(63 µg/L). We added 300 µL of the patient’s serum to 600 µL of a suspension of protein A-Sepharose from Pharmacia (binding capacity, 16 mg human IgG/mL) and incubated the mixture with rotation at room temperature for 20 h. This treatment had no effect on the interference exhibited in the VIDAS assay. This finding suggested that the interference might be attributable to IgM antibodies.

Finally, the serum was treated in heterophilic blocking tubes (HBTs) from Scantibodies, according to the manufacturer’s instructions. According to Scantibodies, these tubes contain a predispensed lyophilized heterophilic antibody-blocking active agent targeted to neutralize heterophilic antibody interference in immunoassays. This agent is a murine IgG directed against human IgM (9). After the treatment in HBTs, the prolactin concentration measured by the VIDAS decreased from >200 µg/L to 17.4 µg/L, whereas the Elecsys result remained practically unchanged (23.3 vs 21.6 µg/L). The finding that the prolactin concentration measured by the VIDAS decreased after treatment with the HBT blocking agent supports the hypothesis that the interfering antibody was an IgM, an idiotypic antibody that binds to a unique idiotype on the VIDAS antibodies. This epitope was apparently not present on the antibodies used with the other assays or on the immunoglobulins probably incorporated in the VIDAS assay reagents as blocking agents. Another hypothesis might be that the Elecsys, Vitros, and AxSYM reagents contain the HBT blocking agent. The interfering antibody is likely to be a natural IgM idiotypic antibody (10). Unfortunately, no more serum was available to confirm the IgM nature of the interfering antibody.

Macroprolactinemia is a frequent source of falsely increased prolactin values. Interference from antibodies is much less common, but may still occur with current prolactin IMAs. Among the four assays used in this study, three were efficiently protected against the antibodies contained in the patient’s serum. As described previously with human chorionic gonadotropin assays (11), treatment with HBTs was able to obviate the interference. In similar cases of hyperprolactinemia without image of pituitary tumor, after macroprolactin screening prolactin should be assayed by two or more IMAs. The use of HBT blocking agents should also be encouraged to prevent unnecessary anxiety and costly medical procedures and to reinforce clinicians’ and patients’ confidence in diagnostic assays.

This study was supported in part by a grant from the Hôpitaux Universitaires de Strasbourg, S. Doffoel and C. Bensimon are acknowledged for the Vitros ECI and AxSYM prolactin measurements, respectively. We thank N. Heider for carefully reviewing the English in the manuscript, bioMérieux for providing the VIDAS analyzer and prolactin assay reagents, and Immuno-tech for providing the protein A suspension.

References


Rémy Sapin*
Chantal Simon*
1 Laboratoire Universitaire de Biophysique Unité d’Analyse Endocrinienne CNRS UIPRES-A 7004 Université Louis Pasteur Faculté de Médecine 67085 Strasbourg Cedex, France
2 Service de Médecine Interne et Nutrition Hôpital de Hautepierre 67098 Strasbourg Cedex, France

*Address correspondence to this author at: Institut de Physique Biologique, Faculté de Médecine, F-67085 Strasbourg Cedex, France. Fax 33-3-90-24-40-57; e-mail sapin@ipb.u-strasbg.fr.

To the Editor:
A recent opinion article (1) in this journal presented sound metrologic principles and a plan of action to improve analytical accuracy in medical laboratories. In addition to the use of Certified Reference Materials outlined by Müller (1), use of proficiency testing and external quality-control programs can have a broader impact on interlaboratory comparability. Ideally, target values should be derived from reference methods and reference materials, but both are limited in their availability, and the mean (or other indicator of central tendency) of participant data is frequently used. This has been shown to provide a basis of comparison often comparable to reference methods (2).

The means of subsets of methods (peer-group means) are often reported as target values, a practice usually ascribed to method-dependent behavior of proficiency test specimens and so-called matrix effects (3, 4); this use of multiple “true values” in proficiency testing has been criticized (5, 6).

Target Values and Method Evaluation in Proficiency Testing Programs

Downloaded from https://academic.oup.com/clinchem/article-abstract/47/12/2185/5639481 by guest on 01 February 2020
In the US, the CLIA’88 serves as the basis for any accredited external quality-control program. The current regulations (7) state that only when “a specific method’s results demonstrate bias that is not observed with actual patient specimens, as determined by a defensible scientific protocol, a comparative method or a method group (‘peer’ group) may be used.”

Nonetheless, peer group evaluation is commonplace, perhaps reflecting potential for sanctions associated with failure in proficiency testing. Here we share two examples where this practice would have masked true analytical bias.

Case 1. In a recent proficiency test, the results for thyroxine from a participant laboratory demonstrated a high bias relative to both the robust mean (8) of all participants and that of the instrument peer group (Fig. 1). The degree of the high bias was less noticeable within the peer group; had the results been evaluated against the peer group, a CLIA score of 100% would have been obtained. The results were not peer graded, and three of the five results fell out of the allowable range and produced an unsatisfactory score of 40% (Fig. 1A, specimens 1–3). In spite of the bias demonstrated by this method, other laboratories using it attained a CLIA score of 100% when evaluated against the overall range or against the peer group-specific range, had the latter been used.

Examination of this failure indicated that the bias was attributable to the precipitation, over time, of magnetic particles in the reagent solution onto the side of the reagent container, producing a shift in reported values. Precipitation would not have occurred had the reagent been consumed within 24 h. The failing laboratory had a low test volume and stored the unused reagent at 4°C between analyses. The identified source of the bias was verified by reproducing the high-bias phenomenon with the stored reagent and by eliminating the high bias using the stored reagent but mixing it just before use (Fig. 1). After this investigation, the manufacturer issued a service notice in February 2001 and initiated activities to improve the reagent quality.

Case 2. Calcium is a non-peer-graded analyte in our proficiency test program with an allowable limit of ±0.25 mmol/L (1.0 mg/dL) from the all-method mean value (7). Methods agree well, and the failure rate in the New York State program is low [in the year 2000, 43 results of 6185 (0.7%) were unsatisfactory]. Over the course of several test events, a negative bias was consistently present for a particular instrument model with the o-cresolphthalein complexone method. The reported calcium values were either at or below the lower acceptable limit. The results within this instrument group were similar, and had peer group-specific target values been used, laboratories using this analyzer would have been awarded passing scores.

The manufacturer indicated that they accommodated the known bias by a commensurate downward adjustment of the reference interval [1.85–2.30 mmol/L (7.4–9.2 mg/dL)]. This range was lower than that typically used by clinical laboratories [2.10–2.55 mmol/L (8.4–10.2 mg/dL)] or found by reference methods (9).

Further investigation by the manufacturer led to changes in the calcium calibrator set point values. The set point adjustment (~9%) produced an increase in the recommended reference interval; proficiency test results for the analyzer now appear consistent with all-participant mean values.

These two examples demonstrate positive outcomes that may occur by avoiding peer-group evaluation in external quality assessment programs. Problems were identified that affected the quality of patient testing, resolutions were found, and comparability among methods was advanced. This supports the view of Büttner (5) that in “proficiency testing, so called ‘peer group mean values’ . . . do not lead to any improvement of the truefulness or therefore of the comparability”.

References

Robert Rej1,2*
Carol S. Norton-Wenzel1
(Tim) Zhimin Cao1

1 Wadsworth Center for Laboratories and Research
New York State Department of Health
Albany, NY 12201
2 Department of Biomedical Sciences
School of Public Health
State University of New York
Albany, NY 12201

*Author for correspondence. Fax 518-474-9185; e-mail clinchem@wadsworth.org.