Patient pools and the use of "patient means" are valuable tools in quality control illustrated by a bone-specific alkaline phosphatase assay

Hinge, Maja; Lund, Erik D.; Brandslund, Ivan; Plesner, Torben; Madsen, Jonna S

Published in:
Clinical Chemistry and Laboratory Medicine

DOI:
10.1515/cclm-2015-0308

Publication date:
2016

Document version
Final published version

Maja Hinge, Erik D. Lund, Ivan Brandslund, Torben Plesner and Jonna S. Madsen*

Patient pools and the use of "patient means" are valuable tools in quality control illustrated by a bone-specific alkaline phosphatase assay

DOI 10.1515/cclm-2015-0308
Received March 31, 2015; accepted July 8, 2015; previously published online August 11, 2015

Abstract

Background: Quality control (QC) is an essential part of clinical biochemistry to ensure that laboratory test results are reliable and correct. Those tests without a defined reference method constitute a special challenge, as is the case with bone-specific alkaline phosphatase (BAP).
Methods and Results: The present study reports an example where a shift in a BAP assay was detected by use of a patient pool and supported by a retrospective calculation of "patient mean", while the external QC and specific assay control material were unaffected by the shift.
Conclusions: Patient pools and the use of patient means remain a useful and inexpensive procedure for internal QC.

Keywords: analytical bias; bone-specific alkaline phosphatase; patient means; patient pool; quality control.

Introduction

For the clinical use of blood tests, it is essential that results are reliable and accurate, which is why quality control (QC) has become the cornerstone of clinical biochemistry. QC must be able to detect analytical errors that can potentially lead to harm in patients [1–3].

Bone-specific alkaline phosphatase (BAP) is a bone formation marker frequently used in clinical trials of diseases that affect the bone, such as multiple myeloma and osteoporosis. However, bone markers are still not a part of daily clinical practice. One reason for this might be explained by the many different challenges concerning the analysis of serum BAP, including: no established reference method, a lack of available primary calibration material [4] and a high degree of inter-method variation [5].

We recently discovered an analytical error due to a shift in an assay of serum BAP that was not detected by either the control material included in the assay or the external QC system, but was brought to our attention by the use of a patient pool (PP). This demonstrates that analytical errors can sometimes be difficult to detect and highlights the importance of a multi-tiered QC system.

Materials and methods

Samples of a PP are routinely used as part of the internal QC for a number of biochemical tests at the Department of Immunology and Biochemistry, Lillebaelt Hospital, Denmark. During the analysis of BAP, samples from the PP were used as part of the QC, as no independent control material was available. During autumn 2013, increasing values of serum BAP were noticed while analyzing the PP. Therefore, we decided to review the results, beginning retrospectively from September 2009 and up to June 2014. This review supported the suggestion of a shift in the results of the assay. In order to investigate this further, the "patient mean" for each month was calculated over the same period.

Serum BAP was analyzed during the whole period with an immunoassay measuring enzyme activity (MicroVue BAP EIA, Quidel Corporation, San Diego, CA, USA). There was no change in the instructions for use of the assays from the manufacturer during the period. Eleven different lot numbers were used during the period. QC material was provided by the supplier of the assay and UK NEQAS was used as an external control throughout. In addition, tests were also exchanged with a Danish university hospital laboratory, from August 2013, using the same BAP assay.

The same PP was used throughout the study period. The PP was produced on 26 May, 2009 as a pool of residual serum from 125 to 150 blood samples analyzed in the laboratory. After mixing, the material

***Corresponding author: Jonna S. Madsen,** Department of Clinical Immunology and Biochemistry, Lillebaelt Hospital, Kabbeltoft 25, 7100 Vejle, Denmark; and Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark, Phone: +45 79406501; E-mail: Jonna.Skov.Madsen@rsyd.dk
Maja Hinge and Torben Plesner: Department of Internal Medicine, Section of Hematology, Lillebaelt Hospital, Vejle, Denmark; and Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark
Erik D. Lund: Department of Clinical Immunology and Biochemistry, Lillebaelt Hospital, Vejle, Denmark
Ivan Brandslund: Department of Clinical Immunology and Biochemistry, Lillebaelt Hospital, Vejle, Denmark; and Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

was aliquoted and stored at –80 °C. Every time the serum BAP test was run, a double determination of the PP was made. Once during the study period we adjusted the limits of the PP. The mean values of serum BAP of the PP were calculated for each month. The monthly patient median was calculated for illustrative purposes as shown in Figure 1.

The standard pre-analytical handling of the test material in the laboratory remained unchanged throughout the period.

All analyses performed on QC material provided by the supplier successfully passed the control test (Figure 2). The external QC, carried out by participation in the assessment program from UK NEQAS, also showed acceptable results throughout the study period. The samples that were exchanged with another Danish laboratory using the same assay also recognized the dramatic decrease in the BAP levels in autumn 2013.

Results

From September 2009 until August 2010 the PP and patient mean remained stable and provided a baseline level. However, from September 2010 to October 2011 there was a slow but steady increase in the results from analysis of the PP, after which they stabilized to a level approximately 25% above the baseline level. In the subsequent period from October 2011 to June 2012, the stability was maintained, but then a continuous increase in results from the PP, as well as patient mean was seen until August 2013 where it reached a level of 100% above the baseline level of 2009.

In December 2013, the level suddenly dropped to the baseline found in 2009 as a 50% decrease in the level of the PP was observed. This level was maintained until end of the observation period (Figure 1).

The mean number of serum BAP patient tests performed was 75 (10–210) per month, which fell to a mean of 26 (10–72) tests per month between February 2013 and June 2014.

Discussion

PP prepared of human serum has long been known as a useful and inexpensive control material [6].

For the analysis of BAP we used the PP as QC, in addition to the control material supplied by the manufacturer as no third party control material was available [7].

However, stability of the measured biochemical marker is essential for the use of PP as QC. The exact long-term stability of BAP in our pool is not known, but maintained stability of biochemical markers in serum pools stored at –80 °C for more than five years has been reported, however, BAP was not included in this study [8]. Others have found BAP to be stable for one year [9], while a study investigating the stability of several biochemical markers including alkaline phosphatase demonstrated that most of the analyses showed variability over time, especially

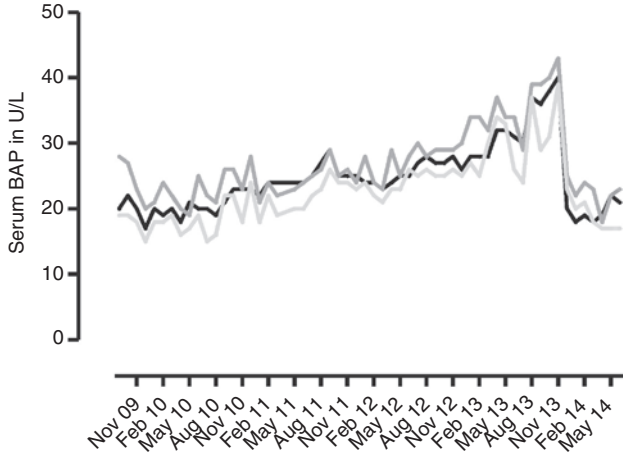

Figure 1: Bone-specific alkaline phosphatase (BAP) patient values in patient pool, "patient mean" and "patient median".
BAP levels measured in the PP (black), the calculated "patient mean" (gray) and the calculated "patient median" (light gray). The mean of the PP, the "patient mean" and the "patient median" were calculated monthly. Note: the steady increase from September 2010 and the marked decrease in December 2013.

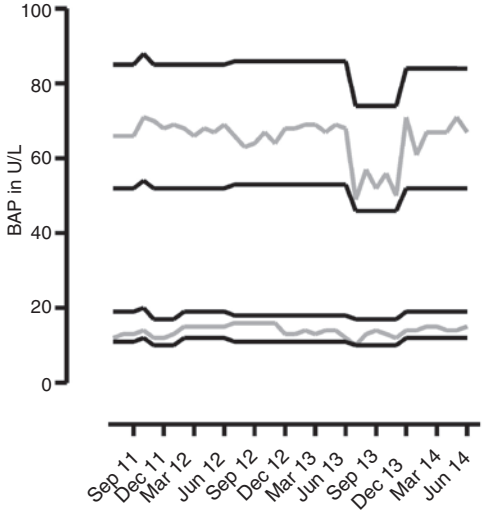

Figure 2: Results of the quality control material provided by the supplier.
Test results from the high (gray) and low (gray) control material was within their respective control limits (black), and did not indicate any problems with the assay. The quality control material from the supplier is specific for each lot number and therefore do the control limits vary during the period.

when new lot numbers were introduced. Nonetheless, the PP remained stable for the majority of parameters measured during the 13-month study period [10]. These studies support the long-term stability of PP and we do not believe that the shift in the results found in our measurements of serum BAP is due to instability of our PP. This observation is also supported by the results of the patients mean which was calculated retrospectively.

Various methods of using patient data in laboratory QC have been described [11–13]. In a study testing the use of patient means, medians and "average of normals" as assessments for long-term analytical stability, the authors demonstrated how the patient results of ammonia increased and revealed an inaccurate lot of the assay, while the normal control material consisting of an ammonium sulphate solution remained stable [14]. Our observations with BAP illustrate the same problem. We simply calculated the patient mean, which reinforce our suspicion of a shift in the BAP assay. The use of patient mean as a QC tool has its limitations as it is only sensitive to systematic errors and not random errors. The patient mean will also be affected by changes in the patient populations as well as being sensitive to outliers especially, when working with limited test numbers. Even so, this method offers some advantage. By providing information about change, both in analytical and pre-analytical performance, the method can be used to detected instrument malfunction or assay error. Furthermore, it is a relatively inexpensive method and advances in computer software have resulted in reduced work load. The use of patient mean is also independent of any control material. However, given that BAP is not a frequently used test, our calculation of patient mean relies on relatively few results which exposed a vulnerability to outliers with extremely high or low values. Calculating an average of normals to exclude extreme outliers using the median or a "moving mean", could have minimized that problem and might therefore be the preferred QC tool [12, 13, 15]. In line with this a recent publication stresses the usefulness of using patient data for monitoring analytical stability by calculations of the monthly medians of patient results [16]. Awareness of change in clinical practice for ordering a test or other change in the case-mix of patients may be important when using patient means as QC. Even though the number of measured serum BAP tests decreased toward the end of our study period, the patient case-mix remained unchanged. As such changes did not occur in our study we concluded that the calculation of the patient mean provided further evidence of an analytical error in the serum BAP assay.

One explanation for the failure of the UK NEQAS external quality assessment program to detect the problem may be that our results were only compared with results from other laboratories that used assays from the same manufacturer. If the measurements from this assay had shifted over time toward increasingly higher levels, as we suspected, all other participating laboratories using this particular assay should also have had a shift in their results over time. However, tests performed simultaneously in different laboratories all showed results of a consistently false high level. Consequently, the external control that was used only provides information about a laboratory's handling of the analysis. As the used levels of BAP in the control tests are random and not fixed, it was not possible to identify a shift over time.

Within the external quality assessment program, there was no direct comparison made with the results from laboratories that used other methodologies as other units were used. In a recent published study, different serum BAP assays were compared and the Quidel MicroVue assay achieved the highest equivalent values [5].

The fact that the control material supplied with the assay did not detect the problem is of concern. One likely explanation for this failure could be that the production of control material might not have been totally independent of the calibrator and the assay. We also speculate that the sudden decrease in test results of serum BAP in the PP seen in December 2013, which corresponded to levels in 2009, may represent a re-calibration.

We first notified the manufacturer of the assay with our observations in August 2013, and then in May 2014 we submitted a written review of our worries concerning the assay. The manufacturer did not agree that our observations could be explained by a shift in the BAP assay.

As seen in our results, the bias observed from the test kit may have started three years prior to our recognition of the problem. The subtle increase in values at the start may have contributed to the delay in recognition. However, once the limits for the PP were adjusted, we observed the first sign of an error. However, this error was not investigated further at that time because there was no documentation available on the stability of BAP in our pool. As a result, an assumption was made that the slow increase in BAP results of the pool could be explained by the instability of BAP in the pool. However, once the increase in test results of BAP in the pool accelerated, we reasoned that there must be another explanation, other than the instability of BAP in the pool. The sudden decrease in the BAP results of the pool, observed in December 2013, also

supported our assumption that the instability of BAP in the pool was not responsible for the steady increase in test results. Another factor contributing to our delay in recognizing the increase in BAP results may have been that BAP is not frequently tested and the test was only used for scientific purposes, and consequently not evaluated on regular basis by the clinicians. A routinely calculation of patient mean may probably have helped to recognize the problem sooner.

We have adjusted the test results of patients in clinical protocols for the period September 2010 to November 2013 using the PP as a reference. We believe this gives the best estimate for the exact value. No patients were harmed or adversely affected as a result of the inaccuracies identified, as BAP was solely used for research purposes. In conclusion, this study illustrates the need for an efficient multi-tiered QC system and for standardization and establishment of a primary reference method.

Acknowledgments: We would like to thank Liselotte Bjerregård Lauersen for her attentiveness in the laboratory work and for recognizing the steady increase in the results of BAP in the PP.
Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Employment or leadership: None declared.
Honorarium: None declared.
Competing interests: The funding organization(s) played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

References

1. Ismail AA, Walker PL, Barth JH, Lewandowski KC, Jones R, Burr WA. Wrong biochemistry results: two case reports and observational study in 5310 patients on potentially misleading thyroid-stimulating hormone and gonadotropin immunoassay results. Clin Chem 2002;48:2023–9.
2. Lapworth R, Teal TK. Laboratory blunders revisited. Ann Clin Biochem 1994;31:78–84.
3. Witte DL, VanNess SA, Angstadt DS, Pennell BJ. Errors, mistakes, blunders, outliers, or unacceptable results: how many? Clin Chem 1997;43:1352–6.
4. Vasikaran S, Eastell R, Bruyere O, Foldes AJ, Garnero P, Griesmacher A, et al. Markers of bone turnover for the prediction of fracture risk and monitoring of osteoporosis treatment: a need for international reference standards. Osteoporos Int 2011;22:391–420.
5. Cavalier E, Souberbielle JC, Gadisseur R, Dubois B, Krzesinski JM, Delanaye P. Inter-method variability in bone alkaline phosphatase measurement: clinical impact on the management of dialysis patients. Clin Biochem 2014;47:1227–30.
6. Bowers GN, Burnett RW, McComb RB. Preparation and use of human serum control materials for monitoring precision in clinical chemistry. Clin Chem 1975;21:1830–6.
7. ISO 15189:2012(E) Medical laboratories – Requirements for quality and competence. ISO: Geneva, 2012.
8. Henriksen GM, Pedersen MM, Norgaard I, Blom M, Blou L, Blaabjerg O, et al. Minimally processed fresh frozen human reference sera: preparation, testing, and application to international external quality assurance. Scand J Clin Lab Invest 2004;64:293–308.
9. Broyles DL, Nielsen RG, Bussett EM, Lu WD, Mizrahi IA, Nunnelly PA, et al. Analytical and clinical performance characteristics of Tandem-MP Ostase, a new immunoassay for serum bone alkaline phosphatase. Clin Chem 1998;44:2139–47.
10. Brinc D, Chan MK, Venner AA, Pasic MD, Colantonio D, Kyriakopolou L, et al. Long-term stability of biochemical markers in pediatric serum specimens stored at -80 degrees C: a CALIPER Substudy. Clin Biochem 2012;45:816–26.
11. Kazmierczak SC. Laboratory quality control: using patient data to assess analytical performance. Clin Chem Lab Med 2003;41:617–27.
12. Van Houcke SK, Stepman HC, Thienpont LM, Fiers T, Stove V, Couck P, et al. Long-term stability of laboratory tests and practical implications for quality management. Clin Chem Lab Med 2013;51:1227–31.
13. Wilson A, Roberts WL, Pavlov I, Fontenot J, Jackson B. Patient result median monitoring for clinical laboratory quality control. Clin Chim Acta 2011;412:1441–6.
14. Lott JA, Smith DA, Mitchell LC, Moeschberger ML. Use of medians and “average of normals” of patients’ data for assessment of long-term analytical stability. Clin Chem 1996;42:888–92.
15. Stepman HC, Stöckl D, Stove V, Fiers T, Couck P, Gorus F, et al. Long-term stability of clinical laboratory data: sodium as benchmark. Clin Chem 2011;57:1616–7.
16. Jørgensen LM, Hansen SI, Petersen PH, Sölétormos G. Median of patient results astool for assessment of analytical stability. Clin Chim Acta 2015;446:186–91.