External Quality Control for Semen Analysis in Germany
- Qualitätskontrolle der Deutschen Gesellschaft für Andrologie (QuaDEGA). The First 5 Years

Cooper TG, Hellenkemper B, Nieschlag E

J. Reproduktionsmed. Endokrinol 2007; 4 (6), 331-335

www.kup.at/repromedizin
Online-Datenbank mit Autoren- und Stichwortsuche
ECMO für die Kitteltasche

2. Auflage Jänner 2019
ISBN 978-3-901299-65-0
78 Seiten, div. Abbildungen
19.80 EUR

Bestellen Sie noch heute Ihr Exemplar auf
www.kup.at/cd-buch/75-bestellung.html
External Quality Control for Semen Analysis in Germany – Qualitätskontrolle der Deutschen Gesellschaft für Andrologie (QuaDeGA)  
The First 5 Years

T. G. Cooper, B. Hellenkemper, E. Nieschlag

For an individual clinic, knowledge of the differences between technicians in their assessment of semen is critical for the day-to-day running of the practice. This is achieved by the use of standardised procedures, embodied in the WHO Laboratory Manual for the examination of human semen and sperm-cervical mucus interaction [1, 2] and participation in regular internal quality control (IQC) schemes of semen analysis procedures. Furthermore, in a country such as Germany, the accuracy of semen analysis results is of importance with regard to the indication of, and reimbursement for, assisted reproduction. For these clinical reasons, as patients move between centres, and for research purposes, e.g. participation in multicentre studies, knowledge of how an andrology laboratory compares with other centres is of increasing importance. Comparisons between centres are controlled by external quality control (EQC) schemes that operate in several European countries (the UK, Italy, Spain, Scandinavia, Denmark). In 2001, the German Society for Andrology set up an external Quality Control Programme for Semen Analysis (Qualitätskontrollen der Deutschen Gesellschaft für Andrologie, QuaDeGA), for German-speaking countries. On its 5th anniversary, the results from this programme are examined and the usefulness of the programme is assessed.

Materials and Methods

QuaDeGA Distributions
QuaDeGA supplies two ring trials per year, each comprising two diluted, fixed semen samples in tubes with colour-coded caps for assessment of sperm concentrations and for making smears for sperm morphology assessment. The fixative was initially formalin (100 µl 37% formaldehyde per ml semen) but some samples displayed clumping and aggregation particularly when semen samples from different donors were mixed. Attempts were made to reduce this problem by pooling semen from individual donors or use of a storage medium containing an anti-foam agent [3]. Videotapes of sperm motility gave way with time to CD-ROM and DVDs for most centres. Data from all 10 distributions (2002 to 2006) were examined.

The Use of a Ranking System to Monitor Changes in Performance
Providing two semen samples to assay in each distribution permits the results to be plotted graphically in a Youden plot where the results for the same parameters from each of the two samples are plotted against each other [4]. This enables both random and systematic errors to be detected and allows a ranking system to be developed. For QuaDeGA, Rank 1 is achieved when the results of both samples fall within the target window; Rank 2 when one is in and one out, Rank 3 when both results are outside the target window and a systemic error is indicated by both samples being too high or too low and Rank 4 when random errors (e.g. one sample too high, one sample too low) are indicated (Fig. 1). For each participating centre, the ranks awarded for each semen parameter were examined. These ranks were used in a number of ways to determine if the assessment of semen analysis had changed with time during the course of participation in the QuaDeGA scheme.

For every centre, the ranks apportioned to each semen parameter were listed for every distribution. From this, the longitudinal changes in ranks with time (participation in the scheme) and the cross-sectional average of ranks (for each distribution) could be calculated.
Target Values

The most difficult task for an EQC programme is to set realistic target values. Initially taking data only from the few EAA centres, this was expanded to centres that provided consistent results on occasions when, unknown to them, the same sample had been provided, but at two dilutions, as separate samples for QuaDeGA. Such centres thus provided consistent morphological scores and concentrations but did not necessarily produce consistent assessment of motility. Here other centres’ results were used but there was difficulty in obtaining centres consistently agreeing with all four motility grades. On occasion good agreement with grade d would lead to even better agreement (a small target window) for grade a; this occurred in distribution 6 such that fewer centres were able to achieve Rank 1 with the distortion of results mentioned below.

Even these selected centres with time produced inconsistent results with the target windows being either too large (so that all centres would attain Rank 1) or too small (so that only few would be able to obtain Rank 1) and values from new recruits to the scheme were examined for suitability as reference centres. It must be recognised that with target values set subjectively, some variability in response will inevitably originate from the centre setting the target values [5].

Results

Number of Participants

The number of participants increased from 27 at the initial distribution to 145 by the tenth (Tab. 1). The breakdown of the laboratories participating in the last distribution (2/2, 2006) showed that there were 40 % from Urology practices, 35 % from ART clinics, 13 % from dermatology institutes, 10 % from laboratory medicine laboratories, ~2 % from andrology clinics and one semen bank. Participants also came from outside Germany: 9 from Switzerland, 3 from Austria, 1 from Hungary and 1 from Sweden. Eighteen centres started the QuaDeGA scheme but dropped out; the other centres that remained in the scheme sometimes missed one distribution (34 centres), two distributions (4 centres) or three distributions (2 centres). Participation in QuaDeGA is voluntary, but a fee was charged with a reduction for members of the German Society of Andrology for running the programme.

Choice of Counting Chambers

A large number of counting chambers was used by participating centres for estimation of sperm concentrations, ranging from reusable deep haemocytometers (Neubauer, Fuchs-Rosenthal, Thoma, Bürker, Bürker-Türk/Türk) and shallow chambers (Makler) and disposable chambers (Cell vision, Mika, Fastread, Kovac, MTG and ZMS) as well as CASA systems (IDENT, IVOS) (Tab. 1). From all the centres providing more than one result, 105 remained with the same counting chamber throughout their participation in QuaDeGA, 35 changed counting chambers once and 13 changed twice. Of those changing chambers, 25 (52 %) changed to a chamber recommended by the WHO [1, 2]; nevertheless, by the last distribution, only 46 % of the participants were using the Neubauer counting chamber, as proposed by the WHO [1, 2] (Tab. 1).

Choice of Semen Diluents

A wide range of fixatives and diluents was used for assessment of sperm concentration, including the WHO-recommended formalin, other diluents containing formalin, stains (Carbolfuchsin, eosin), chemicals (triphénylamine, peroxidase media), bleach (chlorate), acids (acetic, HCl), bicarbonate-buffered media, lysing solutions, salt solutions (NaCl, KCl) water or no media at all (for those using the Makler chamber). Of all the centres providing more than one result, 89 remained with the same diluent throughout, 39 changed diluent once and 17 changed twice. Of

Table 1. Characteristics of the methods used by participants in the first 10 QuaDeGA distributions

<table>
<thead>
<tr>
<th>Distribution No.</th>
<th>Chambers</th>
<th>Diluents</th>
<th>Stains</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. Code</td>
<td>Participants</td>
<td>No.1 % WHO</td>
<td>No.1 % WHO</td>
</tr>
<tr>
<td>1 2002/1</td>
<td>27</td>
<td>4 56</td>
<td>5 13</td>
</tr>
<tr>
<td>2 2002/2</td>
<td>48</td>
<td>8 33</td>
<td>6 11</td>
</tr>
<tr>
<td>3 2003/1</td>
<td>68</td>
<td>5 39</td>
<td>8 8</td>
</tr>
<tr>
<td>4 2003/2</td>
<td>99</td>
<td>9 43</td>
<td>11 19</td>
</tr>
<tr>
<td>5 2004/1</td>
<td>102</td>
<td>9 43</td>
<td>10 19</td>
</tr>
<tr>
<td>6 2004/2</td>
<td>110</td>
<td>9 43</td>
<td>12 18</td>
</tr>
<tr>
<td>7 2005/1</td>
<td>115</td>
<td>10 44</td>
<td>9 19</td>
</tr>
<tr>
<td>8 2005/2</td>
<td>128</td>
<td>11 47</td>
<td>9 20</td>
</tr>
<tr>
<td>9 2006/1</td>
<td>132</td>
<td>12 47</td>
<td>10 23</td>
</tr>
<tr>
<td>10 2006/2</td>
<td>148</td>
<td>11 46</td>
<td>10 25</td>
</tr>
</tbody>
</table>

1 number of different types of chambers, diluents or stains used including no diluent (for Makler chambers) or no stains; 2 percentage of centres employing WHO-recommended chambers, diluents or stains; 3 percentage of centres using no diluent (Makler chambers) for concentration or no stain for morphology.

Figure 1. A Youden plot of sperm concentration. For a given distribution, the results of the assessments of each centre for Sample A (abscissa) is plotted against the assessments of the same centre for Sample B (ordinate). The target ranges for the concentration in each sample (provided by reference centres) is shown between the dotted lines on each axis and they overlap to define the target window (solid line). The positions of points outside the target window indicate random errors (R) or systematic errors (S). The graph is thus divided into 9 areas and Ranks are awarded for falling within the target window (Rank 1), for having random errors (Rank 2 in R2), systematic errors (Ranks 3 in S3) and extreme random errors (Rank 4 in R4).
those changing diluents, 19 (34 %) changed to a chamber recommended by the WHO but by the last distribution, no more than 25 % of the participants were using the WHO-recommended diluent (Tab. 1).

**Morphological Stains and Assessment Criteria**

A wide range of stains was used for assessment of sperm morphology, including the WHO-recommended Papanicolaou, Shorr and Diff Quik stains as well as Berenyi, Gram, Giemsa, May-Grünewald, Pappenheim, Mayer, Haematoxylin, Eosin, Cytocolour, SpermAc, aniline blue, methylene blue, toluidine blue, SpermOscan, SpermAc and rapid stains (Sangrodi, Testsimplets), or no stain at all. Three main criteria for assessing morphology were employed: the WHO manual, Düsseldorf [6, 7] and Strict assessment [8, 9], although the WHO manual [1, 2] is meant to be based on the Kruger (strict) scheme. From all the centres providing more than one result, 56 remained with the same stain throughout, 13 changed stains once, 36 changed twice, 11 changed three times and 4 changed 4 times. Of those changing morphological stains, 18 (28 %) changed to a WHO-recommended stain but by the last distribution only 30 % were using WHO-recommended stains.

For all distributions, between 84 and 94 % of all centres claimed to be using WHO criteria for categorising normal morphology. From all the centres providing more than one result, 104 remained with the same criteria throughout, 4 changed criteria once and 6 changed twice. Of these, all 10 (100 %) changed to a method recommended by the WHO [1, 2].

**Overall Trend in Ranks with Time**

The overall mean rank values for each semen parameter were calculated for each distribution from all centres participating at that time (Fig. 2). A gradual improvement (decrease in mean ranking) was observed for assessment of sperm concentration and recognition of normal morphological forms. A sharp improvement of assessment of grade a (fast progressive) motility was evident (the hiatus in distribution 6 being caused by an abnormally small window for grade a sperm from the particular reference centres chosen). Motility grades b (slow progressive) and c (non-progressive motility) likewise improved. Immotile spermatozoa (grade d) were consistently well assessed.

**Changes Between Ranks with Time**

The percentage of participating laboratories receiving the Ranks 1, 2, 3 or 4 were calculated for each distribution for each semen parameter controlled. These results are shown in Figure 3. For all parameters there were very few centres receiving Rank 4 (random errors) and after distribution 4 no Rank 4 needed to be awarded. Systematic errors or bias (Rank 3) starting at around 20 % of centres for morphology, 40 % for concentration and motility grade c and 60 % for progressive motility (grades a and b) declined gradually through the five years of study, with a few reversals in distribution 6. The awarding of Rank 2 persisted throughout for all parameters and showed some fluctuations, particularly with grading progressive sper-
The percentage of centres given Rank 1 increased drastically for the assessment of concentration and motility (grades a, b and c). Grade d was always assessed well.

**Improvement in Ranks with Continued Participation**

Taking data only from centres that had been with the QuaDeGA scheme for at least 6 distributions enabled within-centre improvement with time to be monitored. For this the mean rank for each semen parameter from the first 3 distributions analysed was compared with the mean rank from the last 3 distributions from that centre. An improvement in performance was considered if there was an increase of at least a ranking of 1 between the means of the first and last 3 distributions. If the initial rank was 1, no improvement could be made, so achieving a mean rank of 1 by the last 3 distributions was also counted as an improvement. Figure 4 shows that there was a consistent increase in the percentage of centres improving performance by at least a ranking of 1 during their participation in the EQC scheme. The improvement in assessment of concentration occurred after participating 4 times whereas improvements in assessment of normal morphological forms and motility grade c were more immediate. There were no obvious trends in improvement in the assessment of motility grades a, b or d. Whereas the difficulty distinguishing between grades a and b is well known [10], explaining the fluctuations in response and the implied poor ability to assess immotile spermatozoa was surprising.

**Counting Chambers and Diluents Used for Assessing Sperm Concentration**

As the most-used chambers were the Neubauer and Makler, a comparison of results from their use was made by selecting only those centres that had used these chambers throughout their participation in QuaDeGA (up to 23 or 21 centres, respectively, by QuaDeGA distributions 9 and 10). Data from at least 9 centres per distribution are shown in Figure 5. Use of both chambers resulted in an improvement in performance but lower ranks were consistently found with the Neubauer chamber, with less fluctuation between distributions. Of the 186 samples examined in a Neubauer chamber 73 % were Rank 1, 28 % Rank 2, 10 % Rank 3 and 2 % Rank 4; the corresponding percentages for 144 samples examined in a Makler chamber were 63, 27, 10 and 1 (Fig. 5). As the semen samples distributed to the centres were fixed, the spermatozoa assessed in the chambers were immotile. This contrasts with the situation in the laboratory where motile spermatozoa in undiluted semen is assessed in Makler chambers; thus this does not constitute a control of the quality of sperm concentration assessment made in those laboratories under these different conditions.

**Discussion**

In this paper, the performance of laboratories participating in the external quality control scheme for semen analysis of the German Society of Andrology (QuaDeGA) was monitored. Using a ranking system in which Rank 1 indicates success in reaching the target value for a particular semen variable, improvement was apparent as increased overall rankings, an increased percentage of centres with Rank 1 or as a time-dependent decrease in mean rank with time.

Only a minority of the participants used WHO-recommended methods and few changed to them during their participation in the scheme despite recommendations sent out with the results. A similar lack of compliance with WHO recommendations has been found for morphological assessment by the British Andrology Society [11]. It is difficult to understand why standardised methods are not introduced into semen analysis, especially when the WHO semen analysis handbook is recom-
mended in the letter accompanying the results from every QuaDeGA ring trial. Results from the andrology laboratory are not solely the responsibility of the technicians performing the analysis; laboratory and institute directors should ensure that, as in every other field of medicine, standardised methods are used.

In this study, where a valid comparison could be made with sufficient numbers of participating centres, fewer fluctuations in sperm concentration between ring trials were found by use of the WHO-recommended Neubauer chamber, most likely as a result of the larger number of spermatozoa assessed, and lower counting errors accrued in the large-volume chambers [12].

The plethora of stains used for morphology, together with the notorious difficulty of morphological assessment, made a similar assessment for morphology not meaningful. Until unified approaches to sperm morphology assessment are introduced, together with effective training schemes [13], little headway will be made. A confounding factor is the difference in assessment of sperm morphology by two European schemes (one being more strict in assessing “borderline spermatozoa” as abnormal) which has already been reported [5]. Close attention to the preparation of smears, use of stains and assessment of normal forms is required to reduce variability in this assessment and this can also be achieved by standardisation of techniques.

The centres that showed improvement in assessment of sperm concentrations and morphology took time to do so (participation in 3 to 4 ring trials), suggesting that changes in assessment/method were gradual and in response to continued feedback in the form of results from the scheme. Although certain laboratories starting with poor performance (high ranks) at their entry into the QuaDeGA scheme were able to improve and maintain improvement with time, other centres were erratic and showed no improvement. Whether this reflects the qualification of the staff or staff turnover is not possible to say from the data collected, but that a certain type of technician is better suited to andrological analysis is well known [14].

In conclusion, although the QuaDeGA scheme has failed in introducing WHO-recommended semen analysis protocols to the German andrology scene, continued participation in the scheme has been associated with an improvement in assessment of some semen parameters for some centres. Improvement in other parameters, or in other centres, will rest upon the introduction of unified protocols that should help eliminate variability in assessment between centres. In this regard, a new version of the WHO manual for semen analysis is in preparation that covers in greater detail than before the basic methods of assessing sperm concentration, motility and morphology.

References:

Haftungsausschluss


Bitte beachten Sie auch diese Seiten:

Impressum  Disclaimers & Copyright  Datenschutzerklärung

Mitteilungen aus der Redaktion

Besuchen Sie unsere Rubrik

✔ Medizintechnik-Produkte

Neues CRT-D Implantat Intica 7 HFT QP von Biotronik

Artis pheno Siemens Healthcare Diagnostics GmbH

Philips Azurion: Innovative Bildgebungslösung

Aspirator 3 Labotect GmbH

InControl 1050 Labotect GmbH

e-Journal-Abo

Beziehen Sie die elektronischen Ausgaben dieser Zeitschrift hier.

Die Lieferung umfasst 4–5 Ausgaben pro Jahr zzgl. allfälliger Sonderhefte.

Unsere e-Journale stehen als PDF-Datei zur Verfügung und sind auf den meisten der marktüblichen e-Book-Readern, Tablets sowie auf iPad funktionsfähig.

✔ Bestellung e-Journal-Abo