

J. Clin. Chem. Clin. Biochem.

Vol. 26, 1988, pp. 441–446

© 1988 Walter de Gruyter & Co.  
Berlin · New York

## External Quality Control in the Determination of Neonatal Bilirubin An Approach to the Improvement of Results

By G. Röhle

*Institut für Klinische Biochemie, Universität Bonn*

H. Schlebusch

*Universitätsfrauenklinik, Bonn*

W. J. Geilenkeuser and R. Kruse

*Institut für Klinische Biochemie, Universität Bonn*

(Received December 22, 1987/March 3, 1988)

**Summary:** The reliability of bilirubin analyses is especially important in cases of neonatal hyperbilirubinaemia. However, when the means of the results of external quality control surveys and the method-dependent stated values for control sera were compared with reference method values, differences of up to 10% were found. Further inaccuracy arose from interlaboratory imprecision, which showed coefficients of variation of at least 7%, and from greater or lesser interference from contamination of samples with haemoglobin.

The present work investigates whether the current situation can be improved by available means.

### Introduction

One prerequisite for an unambiguous diagnosis of neonatal hyperbilirubinaemia is an accurate determination of the total bilirubin concentration in serum. Errors can lead to therapeutically unsound decisions, which in turn may result in a more or less pronounced brain defect (1). Thus, exacting demands are made on the reliability of the clinical-chemical analysis, and so far these demands have only been partially met (2–4):

- Current routine methods are not specific. Sources of interference are inherent in the analysis material, especially that of haemoglobin and its derivatives, which affect the principles of analysis in different ways and lead to method-related variations in the results.
- Bilirubin is comparatively unstable. Primary bilirubin standards that are easy to use for regular calibration and applicable for all routine methods are therefore not available. In addition, systematic

errors are responsible for interlaboratory variations in the results, even when these laboratories use the same methods.

The diagrams for exchange transfusions (5–7) and phototherapy (8, 9), which are used as aids in establishing a therapeutic scheme, do not take account of method-dependent bilirubin values. A critical situation therefore arises, in which the unsatisfactory comparability of the results from different laboratories working with varying methods and equipment can lead repeatedly to controversial decisions (10).

“Total bilirubin” consists of up to 5 fractions in the serum, whose relative quantities are so far undefined:

- 1) unconjugated, albumin-bound bilirubin,
- 2) bilirubin-monoglucuronide,
- 3) bilirubin-diglucuronide,
- 4) unconjugated bilirubin which is covalent bound to albumin (“ $\delta$ -bilirubin”) (11) and
- 5) free, unconjugated bilirubin.

The development of a reference method in its strictest sense (12) is hardly possible for such a group determination.

*Doumas* et al. (13) have, however, described a method for the determination of total bilirubin ("candidate reference method") which extensively satisfies the requirements of a reference method. This will, in the following, simply be called the reference method, despite the reservations inherent here. Furthermore, since 1986, a stable, secondary bilirubin standard developed by *Vink* (14) has been available, which contains bovine albumin, and is suitable only for chemical, not for direct photometric, bilirubin determination.

Under the given conditions, an attempt was made to test with optimal reliability the bilirubin concentration of 3 control specimens used for external quality control surveys as well as of several commercially available control samples, and to compare these results with those obtained by routine methods. The results from external quality control surveys for neonatal bilirubin carried out in cooperation with the Deutsche Gesellschaft für Klinische Chemie served as a criterion for the performance of routine methods.

## Materials and Methods

### Sample material

#### Standards

Bilirubin standards 2 and 3 were from the Rijksinstituut voor Volksgezondheid en Milieuhygiene in Bilthoven, the Netherlands.

Standard 2: Lot no. 850723, assigned value 98.8  $\mu\text{mol/l}$ .

Standard 3: Lot no. 850726, assigned value 290  $\mu\text{mol/l}$  (15).

#### Quality control survey samples

Unconjugated bilirubin was dissolved in a solution of albumin, human serum or bovine serum. Identical volumes of each material were placed in small test tubes and lyophilized. The material had partly been produced in our own laboratory and partly provided as a finished product by commercial companies. The composition of the material is specified for 6 quality control surveys, the results of which will be described in detail:

- Samples A and B of quality control surveys 3/83—3/84 each contained the same mixture of albumin and bilirubin; specimen B of quality control survey 3/83 was, however, additionally spiked with 310  $\mu\text{mol/l}$  haemoglobin, and specimen B of quality control surveys 1/84—3/84 with 93  $\mu\text{mol/l}$  haemoglobin.
- Specimens 101, 102 and 103, analysed in the quality control surveys 1/87, 2/87 and 3/87, consisted of human serum and varied concentrations of bilirubin.

#### Other control samples

The commercially available control samples were provided by the following companies:

Sta-Bil (human albumin, No BIC-975) Travenol, München.

Kontrollogen-Bilirubin (human serum, No 61 34 15) Behringwerke, Marburg.

Precibil (human serum No 1-211) Boehringer-Mannheim, Mannheim.

Controbil (bovine serum No 705) Merck, Darmstadt.

### Reference method values

Time of the determinations: February to April 1987.

The reference method values of specimens 101, 102 and 103 had been determined previously under similar, though somewhat less elaborate, conditions. The results of this first determination were slightly higher than the second; only the second determination is used here.

### Participating laboratories

Six laboratories took part in the present inquiry into reference method values. These were headed by *W. Appel*, Karlsruhe; *B. G. Blijenberg*, Rotterdam; *H. Brettschneider*, Penzberg; *G. Brüggemann*, Tübingen; *H. Schlebusch*, Bonn and *H. Wisser*, Stuttgart.

### Determination method

All quantitative determinations to establish reference method values were carried out according to *Doumas* et al. (13).

### Photometer controls

The spectrophotometers used were first tested: the wavelengths with a holmium nitrate solution and the absorption with gray filters, whose absorption for wave-length 598 nm were measured by the Physikalisch-Technische Bundesanstalt, Berlin.

### Volume controls

Each laboratory controlled its volume measurements according to the recommendations issued by the National Committee for Clinical Laboratory Standards, USA (NCCLS, Vol. 4 No 6).

### Experimental design

Each laboratory determined the bilirubin concentration in 2 standard samples, 3 quality control survey samples and 4 commercial control samples. All determinations were performed in duplicate, in 5 series of analyses.

### Establishing the reference method values

After eliminating the results from one reference laboratory, which showed the largest scatter, the median of the remaining 90 values for one given specimen was fixed as the reference method value.

### External quality control surveys

Quality control surveys for the determination of neonatal bilirubin have taken place three times yearly since 1983. The laboratories that applied for survey participation are each, with few exceptions, part of a hospital with facilities for the newborn. For each survey the participants analysed two samples, A and B. In the first survey of 1983, sixty pairs of analytical results were sent for quality control evaluation; in the third survey of 1987, this number had increased to 160.

### Evaluation

After every quality control survey, the results were divided into several subgroups according to the method of analysis. For every subgroup, statistical parameters were calculated: mean value, standard deviation and coefficient of variation.

### Methods taken into account

*Jendrassik-Grof* method (16); 2,5-dichlorophenyldiazonium (DPD)-method (17); direct spectrophotometry without dilution (bilirubinometer). Beginning with quality control survey 2/86, the results of direct spectrometry with dilution (18) were also evaluated separately.

### Transformation of analysis results with the help of a fictive standard

The participants' individual results for sample A of the quality control surveys 1/87 and 3/87 ( $A_i$ ) were multiplied with the quotient from the reference method value for sample B ( $RMV_B$ ) and the individual result for Sample B ( $B_i$ ).

$$\frac{A_i \cdot RMV_B}{B_i} = A_{is}$$

From the values standardized in this way ( $A_{is}$ ), the statistical parameters — mean value and standard deviation — were calculated (fig. 1).

### Criterion for excluding mean values and standard deviations in figure 1

When a first calculation of mean value and standard deviation gave a single value differing from the mean by plus or minus 3 standard deviations, this was eliminated before a second and final calculation.

## Results and Discussion

### Interlaboratory precision

An important purpose of the quality control surveys is to determine the degree to which the results from different laboratories differ from one another. An often very rough but otherwise very graphic estimate of the extent of scatter in quality control results emerges from the standard deviation or the relative standard deviation. It proved repeatedly to be the

case that interlaboratory precision increased more or less significantly after the start of the quality control programme. This was not the case, however, when neonatal bilirubin was determined, or, at best, improvement was only apparent in its initial stage.

In table 1 the mean values of the coefficients of variation are given for each of the 6 control samples that were analysed within a given year.

According to these, during the last 5 years there was no convincing improvement in interlaboratory precision for the 3 analysis methods most often used by the participants. A slight improvement was noticeable only over the last 2 years when all the results were taken into account without differentiating the methods. The available data allow the conclusion that this was due to a reduction in the differences between the mean values of each method collective.

This practically unaltered interlaboratory precision found in participants with the same analysis principles, points to the fact that needed improvement can hardly be expected by increased care on the part of the analysts. Apparently other measures are needed to improve a precision impaired most likely, and to the greatest extent, by systematic errors of differing magnitude.

### Accuracy

Table 2 shows that wide differences can exist between the results obtained by reference methods and those by routine methods. Not only the mean values of the results from quality control surveys but also the stated values cited by the suppliers of the control sera deviated up to 10% from those of reference methods. These findings accord with the results of previous studies (e.g. l.c. (3)). Such a noticeable difference raises the question of whether the calibration of the usual routine methods might not need to be re-examined and corrected.

Tab. 1. Means of some method-dependent data from three external quality control surveys for bilirubin determinations per year, each including the analyses of two specimens: mean of the number of participants (n), mean of the coefficients of variation (CV), mean of the relative differences (D%) between the methods means and the total means of all values for a given specimen.

Year	<i>Jendrassik-Grof</i>			2,5-Dichlorophenyl diazonium reagent			Direct photometry without dilution			All methods CV (%)
	n	CV (%)	D (%)	n	CV (%)	D (%)	n	CV (%)	D (%)	
1983	30	7.3	5.2	14	10.9	3.0	17	7.9	8.2	9.2
1984	40	9.3	3.0	24	7.2	4.7	25	9.1	7.7	10.0
1985	39	5.9	3.8	34	6.6	2.5	29	10.0	8.3	9.5
1986	43	7.2	1.7	45	6.9	2.3	43	7.5	3.5	7.9
1987	43	6.2	3.0	53	6.8	1.5	42	7.7	3.3	7.3

Tab. 2. Results of the determinations of "reference method values" (16% and 84% percentiles in parenthesis) for 9 bilirubin specimens (2 standards, 3 specimens, 101–103, used in external quality control surveys and 4 commercial control specimens) and the means of survey results<sup>1)</sup> and the stated values<sup>2)</sup> of specimen producers, respectively.

Specimen	Reference method value ( $\mu\text{mol/l}$ )	Method dependent values		
		<i>Jendrassik-Grof</i> ( $\mu\text{mol/l}$ )	2,5-Dichlorophenyl- diazonium method ( $\mu\text{mol/l}$ )	Direct photometry without dilution ( $\mu\text{mol/l}$ )
Standard 2 (98.8 $\mu\text{mol/l}$ )	97.9 (96.0–99.1)	—	—	—
Standard 3 (290 $\mu\text{mol/l}$ )	286 (280–289)	—	—	—
101	91.8 (88.5–93.9)	96.1 <sup>1)</sup>	101 <sup>1)</sup>	104 <sup>1)</sup>
102	178 (174–180)	188 <sup>1)</sup>	195 <sup>1)</sup>	201 <sup>1)</sup>
103	266 (262–268)	281 <sup>1)</sup>	288 <sup>1)</sup>	292 <sup>1)</sup>
Sta-Bil	316 (311–321)	325 <sup>2)</sup>	—	—
Kontrollrogen-Bilirubin	333 (327–337)	344 <sup>1)</sup>	356 <sup>1)</sup>	310 <sup>1)</sup>
Precibil	273 (268–276)	300 <sup>2)</sup>	300 <sup>2)</sup>	—
Controtil	303 (297–309)	322 <sup>2)</sup>	311 <sup>2)</sup>	—

### Possible effects from the use of standards

Independent of the possible necessity for new calibration procedures for routine methods, and considering the great diagnostic importance of neonatal bilirubin determination, the calibration of every analytical series should prove more reliable with the help of a supplementary, if only, secondary standard. The material developed in the Netherlands (14) is available as a standard for chemical methods.

To what extent this could improve the reliability of results is shown in an evaluation made by the control survey 1/87 (fig. 1), in which samples 102 and 103, from the same production supply, were studied. The

first group of columns shows the mean values and standard deviations arrived at for the first of the two samples (A). These results would have been much better if the participants had used the second sample (B<sub>1</sub>) as a standard for the first, which was exactly the requirement we made for our later calculation. The resulting second group of columns shows a very good match up between the means and the reference method values. Likewise, a considerable improvement in interlaboratory precision would have occurred for all methods of analysis. Both accuracy and interlaboratory precision would, in practice, be feasible if systematic intralaboratory errors of varying degree were eliminated by the use of suitable standards.

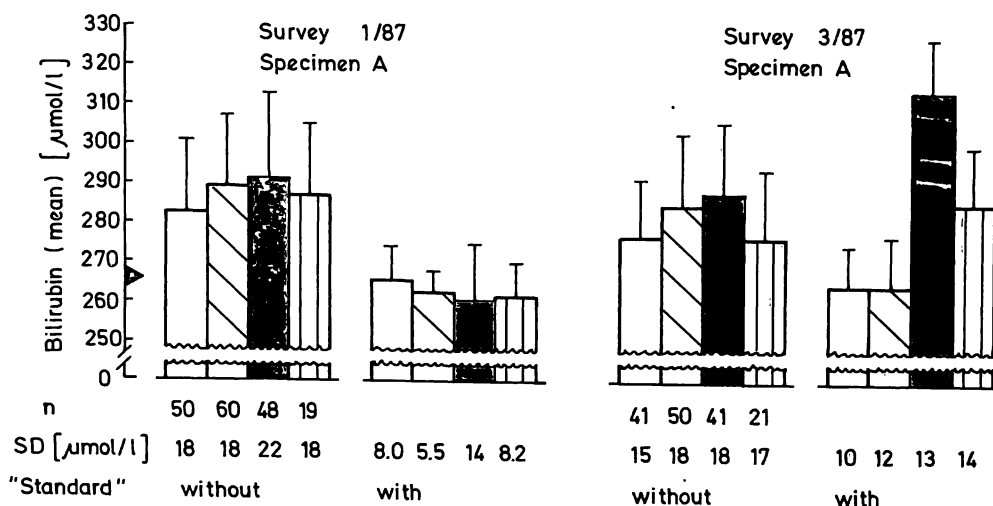


Fig. 1. Means and standard deviations (SD) of bilirubin determinations in specimen A of the surveys 1/87 and 3/87, and transformed means and standard deviations for the same specimen calculated on the basis of the individual results for the samples B<sub>1</sub> and B<sub>2</sub> when these were regarded as fictive standards with 178  $\mu\text{mol/l}$  and 333  $\mu\text{mol}$  bilirubin respectively.

- *Jendrassik-Grof* method
- ▨ 2,5-Dichlorophenyldiazonium method
- Direct spectrophotometry without dilution
- ▤ Direct spectrophotometry with dilution
- T Standard deviation
- n = number of results
- ▶ reference method value

Again, specimen 103 of the control samples from the quality control survey 3/87 served as sample A, whereas the Kontrollrogen-Bilirubin (No. 61 34 15) was sample B (B<sub>3</sub>). The specimens were from different suppliers and therefore had dissimilar matrices. The third column group in figure 1 shows, with no major difference from the first, the statistical parameters of the original results for sample A. The fourth group shows these same parameters after transformation with the individual results and with the reference method values for sample B<sub>3</sub>.

Again, with all methods there was a marked, if somewhat slighter, improvement in interlaboratory precision because of the transformation. The conversion using the diazo-methods even matched the reference method value almost exactly. But in the case of direct photometry, with and without dilution, the opposite effect occurred, i. e. accuracy decreased. This phenomenon can be partly explained if it is assumed that the human serum matrix of sample A contained carotenoids which could have influenced the photometric determinations more than the chemical ones. On the other hand, it is not clear why the mean of the photometric determinations of sample B<sub>3</sub> (Kontrollrogen-Bilirubin) was significantly lower than the reference method value (tab. 2). Control samples for which reference method values are available are therefore not automatically suitable as calibrators for bilirubin determination.

In connection with this observation it is worth noting that the specificity of direct photometry is called into question by the varying absorption spectra of the individual bilirubin fractions (11). This is surely one source of the criticism that has been levelled at direct photometry, especially when applied without dilution of the samples (3, 4, 19).

#### Haemoglobin interference

In four of the quality control surveys from 1983 and 1984, two specimens each were studied that were identical in matrix and bilirubin concentration. One of the two samples, however, was mixed with haemoglobin to simulate a medium or light haemolysis. These experiments took account of the experience that in determining neonatal bilirubin in patients' samples, an analysis often must be made with serum or plasma that is more or less haemolytic.

With one exception all analysis methods gave a median for the sample with haemoglobin that was lower than for those without (fig. 2). In the case of routine methods using the diazo-technique, this difference must be noted as an effect due to lack of specificity of known origin. Direct photometry without dilution, however, which is supposed to eliminate the interference from haemoglobin by measurements with 2 different wavelengths, does not seem to satisfactorily fulfil this requirement.

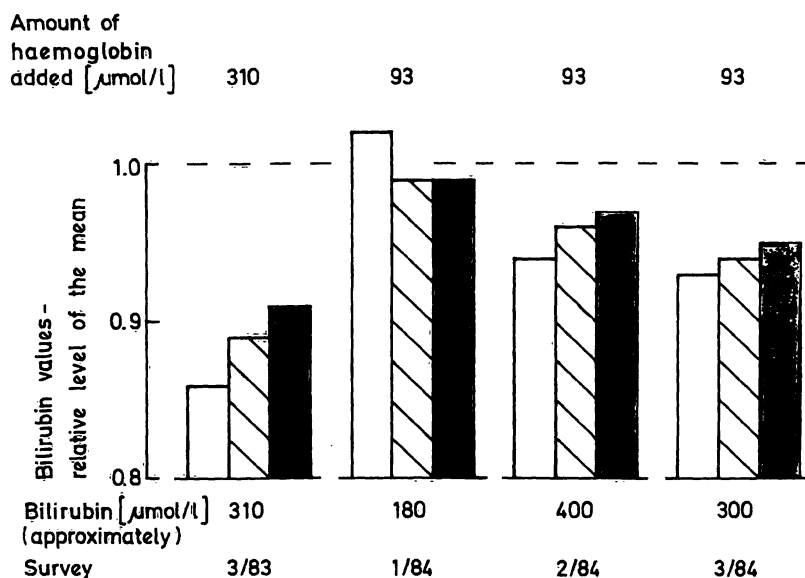


Fig. 2. Relative levels of the means of the bilirubin determinations of four quality control surveys in which the specimens B were spiked with haemoglobin, compared with the means for the specimens A (1.0) which were identical except for the haemoglobin contamination.

- Jendrassik-Grof method
- ▨ 2,5-Dichlorophenyldiazonium method
- Direct spectrophotometry without dilution

## Conclusions

Considering the special importance of bilirubin determination for neonatology, the state of analysis today is unsatisfactory. Substantial improvement could probably be brought about by means that are already available or could, in all probability, shortly become available.

1. Accuracy and precision of the results would be decisively improved if their calculation could be examined with the additional help of secondary standards. Samples developed as reference material in the Netherlands (15) are available for use as secondary standards for diazo-methods (cf. Material and Methods). Despite all reservations about using control samples as calibrators, commercial control series might also serve the same purpose. The prerequisite would be, however, that they are tested for suitability as calibrators for defined methods and that the reference method values for bilirubin concentrations are available.

As a result of being easy to work with, direct photometry is often used and, by a standardized method suggested by Blijenberg et al. (20), could prove to be a pragmatic step towards more reliable results.

2. Normal unspecific analysis methods could be replaced by the specific ones of Doumas et al. if they could be modified so that smaller sample volumes can be handled. According to the authors' test results (13), the problem of haemoglobin interference would thus be solved.

Improved analysis could bring about a situation in which a revision of the decision limits given in scientific publications (5–7) will need to be discussed.

## Acknowledgement

We would like to thank

Dr. A. B. T. J. Boink, Rijksinstituut voor Volksgezondheid en Milieuhygiene, Bilthoven, for allowing us to use the bilirubin standards.

Prof. Dr. W. Appel, Karlsruhe; Dr. B. G. Blijenberg, Rotterdam; Dr. H. Brettschneider, Penzberg; PD Dr. G. Brüggemann, Tübingen and Prof. Dr. Dr. H. Wisser, Stuttgart, for their commitment to establishing reference method values, a commitment that demanded a great deal of work from each of them.

Prof. Dr. R. Nink, Physikalisch-Technische Bundesanstalt, Berlin, for lending us grey filters for the photometry controls.

The companies Travenol, Behringwerke, Boehringer-Mannheim and E. Merck who made their own manufactured control series available to us.

## References

1. Deutsche Gesellschaft für Gynäkologie und Geburtshilfe, Deutsche Gesellschaft für Kinderheilkunde & Berufsverbände (1983) *Gynäkologie und Geburtshilfe* 3/83, 24–25.
2. Sitzmann, F. C., Polacek, K., Colombo, J. P., Schall, H., Schellong, G., Bidlingmaier, F., Kübler, W., Witt, I., Windorfer, A., Gleispach, H. & Mönch, E. (1975) *Pädiat. Praxis* 15, 295–301.
3. Blijenberg, B. G., Roetering, H. A. & Leijnse, B. (1987) *J. Clin. Chem. Clin. Biochem.* 25, 177–181.
4. Blijenberg, B. G., Brouwer, H. J., Roetering, H. A. & Leijnse, B. (1984) *J. Clin. Chem. Clin. Biochem.* 22, 609–612.
5. Polacek, K. (1963) *M Schr. Kinderheilkd.* 111, 6–11.
6. Brüster, H. & Wirtz, P. (1969) *Biol. Neonat.* 32, 332–339.
7. Wille, L. & Obladen, M. (1979) In: *Neugeborenen-Intensivpflege*, pp. 250–251, Springer Verlag.
8. Wiese, G. & Ballowitz, L. (1982) *Pädiat. Praxis* 26, 13–19.
9. Maisels, M. J. (1972) *Pediatr. Clin. North Am.* 19, 447–454.
10. Keller, H. (1986) In: *Klinisch-chemische Labordiagnostik für die Praxis*, p. 249, Georg Thieme Verlag, Stuttgart–New York.
11. Wu, T. W. (1984) *Clin. Biochem.* 17, 221–229.
12. Büttner, J., Borth, R., Boutwell, J. H., Broughton, P. M. G. & Bowyer, R. C. (1979) *Clin. Chim. Acta* 98, 129F–186F; (1980) *J. Clin. Chem. Clin. Biochem.* 18, 69–88.
13. Doumas, B. T., Kwok-Cheun, P. P., Perry, B. W., Jendrzek, B., McComb, R. B., Schaffer, R. & Hause, L. L. (1985) *Clin. Chem.* 31, 1779–1789.
14. Vink, C. L. J., Schuurman, W. & Van Gansewinkel, R. (1986) *Clin. Chem.* 32, 1389–1393.
15. National Institute of Public Health (The Netherlands) Report nr. 378507002 (1986).
16. Jendrasik, L. & Grof, P. (1938) *Biochem. Z.* 297, 81–89.
17. Wahlefeld, A. W., Herz, G. & Bernt, E. (1972) *Scand. J. Clin. Lab. Invest.* 29 (Suppl.) 126.
18. Hertz, H., Dybkaer, R. & Lauritzen, M. (1974) *Scand. J. Clin. Lab. Invest.* 33, 215–230.
19. Schlebusch, H., Liappis, N. & Niesen, M. (1988) *Pädiat. Praxis*, in press.
20. Blijenberg, B. G., Roetering, H. A., de Vos, A. & Leijnse, B. (1987) *J. Clin. Chem. Clin. Biochem.* 25, 737–741.

Priv.-Doz. Dr. G. Röhle  
Institut f. Klinische Biochemie  
Sigmund-Freud-Str. 25  
D-5300 Bonn 1