens, the reproducibility of urine NTX and serum BAP was highly variable at US commercial labs. On the one hand, several labs were quite precise in their results longitudinally (between runs separated in time) and within a given run: for example, Esoterix produced five identical measurements for serum BAP within one run. On the other hand, other labs were imprecise: for example, LabCorp's CVs were greater than

Table 3 Within-run reproducibility of urine NTX

Units for reference ranges, means and SDs: nM BCE/mM Cr

[a] Reference ranges, provided by each laboratory, are for postmenopausal women for ARUP and Esoterix, premenopausal women for Quest and Mayo, and not specified for LabCorp and Specialty

Lab	Assay	Reference range[a]	Mean ± SD	CV, % (95% CI)
ARUP	Vitros ECi	26–124	36.4±0.5	1.5 (0.9–4.3)
Esoterix	Vitros ECi	25–110	34.0±1.4	4.2 (2.5–12.0)
LabCorp	Osteomark	19–63	59.0±4.2	7.1 (4.2–20.6)
Mayo	Vitros ECi	4–64	40.0±1.6	4.0 (2.4–11.4)
Quest	Vitros ECi	5–65	34.0±1.2	3.6 (2.2–10.4)
Specialty	Osteomark	14–74	52.8±9.1	17.2 (10.2–52.9)
Vitros ECi (all)			36.1±1.3	3.5 (2.6–5.1)
Osteomark (all)			55.9±7.1	12.7 (8.7–23.5)

20% for longitudinal specimens for both urine NTX and serum BAP, with the lower ends of its 95% CIs greater than 13%, and its CV for within-run BAP measurements was 15.5% (CI 9.2–47.1).

Of important note is the difference in reproducibility of urine NTX measurements when labs using the Osteomark assay (Wampole Laboratories), an ELISA, are compared to those using the Vitros ECi assay (Ortho-Clinical Diagnostics), a fully automated chemiluminescence test. When longitudinal and within-run reproducibility data were compared in this study, the collective CVs for the Vitros ECi assay were significantly lower than the collective CVs for the Osteomark assay. This finding is consistent with the findings of other studies comparing automated and manual assays, such as an examination of urinary free deoxypyridinoline assays that showed the precision of the automated techniques studied was superior to that of the manual immunoassays studied [10].

In fact, one interpretation of the significance of the present study is not the overall inconsistent reproducibility of urine NTX and serum BAP but rather the marked relative success of the newer, automated assays in minimizing analytical variability. A limitation of the present study is the small number of labs evaluated, as a larger number using each type of assay would help support this interpretation; the labs evaluated, though, represent high-volume, well-known commercial labs collectively responsible for a significant proportion of the urine NTX and serum BAP assays conducted in the USA.

Another limitation of the present study is the testing of a single pooled sample for each marker, rather than the testing of multiple pooled samples representing high, normal, and low marker values. However, it is likely that the reproducibility of measurements at the extremes of or outside the normal range would show even greater variability. As each lab determines its own reference ranges, reference ranges varied, but this should not affect measurement reproducibility. In addition, the assay used or the reference range cited by each lab may have changed after the completion of this study.

Clinical laboratories evaluate the quality of their results through proficiency testing, which is required by the Clinical Laboratory Improvement Amendments and performed by organizations including the College of American Pathologists, but survey results are not easily available to practicing clinicians. These and other evaluations of marker assays, such as one conducted as a part of a Centers for

Table 4 Within-run reproducibility of serum BAP

Lab	Assay	Reference range[a]	Mean ± SD	CV, % (95% CI)
ARUP	Ostase	7.0–22.4	15.6±0.6	3.8 (2.3–11.1)
Esoterix	Ostase	≤22.4	14.0±0.0	0 (0–0)
LabCorp[b]	Ostase	0.0–21.3	11.3±1.8	15.5 (9.2–47.1)
Mayo	Ostase	≤22	13.2±1.1	8.3 (5.0–24.2)
Quest	Ostase	5.6–29.0	14.2±0.3	2.2 (1.3–6.3)
Specialty	Metra BAP	14.2–42.7	25.8±0.9	3.5 (2.1–10.1)
Ostase (all)			13.8±0.9	6.6 (5.2–9.3)
Metra BAP			25.8±0.9	3.5 (2.1–10.1)

Units for reference ranges, means and SDs: μg/L, except U/L for Metra

[a] Reference ranges, provided by each laboratory, are for postmenopausal women for ARUP, Mayo, and Esoterix, and not specified for Quest, LabCorp, and Specialty

[b] Of the five identical serum specimens sent on one date to LabCorp, one was not processed, cited as "quantity not sufficient"

Disease Control study to develop a reference system to standardize the measurements of bone resorption markers pyridinium crosslinks pyridinoline and deoxypyridinoline [11], invite labs to participate and announce the tested specimens. While the results provide valuable information, the concern exists that reproducibility may be at its best during an announced test. The present study is important in that the serum and urine specimens submitted to the six high-volume US clinical labs investigated were processed as routine clinical specimens ordered by clinicians would be processed: the labs were unaware of the investigation, fictional identifiers were used, and the specimens were sent by the authors' institutional clinical laboratory, so the specimens were indistinguishable from routine clinical specimens. This element of the study's design was considered extremely important, even though it prevented the direct observation of potential factors that might have explained some of the variability in lab reproducibility, such as the handling of specimens by different labs. In the past, some published studies comparing laboratory performance have published data without naming the laboratories [12, 13], but reaction in the literature has included the belief that the laboratories should be identified [14]; the present study provides laboratories' names in order that the results and discussion generated be as useful as possible to clinicians. The identification of laboratories by name is similar to the identification of commercial assays by name when such assays are compared, and this is not uncommon in the literature [15–17].

Inconsistent reproducibility is a barrier to the use of biochemical markers of bone turnover in clinical practice, particularly if clinicians do not consistently use the same assay and laboratory. The challenge of consistent use is heightened by the fact that many institutional labs send specimens out to higher-volume "send-out" labs (including and especially those investigated here), and clinicians may not be aware to which lab a specimen is being sent. Further, and perhaps more importantly, information about the particular assay used by a given lab is often difficult to find: the type of assay (for example, "chemiluminescent immunoassay") is often listed in a lab's on-line catalog, but none of the faxed reports of urine NTX results identified whether the Vitros ECi or Osteomark assay had been used. Of the faxed reports of serum BAP results, only the Esoterix and LabCorp reports indicated the assay employed, and even then, LabCorp referred to an outdated form of the Ostase test.

The findings of the present study support the call for urgent improvement in analytical precision for these two biochemical markers of bone turnover. Laboratory performance data should be made widely available to clinicians, institutions, and payers, and proficiency testing and standardized guidelines should be strengthened to improve marker reproducibility at those labs currently performing poorly.

Conflicts of interest Dr. Bauer is a consultant for Tethys Bioscience and Roche Diagnostics. The other authors declare that they have no conflicts of interest or disclosures.

References

1. Garnero P, Shih WJ, Gineyts E, Karpf DB, Delmas PD (1994) Comparison of new biochemical markers of bone turnover in late postmenopausal osteoporotic women in response to alendronate treatment. J Clin Endocrinol Metab 79:1693–1700
2. Ravn P, Hosking D, Thompson D, Cizza G, Wasnich RD, McClung M, Yates AJ, Bjarnason NH, Christiansen C (1999) Monitoring of alendronate treatment and prediction of effect on bone mass by biochemical markers in the early postmenopausal intervention cohort study. J Clin Endocrinol Metab 84:2363–2368
3. Eastell R, Barton I, Hannon RA, Chines A, Garnero P, Delmas PD (2003) Relationship of early changes in bone resorption to the reduction in fracture risk with risedronate. J Bone Miner Res 18:1051–1056
4. Reginster JY, Sarkar S, Zegels B, Henrotin Y, Bruyere O, Agnusdei D, Collette J (2004) Reduction in PINP, a marker of bone metabolism, with raloxifene treatment and its relationship with vertebral fracture risk. Bone 34:344–351
5. Bauer DC, Black DM, Garnero P, Hochberg M, Ott S, Orloff J, Thompson DE, Ewing SK, Delmas PD, Fracture Intervention Trial Study Group (2004) Changes in bone turnover and hip, non-spine, and vertebral fracture in alendronate-treated women: the fracture intervention trial. J Bone Miner Res 19:1250–1258
6. Fraser WD, Anderson M, Chesters C, Durham B, Ahmad AM, Chattington P, Vora J, Squire CR, Diver MJ (2001) Circadian rhythm studies of serum bone resorption markers: implications for optimal sample timing and clinical utility. In: Eastell R, Baumann M, Hoyle NR, Wieczorek L (eds) Bone markers: biochemical and clinical perspectives. Martin Dunitz, London, pp 107–118
7. Seibel MJ, Lang M, Geilenkeuser WJ (2001) Interlaboratory variation of biochemical markers of bone turnover. Clin Chem 47:1443–1450
8. Vangel MG (1996) Confidence intervals for a normal coefficient of variation. Am Stat 50:21–26
9. Feltz CJ, Miller GE (1996) An asymptotic test for the equality of coefficients of variation from k populations. Stat Med 15:647–658
10. Seibel MJ, Woitge HW, Farahmand I, Oberwittler H, Ziegler R (1998) Automated and manual assays for urinary crosslinks of collagen: which assay to use? Exp Clin Endocrinol Diabetes 106:143–148

11. Vesper HW, Smith SJ, Audain C, Myers GL (2001) Comparison study of urinary pyridinoline and deoxypyridinoline measurements in 13 US laboratories. Clin Chem 47:2029–2031

12. Binkley N, Krueger D, Cowgill CS, Plum L, Lake E, Hansen KE, DeLuca HF, Drezner MK (2004) Assay variation confounds the diagnosis of hypovitaminosis D: a call for standardization. J Clin Endocrinol Metab 89:3152–3157

13. Binkley N, Krueger D, Gemar D, Drezner MK (2008) Correlation among 25-hydroxy-vitamin D assays. J Clin Endocrinol Metab 93:1804–1808

14. Hollis BW (2004) The determination of circulating 25-hydroxyvitamin D: no easy task. J Clin Endocrinol Metab 89:3149–3151

15. Tortajada-Genaro LA, Cózar MP, Frigols JL, de Avila CR (2007) Comparison of immunoradiometric assays for determination of thyroglobulin: a validation study. J Clin Lab Anal 21:147–153

16. Holvoet P, Macy E, Landeloos M, Jones D, Jenny NS, Van de Werf F, Tracy RP (2006) Analytical performance and diagnostic accuracy of immunometric assays for the measurement of circulating oxidized LDL. Clin Chem 52:760–764

17. Lee JS, Ettinger B, Stanczyk FZ, Vittinghoff E, Hanes V, Cauley JA, Chandler W, Settlage J, Beattie MS, Folkerd E, Dowsett M, Grady D, Cummings SR (2006) Comparison of methods to measure low serum estradiol levels in postmenopausal women. J Clin Endocrinol Metab 91:3791–3797