

External Quality Control of Semen Analysis Reveals Low Compliance with WHO Guidelines

E. Nieschlag, T. Pock, B. Hellenkemper,

QuaDeGA (Quality Control Programme of the German Society of Andrology)

Semen analysis is the oldest and most commonly used surrogate parameter for male fertility. Like all other clinical laboratory tests, semen analysis should be under strict internal and external quality control. WHO has pioneered standardisation and quality control of semen analysis. The external Quality Control Program of the German Society of Andrology (QuaDeGA) is based on WHO guidelines and since 2011 participation in ring trials is compulsory for all laboratories performing semen analysis in Germany. However, many laboratories fail to follow these guidelines so that it remains unclear whether the wide scatter of results from ring trials is caused by protocol failures or is inherent to semen analysis. In order to resolve this question we performed a survey among 624 participants and received valid answers from 256 (42.5%). The answers revealed that many laboratories lack basic equipment such as mixers, vortex, positive displacement pipettes and appropriate microscopes, do not use the recommended improved Neubauer chamber for counting, use inappropriate staining methods and do not evaluate sufficient sperm. Most surprisingly, 23% of the laboratories apply no internal quality control and these have a higher failure rate than those with internal control (28 vs 8%). Laboratories performing more than 20 semen analyses per month score better than those with lower sample frequency. Strict observation of WHO guidelines and more intensive teaching and practical training of technicians appear to be the most important measures to improve results. Until better agreement of results within and between laboratories is achieved, arguing over normal reference and threshold values remains a secondary problem. **J Reproduktionsmed Endokrinol_Online 2017; 14 (6): 306–10.**

Key words: Quality Control Program of the German Society of Andrology, QuaDeGA, semen analysis

Introduction

In 2011 the German Federal Medical Board (BÄK) integrated semen analysis into its compulsory external quality control program for medical laboratories [1]. Since then all laboratories performing semen analysis for human medical purposes are obliged to participate twice yearly in ring trials (“Ringversuche”) to obtain a certificate as a pre-requisite for charging patients or insurances for semen analysis. The BÄK guidelines require participating laboratories to use methodology as described in the current World Health Organization (WHO). Lab-

oratory Manual for the Examination and Processing of Human Semen. At the present time this is the 5th edition of WHO Manual [2] for which a German translation has been published [3]. The Quality Control Programme of the German Society of Andrology (QuaDeGA GmbH) was licensed by the BÄK as official Reference Institution to conduct the compulsory ring trials.

QuaDeGA was established in 2002 and had performed ring trials on a voluntary basis until 2011. Before participation became mandatory in 2011 250 laboratories had taken part in this external quality

control program. Since then the number has increased to over 700 (Fig. 1).

While the number of participants in the programme has almost tripled since it became compulsory, over the years the Youden plots continued to result in broad windows and the rate of obtaining the certificate has remained around 80 % with only a small tendency for improvement (Fig 2). In an attempt to find out why the failure rate remains high, we conducted a survey among the participants asking for details of their laboratory techniques, especially in regard to guidelines provided by the WHO Manual [2, 3].

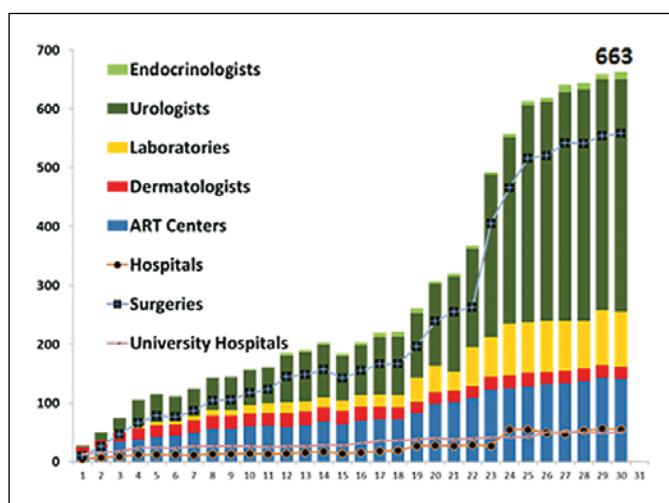


Figure 1. Participants in the QuaDeGA program 2002–2016 (ring trial 1–30).

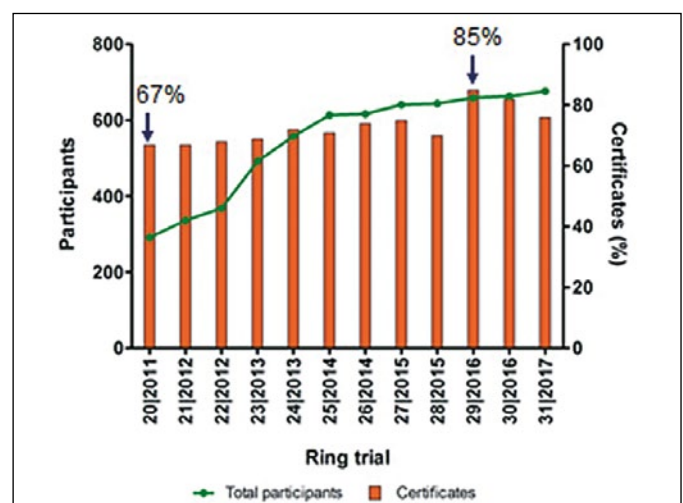


Figure 2. Development of total number of participants and the percentage of those obtaining a certificate from ring trial 20/2011 to 31/2017.

Received: October 12, 2017; accepted after revision: October 26, 2017 (responsible Editor: Prof. H. Behre, Halle)

From the Center of Reproductive Medicine and Andrology, University Hospital Münster, Germany

Correspondence: Univ.-Prof. em. Dr. med. Dr. h.c. Eberhard Nieschlag, FRCP, Center of Reproductive Medicine and Andrology, Domagkstraße 11, D-48149 Münster, e-mail: Eberhard.Nieschlag@ukmuenster.de

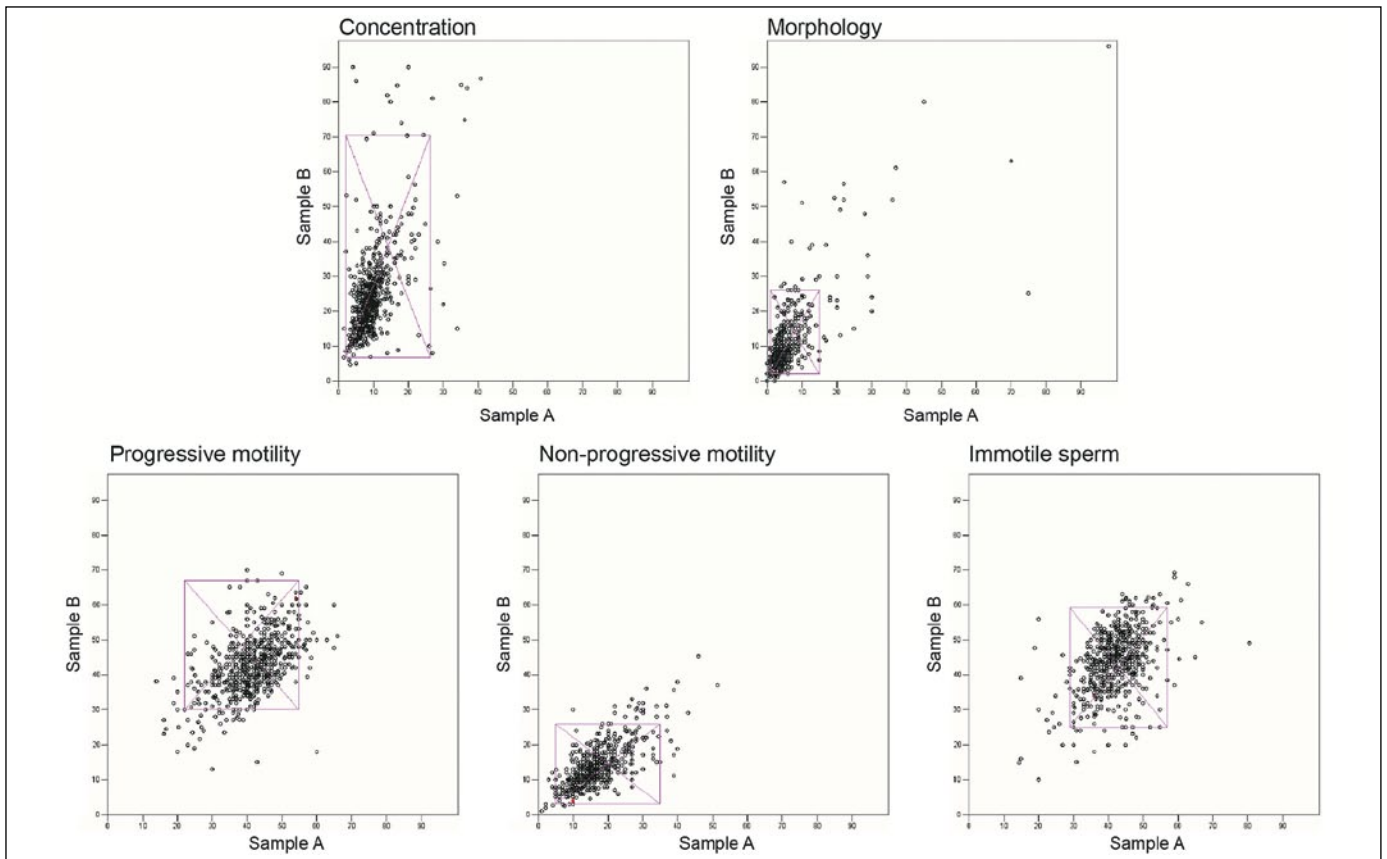


Figure 3. QuaDeGA ring trial 30/2016: Youden plots for sperm parameters in samples A and B of 663 participants constructed from medians of the WHO compliant participants (2.5 to 97.5 percentiles).

■ Methods

Ring trials

The QuaDeGA program has been described in previous publications [4, 5] and the reader is referred for details to these publications. In short, QuaDeGA carries out external quality control trials twice a year, consisting of two fixed sperm preparations (sample A and B). These samples allow for the measurement of sperm concentration as well as for preparation and staining of a smear for the assessment of morphologically normal forms. Film sequences of two different native semen samples are provided on the QuaDeGA platform for analysis of sperm motility. Each participant inserts the results on the online platform.

For the three parameters (1) sperm concentration, (2) normal morphology and (3) progressive motility medians of the values obtained by those laboratories indicating that they adhere strictly to WHO guidelines the 2.5 and 97.5 percentiles are used to construct Youden plots (Fig. 3). Those participants whose results for all parameters lie within the Youden plot windows receive certificates of passing the external quality control. In

addition, QuaDeGA provides a ranking for each result indicating whether results lie within the Youden plot (rank 1), or whether a systematic (rank 2) or a non-systematic (rank 3) or a random error (rank 4) has been noted.

Method used for the Survey

A questionnaire comprising 35 items concerning technical and methodological details of semen analysis as performed in the individual laboratories was drafted in German and sent electronically to participants using the online survey platform SurveyMonkey (Registered trade mark). The 624 participants in the ring trial 30/2016 in Germany, Austria and Switzerland were addressed, and a deadline of 3 weeks was set for returning answers. The questions are not described here in detail as their content becomes evident from the results in the next section.

■ Results

Responders

273 answers were received of which 256 (93.8%) could be evaluated. Data were saved and summarized by the online platform for further analysis. The 256

laboratories represent 42.5 % of those who had received the questionnaire and reflected the spectrum of participating laboratories (andrologists, urologists, ART centers, general clinical laboratories, hospitals, university clinics and private surgeries). Data were saved for further analysis with Microsoft Excel.

In order to find out whether the responders were biased concerning their performance in the ring trials tests, their results in run 30/2016 were compared with those from the non-responders. While 83% of the 624 addressed participants had received a certificate, 86% of the 256 responders and 82% of the 359 non-responders had received certificates, indicating that there was no significant difference in performance between responders and non-responders. As not all responders answered all questions, the number of replies varies from question to question. On average 7% of the questions were not answered.

Availability of the WHO Manual, Lab equipment and Techniques

When asked whether the WHO Manual was available in the individual labora-

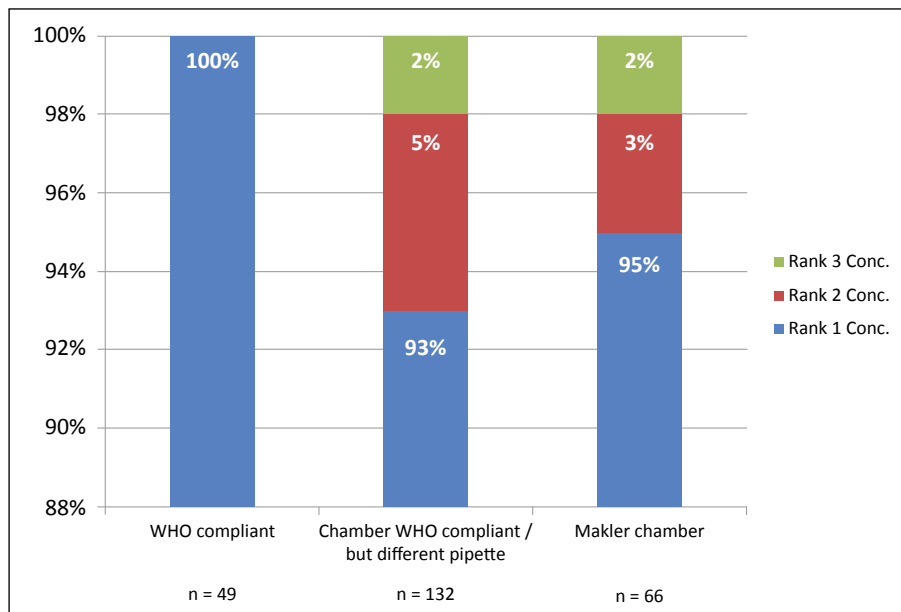


Figure 4. Impact of counting chambers and pipettes on determination of sperm concentrations (from 247 responses) as reflected by results from ranking. **Left:** using improved Neubauer chamber and positive displacement pipettes; **middle:** using improved Neubauer chamber, but no positive displacement pipettes; **right:** using Makler chamber.

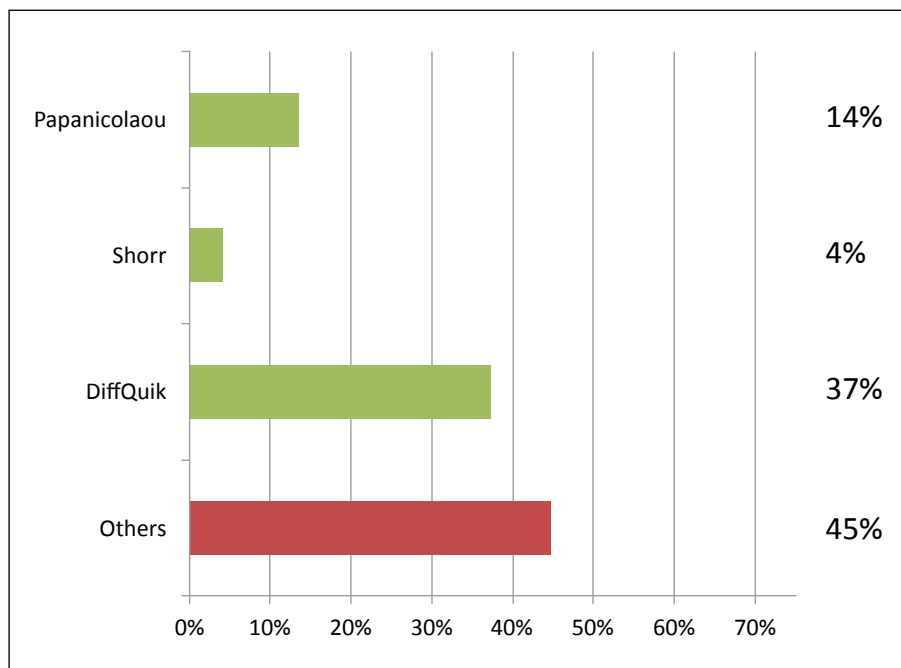


Figure 5. Different stainings used by 257 laboratories for sperm morphology. Green: WHO compliant; red: Not WHO recommended.

tory, 93% answered “Yes”. When asked whether they followed WHO guidelines for the individual parameters, 93% answered “Yes” for sperm concentration, 94% for motility and 90% for morphology.

However, when asked for the available basic equipment in their laboratory, 28% had no Vortex (or comparable mixing device), 45% no simple counting aid and 34% no laboratory counter. Of 254 answering laboratories 36% used a phase contrast microscope and 22% a high-

quality light microscope; the rest used medium (33%) or standard (9%) microscopes not considered sufficient by WHO guidelines.

Only 38.2% used the WHO recommended diluent, 23.1% no diluent and the remaining 38.7% used either NaCl, water or other media.

Only 59% of 249 laboratories used the WHO recommended improved Neubauer chamber for counting sperm. 27% used the Makler chamber, and the re-

maining 14% various but not recommended chambers.

Only 55 of 256 laboratories (22%) used the WHO recommended positive displacement pipettes for concentration. All 49 laboratories using the correct chamber and correct pipettes achieved rank 1, while only 93% of those using the correct chamber, but incorrect pipettes, and 95% of those using Makler chambers achieved rank 1 (Fig. 4).

Only 55% of 256 labs used the WHO recommended Papanicolaou, Shorr or DiffQuik staining method (Fig. 5). The 61 of 244 (25%) laboratories evaluating more than 200 sperm cells achieved the best results.

Internal Quality Control

The WHO manual considers it mandatory that all laboratories perform internal quality control and the BÄK provided Excel tables for this purpose which can be obtained from QuaDeGA upon request. However, only 35% perform their internal quality control by using these tables, while 42% use other programs for internal quality control. Surprisingly 23% of the 244 laboratories answering this question used no internal quality control. While 92% of those practicing internal quality control obtained a certificate, only 74% of those without internal quality control reached that goal.

Frequency of Semen Analysis

Investigating the impact of the number of semen analyses on the results of the individual laboratories, it became clear that those laboratories performing more than 20 semen analyses per month obtain certificates to a higher proportion than those laboratories with fewer than 20 semen samples (Fig. 6).

Discussion

Despite a battery of sophisticated sperm function tests, semen analysis remains the basic standard investigation to assess male fertility and infertility. Semen analysis plays a central role in the work-up of the infertile couple, but it is also important in toxicological, ecological and epidemiological studies. Recently, semen parameters were even found to reflect general health and – as a biomarker – to predict life expectancy [6].

In the light of this eminent role of semen parameters, it is surprising that semen analysis has long escaped quality assurance programs which are an obligatory exercise for all other measurements in the medical laboratory. Even the WHO manual (1st edition 1980) was not overly concerned with external quality control, and it was only in the 4th edition in 1990 that a small paragraph was dedicated to this topic. The 5th edition of 2010 [2, 3] expanded external quality control into a chapter. In parallel, several national and international external quality control programs have been developed in order to improve the validity and inter-laboratory comparability of results. In a few countries participation in external quality control programs became legally obligatory, and in Germany health insurances will not refund costs for semen analysis without a valid certificate from the quality control reference institution issued to the investigating laboratory.

Nevertheless, some sceptics continue to doubt the value of quality control and the adherence to generally accepted guidelines in order to guarantee reliable and reproducible results for sperm counts, motility and morphology [7, 8]. Others question the predictive value of sperm parameters in terms of chances for conception and pregnancy [9–13]. However, this remains an elusive discussion as long as the suggested and largely accepted guidelines are not strictly followed. With its manual, WHO provides such guidelines which have been accepted as the international standard, although they remain not undisputed.

As long as participation in an external quality control program for semen analysis was voluntary, we recognized that only 8% of laboratories participating in the QuaDeGA program adhere strictly to the WHO manual [5], and other external quality control schemes reported a similar low adherence to WHO guidelines [14–16]. Since external quality control and use of WHO guidelines became compulsory in Germany over 90% of labs claim to adhere to WHO guidelines, but as our current survey has shown, in reality the proportion of WHO followers is much lower. This failure to adhere to the guidelines ranges from inappropriate equipment and techniques to the lack of internal quality control. Until this situation changes and uniform methodology

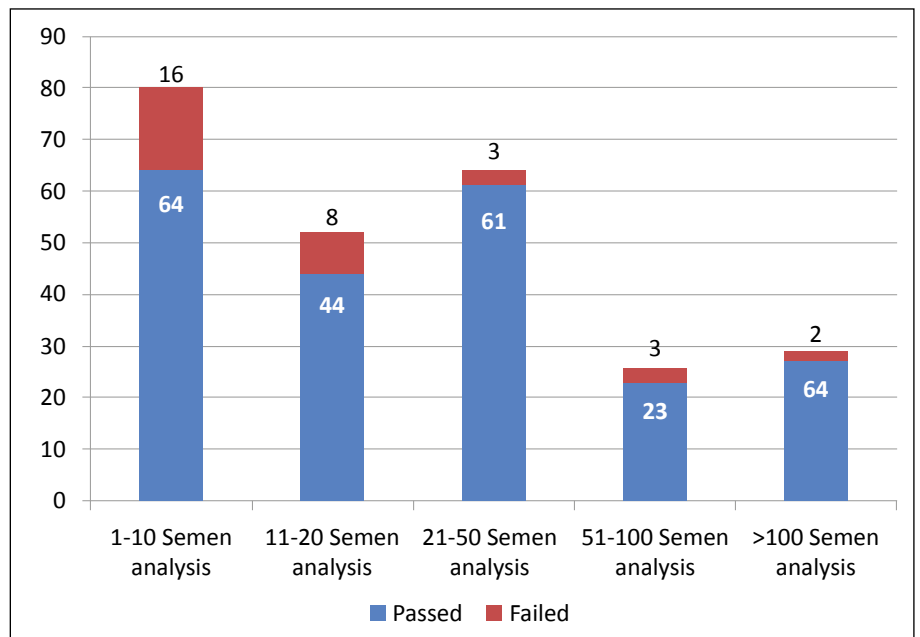


Figure 6. Impact of number of semen analyses per lab and month on awarding the certificate based on answers from 218 participants.

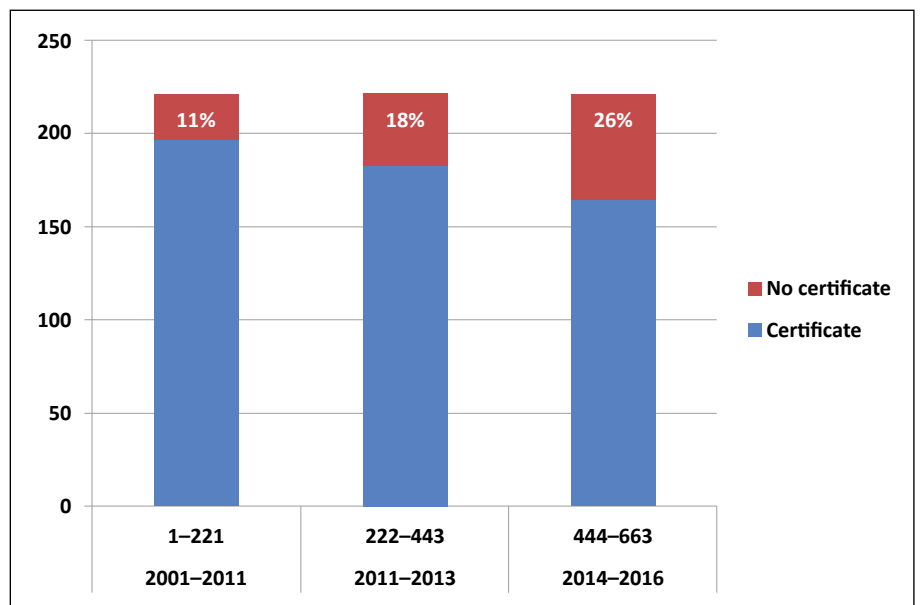


Figure 7. Performance in QuaDeGA ring trial 30/2016 with 663 laboratories divided into three groups of 221 each according to time of starting participation between 2002 and 2016.

is used, it will be impossible to judge the value of quality control programs and the predictive value of semen parameters as such. The high failure rate in obtaining the certificate by those not performing internal quality control, and high success rates of those using proper counting chambers and pipettes, provide visible examples of how adherence to guidelines can impact results positively. Also the impact of different staining techniques on the evaluation of sperm morphology has been well documented, as not only the chemicals used for coloring, but also the osmolarity of the solutions strongly influence sperm appearance and pre-

dispose to divergent results. In order to overcome this problem, use of only one staining technique to be used in all laboratories has been suggested [17, 18], but the editors of the WHO manual could not agree to such a strict requirement.

Only if all participants adhere to the same technical template, can the magnitude of an adherence problem of semen analysis be properly assessed. It is a fact that since the first discovery of sperm under the microscope of **Anthony Leeuwenhoek** in 1678, semen analysis has remained a subjective method depending on the training and the skills of

the investigator. As most semen analyses are performed by medical laboratory technicians it is deplorable that at most of their schools semen analysis is not included in their curriculum and most undergo training on the job when confronted with semen analysis. Therefore post-graduate semen analysis courses are of utmost importance for improvement of this situation [13]. In addition, participation in external quality control programs has an educational effect on its own [19], as demonstrated by an increasing rate of obtaining certificates with the duration of participating in the QuaDeGA ring trials (Fig. 7). In support of stricter use of guidelines, journals should request proof of proficiency by the laboratories submitting data from semen analysis [20].

Furthermore it is astonishing that despite 30 years of computer-assisted semen analysis (CASA) research, semen analysis remains a subjective method. Although technology has advanced to the extent that individual human faces can be identified among thousands of subjects, it remains a puzzle why sperm at low concentrations cannot be differentiated exactly from debris and sperm morphology cannot be recognized accurately by electronic means. Hopefully, once the necessary technology has been developed, all quality control problems will be resolved – or not, if it should then become evident that reproducibility and interlaboratory comparability of results from semen analysis depend on other factors intrinsic to the object under investigation.

Acknowledgements

The authors gratefully acknowledge cooperation with Dr. T. J. Panholzer for maintaining the QuaDeGA platform at the Institute for Medical Biometry, Epidemiology and Information Technology of the University Medicine Mainz, as well as language editing of the manuscript by Susan Nieschlag MA.

Conflict of Interest

The authors are employed part-time by the Quality Control Programme of the German Society of Andrology (QuaDeGA GmbH).

References:

1. BÄK. Richtlinie der Bundesärztekammer zur Qualitätssicherung laboratoriumsmedizinischer Untersuchungen: Ejakulatuntersuchungen. Dt Ärztebl 2011; 108: 2298–304.
2. World Health Organization (WHO). Laboratory manual for the examination and processing of human semen. 5th Edition, Geneva, 2010.
3. WHO Laborhandbuch zur Untersuchung und Aufarbeitung des menschlichen Ejakulates. Übersetzt von Nieschlag E, Schlatt S, Behre HM, Kliesch S (Hrsg.) unter Mitarbeit von Bongers R, Gottardo F, Greither T, Hellenkemper B, Nieschlag S, Nordhoff V, Schalkowski M, Zitzmann M. 5. Auflage, Springer Heidelberg, 2012.
4. Cooper T, Hellenkemper B, Nieschlag E. External Quality Control for Semen Analysis in Germany - Qualitätskontrolle der Deutschen Gesellschaft für Andrologie (QuaDeGA). The First 5 Years. J Reprod Med Endocrinol 2007; 4: 331–5.
5. Mallidis C, Cooper TG, Hellenkemper B, Lablans M, Ückert F, Nieschlag E. Ten years' experience with an external quality control program for semen analysis. Fertil Steril 2012; 98: 611–6.
6. Latif T, Kold Jensen T, Mehlsen J, Holmboe SA, Brinth L, Pors K, et al. Semen quality as a predictor of subsequent morbidity: A Danish cohort study of 4,712 men with long-term follow-up. Am J Epidemiol. 2017; 279:1–8.
7. Tomlinson MJ. Uncertainty of measurement and clinical value of semen analysis: has standardisation through professional

guidelines helped or hindered progress? Andrology 2016; 4:763–70.

8. Jequier AM. Is quality assurance in semen analysis still really necessary? A clinician's viewpoint. Hum Reprod 2005; 20: 2039–42.
9. Van Der Steeg JW, Steures P, Eijkemans MJC, Habbema JDF, Hompes PGA, Kremer JAM, et al. Role of semen analysis in subfertile couples. Fertil Steril 2011; 95:1013–9.
10. Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HWG, Behre HM, et al. World Health Organization reference values for human semen characteristics. Hum Reprod Update 2009; 16: 231–45.
11. Van Geloven N, Van Der Veen F, Bossuyt PMM, Hompes PG, Zwiderman AH, Mol BW. Can we distinguish between infertility and subfertility when predicting natural conception in couples with an unfulfilled child wish? Hum Reprod 2013; 28: 658–65.
12. Hamilton JAM, Cissen M, Brandes M, Smeenk MJM, De Bruin JP, Kremer JAM, et al. Total motile sperm count: A better indicator for the severity of male factor infertility than the WHO sperm classification system. Hum Reprod 2015; 30: 1110–21.
13. Barratt CLR, Björndahl L, Menkveld R, Mortimer D. ESHRE special interest group for andrology basic semen analysis course: A continued focus on accuracy, quality, efficiency and clinical relevance. Hum Reprod 2011; 26: 3207–12.
14. Penn HA, Windsperger A, Smith Z, Parekattil SJ, Kuang WW, Kolettis PN, et al. National semen analysis reference range reporting: Adherence to the 1999 World Health Organization guidelines 10 years later. Fertil Steril 2011; 95: 2320–3.
15. Björndahl L. Compliance with recommendations for reliable semen analysis results – a matter of importance for patients and scientific development. Andrology 2016; 4: 771–2.
16. Punjabi U, Wyns C, Mahmoud A, Vernelen K, China B, Verheyen G. Fifteen years of Belgian experience with external quality assessment of semen analysis. Andrology 2016; 4: 1084–93.
17. Gatimel N, Moreau J, Parinaud J, Léandri RD. Sperm morphology: Assessment, pathophysiology, clinical relevance, and state of the art in 2017. Andrology 2017; 1–18.
18. Meschede D, Keck C, Zander M, Cooper TG, Yeung CH, Nieschlag E. Influence of three different preparation techniques on the results of human sperm morphology analysis. Int J Androl 1993; 16: 362–9.
19. Brüggemann M. Qualitätssicherung laboratoriumsmedizinischer Untersuchungen. In: Jonitz G, Mansky T, Scriba PC (Hrsg). Ergebnisverbesserung durch Qualitätsmanagement. Deutscher Ärzteverlag, Köln, 2014; 109–114.
20. Carrell DT, De Jonge CJ. The troubling state of the semen analysis. Andrology 2016; 4: 761–2.