

# DESIGNING QUALITY CONTROL FOR NEONATAL SCREENING ASSAYS

Else-Maj R Suolinna<sup>1</sup>, Toni E Torresani<sup>2</sup>, James O Westgard<sup>3</sup>

<sup>1</sup>PerkinElmer Life Sciences Wallac, Turku, Finland; <sup>2</sup>University Children's Hospital, Department of Endocrinology, Zürich, Switzerland; <sup>3</sup>Department of Pathology and Laboratory Medicine, University of Wisconsin Medical School, Madison, WI, USA

**Abstract.** The purpose of neonatal screening is to find those that have a high risk for a disorder and therefore need further action for diagnosis and treatment. The separation of the high-risk and low-risk groups is typically achieved by establishing cut-off values for interpretation of the test result. The guidelines should preferably include both a low cut-off and a high cut-off value, with a grey zone between them. The width of the grey zone can be used for defining the quality required by the assay. The present work is based on the use of the EZ Rules™ software (from Westgard QC, Madison, WI), which automatically selects control rules. The user enters the needed parameters, such as the precision (as % CV) and bias of the assay. By using the grey zone as the medical decision interval the program will calculate possible control rules for the assay. The program was used to calculate the control rule for the AutoDELFIA™ neoTSH assay (from PerkinElmer Wallac, Turku, Finland). The grey zone was taken as 10 – 20 mU/L TSH, which is the recommendation of the American Academy of Pediatrics (Pediat. 91, 1203 – 1209). The entered parameters were: a total imprecision of 9%, which is typically seen with the AutoDELFIA neoTSH kit, a bias of 0% and a preanalytical variation of 20%. With the number of controls chosen as two, as often is used, a  $1_{3.0S}$  rule can be applied. The program also gives alternative control rules. Many laboratories lack a documented definition of the required quality, and tend to use a 2 SD control rule, which however leads to many unnecessary rejections. The EZ Rules program provides a tool for selection of QC rules. With the quality of the AutoDELFIA neoTSH kit two controls and a  $1_{3.0S}$  rule is sufficient. Runs are rejected only if one control out of two exceeds the 3.0 SD limit.

## INTRODUCTION

In neonatal screening all newborns are tested with the purpose of finding those at risk for a certain disorder within an apparently healthy population. In most countries the screening is centralised and the laboratories should be able to handle high sample volumes with high quality and low cost. The whole screening process requires careful quality management, so that all affected infants are identified and any needed treatment started as soon as possible. However, information on the quality requirements for neonatal assays is limited. Neonatal screening started with semi-quantitative assays, and most laboratories have based their quality control on a long knowledge of the performance of the assays. With more precise and automated assays a quantitative approach is essential. The analytical quality control in the laboratory can be managed by statistical QC procedures that take into account the quality required and the performance of the method (ie imprecision and inaccuracy).

The questions that must be answered in quality planning are:

- What is the required quality for the assay?
- How are the quality requirements translated into operating specifications for allowable imprecision and inaccuracy?
- What quality control is needed for the detection of medically important errors with minimum false rejections of assay runs?

**Quality requirement.** For neonatal screening tests, the quality requirement can be based on the purpose of the test, which is to classify the concentrations into one group that is below a certain cut-off level and another group that is above it, and preferably defining a grey zone<sup>1</sup> between them.

The so-called grey zone is a “clinical decision interval”, which can be used to derive the specifications for imprecision, inaccuracy and QC, as has been described (Westgard *et al*, 1991). Many neonatal screening laboratories have guidelines for interpretation that include a grey zone. Professional organisations also make recommendations that include a grey zone, eg the American Academy of Pediatrics guidelines for congenital

hypothyroidism screening define a grey zone of 10 – 20 mU/l TSH in blood (American Academy of Pediatrics 1993). The clinical decision interval requirement can then be used with the quality planning tool “OPSpecs charts”, which was introduced by Westgard in 1992 (Westgard, 1992a). The chart of operating specifications (OPSpecs chart) provides a practical tool for displaying the relationship between the quality required for the test, the inaccuracy and imprecision of the method, and different statistical control rules and the number of control measurements (Westgard, 1992, a,b).

## MATERIALS AND METHODS

The present work is based on the use of the EZ Rules™ software (Westgard QC, Inc, Madison WI) for constructing OPSpecs charts. After entering the data on the quality requirement and the performance characteristics of the assay, the program automatically selects QC rules. The selection can be based on either traditional statistical rules or average-of-normal results. The data presented here are from the routine neonatal screening program in Zürich, Switzerland (University Children’s Hospital, Dept. of Endocrinology, Zürich), using the PerkinElmer Wallac AutoDELFIA™ neoTSH kit (PerkinElmer Life Sciences, Wallac Oy, Turku, Finland).

The program uses an *interview mode* (Figs 1 and 2). After entering the test name the first question is about the *type of QC rule* to use. In the case of TSH the choice is usually for traditional statistical rules. For T4 or 17OHP the “average of normals” may be an alternative if large enough number of samples are analysed. The interview mode of the software then asks for *the quality requirement* as mentioned above. The use of the neonatal tests is best based on the clinical decision interval (grey zone). For the neoTSH assay the Zürich screening laboratory uses a grey zone of 15 – 25 mU/L blood. Therefore the decision level is 15 mU/l TSH and the interval (10 mU/l) is 66.6% of the decision level.

The next question is about *pre-analytical factors*, eg biologic variation, sampling variation and specimen bias. Biologic variation can account for a known group or within-subject variation. Data on TSH within-subject variation for neonates is not available, but for adults it is given as 19.7%, and is used in this example. The other pre-analytical factors are difficult to estimate. However, the performance data for the kit have been obtained with controls on filter paper, ie the same matrix as the actual specimens, and thus the sampling variation and specimen bias are already included in them. This may not be entirely true, as all heel prick specimens are not of equal quality.

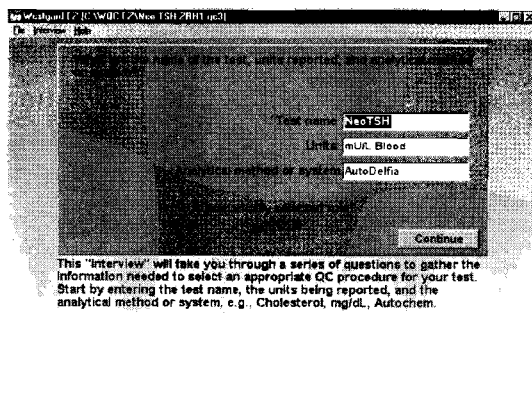


Fig 1. Interview mode.

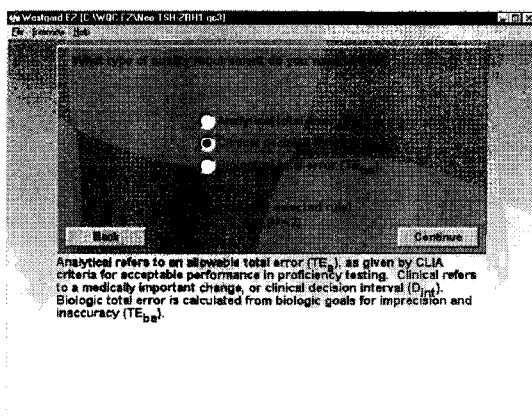


Fig 2. Quality requirement.

The *performance data* are then entered. These can be either the data provided by the manufacturer in the kit insert or preferably, as here, from data generated in the laboratory. In the Zürich screening laboratory the CV for the neoTSH assay at a level of 16 mU/l TSH is 7.3% (derived from 150 assays), and the estimated bias 4.6%. The bias can be estimated from eg external quality assessment programs. If the laboratory determines the cut-off values with the same kit, the bias can be taken as 0. *The instability of the assay* is a measure of overall problems encountered with the kit – if the kit performance is stable with infrequent problems, < 2% can be entered.

The program then asks for *the number of controls* that is used – the number of controls supplied with the neonatal kits is usually 2 (a low and high, representing a normal and abnormal value), but the laboratory can of course include in-house controls. After entering the number of controls, the program automatically generates a *QC rule* and the *OPSpecs chart*.

## RESULTS

Fig 3 below shows the generated OPSpecs chart, with the allowable imprecision (as %CV) on the x-axis and allowable inaccuracy (as % bias) as the y-axis. The "Operating Point" shows the position of the method in use, ie a CV of 7.3% and a bias of 4.6% in the present example. The chosen number of controls is 2, and the rule selected by EZ Rules program is  $1_{3.0s}$ , which is shown as the bold line. However, the program generates a series of alternative QC procedures and gives the corresponding control rules in the table on the right in the same order as in the graph (from down up).

The two best control rules (giving 90% error detection and no false rejections) are:

- 2 controls and  $1_{3.0s}$  rule, ie a run is rejected if one of the two control values is outside  $\pm 3.0$  SD limits
- 4 controls and a  $1_{3.5s}$  rule, ie a run is rejected if one of the four control values is outside  $\pm 3.5$  SD limits

Both rules give a 90 % error detection (Analytical Quality Assurance) and no false rejections. Many laboratories may prefer to use only two controls, as it leaves more space for samples.  $N = 2$  (or 4) can be either the same control run twice (or four times) or separate controls. R is the number of runs that the control can be applied to, and would in the AutoDELFA assay be one plate, as each plate should have its own controls.

## DISCUSSION

With the more precise and automated neonatal screening assays a quantitative approach to quality control becomes both possible and essential. Defining a grey zone between the lower and upper cut-off values is one way to approach the quality requirement. The purpose of the screening assay is a correct classification of the result into a low- and a high-risk group. Defining only one set cut-off does not allow for any variation in the assay, or if defined conservatively to allow for variation, it leads to many unnecessary false positives. Results in the grey zone are confirmed by additional testing before classification into the low- or high-risk group.

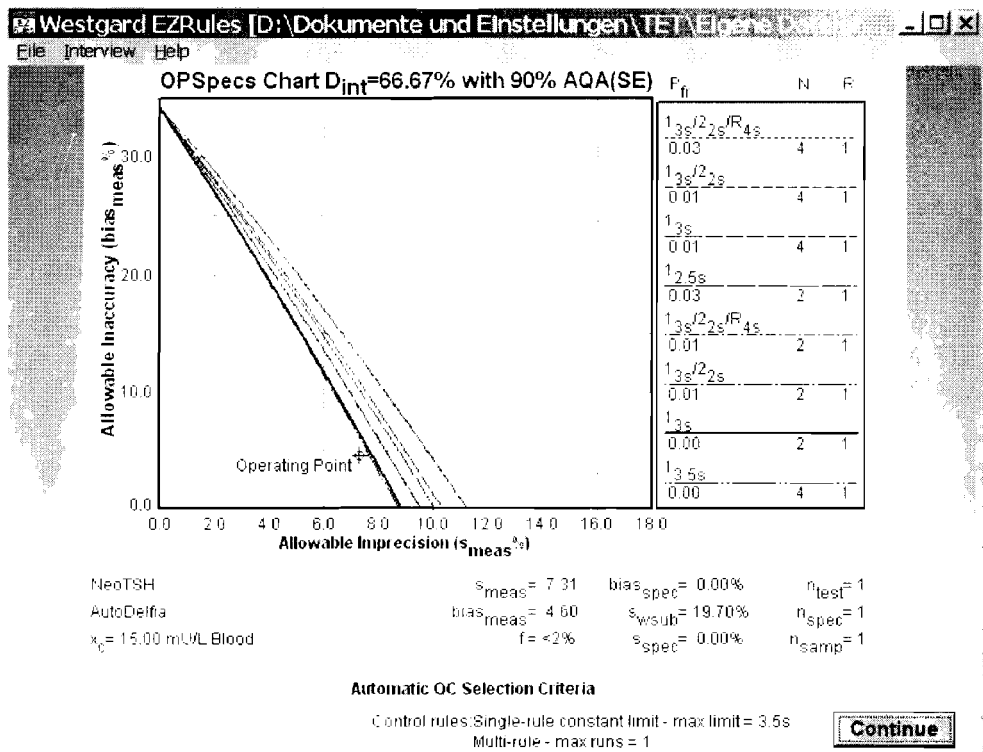


Fig 3. OPSpecs chart.

The example shows that with the performance of the AutoDELFIA neoTSH assay the needed quality control can be 2 controls and a 1<sub>3.05</sub> rule. The two controls can be either the same control twice or two different controls. Runs should be rejected if one of the two controls is outside the 3.0 SD limits. An additional benefit of the program is the QC validation report, to which comments can be added. It can be printed for documentation for eg accreditation purposes.

#### CONCLUSION

Many laboratories may lack a documented definition of the quality required for each assay, and quite often a 2 SD rule is in use. A 2 SD rule leads to many unnecessary rejections as statistically 5 % of the runs will be outside even when everything is OK. This means too many false alarms, too many rejections of runs and too many repeats. The EZ Rules program provides a tool for a solid basis for selection of QC rules, more security and a documented report.

The presented approach gives QC with scientific basis, and gives the laboratory more confidence, less rejected runs and cost savings.

#### REFERENCES

- American Academy of Pediatrics. *Pediatrics* 1993;91:1203-9.
- Westgard JO, Petersen PH, Wiebe DA. Laboratory process specifications for assuring quality in the U.S. National Cholesterol Education Program. *Clin Chem* 1991; 37:656-61.
- Westgard JO. Charts of operational process specifications ("OPSpecs Charts") for assessing the precision, accuracy, and quality control needed to satisfy proficiency testing criteria. *Clin Chem* 1992a; 38:1226-33.
- Westgard JO. Assuring analytical quality through process planning and quality control. *Arch Pathol Lab Med* 1992b; 116:765-9.

<sup>1</sup> Why grey zone? Take the actual clinical cut-off value for congenital hypothyroidism eg as 15 mU/l TSH. A method with a 10% CV will give values of 12 – 18 mU/l (95% of results), even when everything is running well. Thus one would either have to reduce the cut-off, which would increase the false positives, or define a grey zone, for which eg repeating the sample in duplicate would confirm the result.