Standards in semen examination: publishing reproducible and reliable data based on high-quality methodology

Lars Björndahl, Christopher L.R. Barratt, David Mortimer, Ashok Agarwal, Robert J. Aitken, Juan G. Alvarez, Natalie Aneck-Hahn, Stefan Arver, Elisabetta Baldi, Lluís Bassas, Florence Boitrelle, Riana Bornman, Douglas T. Carrell, José A. Castilla, Gerardo Cerezo Parra, Jerome H. Check, Patricia S. Cuasnicu, Sally Perreault Darney, Christiaan de Jager, Christopher J. De Jonge, Joël R. Drevet, Erma Z. Drobnis, Stefan S. Du Plessis, Michael L. Eisenberg, Sandro C. Esteves, Evangelini A. Evgeni, Alberto Ferlin, Nicolas Garrido, Aleksander Giwercman, Ilse G.F. Goovaerts, Trine B. Haugen, Ralf Henkel, Lars Henningson, Marie-Claude Hofmann, James M. Hotaling, Piotr Jedrzejczak, Pierre Jouannet, Niels Jørgensen, Jackson C. Kirkman Brown, Csilla Krausz, Maciej Kurpisz, Ulrik Kviat, Dolores J. Lamb, Hagai Levine, Kate L. Loveland, Robert I. McLachlan, Ali Mahran, Liana Maree, Sarah Martins da Silva, Michael T. Mbizvo, Andreas Meinhardt, Roelof Menkveld, Sharon T. Mortimer, Sergey Moskowtsev, Charles H. Muller, Maria José Munuce, Monica Muratori, Craig Niederberger, Cristian O'Flaherty, Rafael Oliva, Willem Ombelet, Allan A. Pacey, Michael A. Palladino, Ranjith Ramasamy, Liliana Ramos, Nathalie Rives, Eduardo R. Roldan, Susan Rothmann, Denny Sakkas, Andrea Salonia, Maria Cristina Sánchez-Pozo, Rosanna Sapiro, Stefan Schlatt, Peter N. Schlegel, Hans-Christian Schuppe, Rupin Shah, Niels E. Skakkebæk, Katja Teerds, Igor Toskin.

†These authors contributed equally to this study.

Disclaimer: Some of the authors are present or former staff members of the World Health Organization. The authors alone are responsible for the views expressed in this publication and they do not necessarily represent the views, decisions, or policies of the institutions with which they are affiliated.

© The Author(s) 2022. Published by Oxford University Press on behalf of European Society of Human Reproduction and Embryology.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (https://creativecommons.org/licenses/by-nc/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com
ABSTRACT: Biomedical science is rapidly developing in terms of more transparency, openness and reproducibility of scientific publications. This is even more important for all studies that are based on results from basic semen examination. Recently two concordant documents have been published: the 6th edition of the WHO Laboratory Manual for the Examination and Processing of Human Semen, and the International Standard ISO 23162:2021. With these tools, we propose that authors should be instructed to follow these laboratory methods in order to publish studies in peer-reviewed journals, preferable by using a checklist as suggested in an Appendix to this article.

Key words: reproducibility / basic semen examination / standardized laboratory procedures / andrology / reproductive medicine / laboratory training / quality control / patient security / science development / journal requirements

Appeal to the scientific society involved in Andrology and Reproductive Medicine

As scientists are aware, there has been much discussion about the transparency, openness and reproducibility of science. This is not a new issue. Ten years ago, Begley and Ellis proposed a series of recommendations to improve the reliability of studies in preclinical cancer research (Begley and Ellis, 2012) that helped to initiate a series of developments to address and improve reproducibility. These have included more detailed reporting and transparency of methods such as the STAR Methods for Cell Press journals https://www.cell.com/star-authors-guide. Concomitant with these developments, national programmes, such as The MDAR (Materials Design Analysis Reporting) Framework, for transparent reporting in the life sciences have been launched (MacLeod et al., 2021) and specific consortia have been developed to repeat key published experiments, e.g. Reproducibility Project: Cancer Biology (RP: CB) (https://elifesciences.org/collec-tions/9b1e83d1/reproducibility-project-cancer-biology) (Rodgers and Collings, 2021). Furthermore, there are significant resources available such as EQUATOR guidelines (https://www.equator-network.org/).

The clear direction of travel is to improve standards and have transparent reporting of research (Amara, 2022). There are challenges, however. For example, in the RP:CB project, insufficient information was available to repeat selected experiments published in high impact journals. Furthermore, in the experiments that could be repeated (50/193), fewer than half yielded similar results. As such, the final report of the RP: CB consortia suggested that ‘it is hard to assess whether reported findings are credible’ (Errington et al., 2021).

In our own discipline of Andrology and Reproductive Medicine, there is a plethora of evidence to show that using non-standardized methods produces unreliable data including, for example, for human sperm concentration and sperm motility assessments. This has created significant problems for the field, making it difficult to determine the scientific accuracy of many studies and ultimately establish their real clinical and public health impact. A recent example of this is the study of Campbell et al. where they updated the World Health Organization (WHO) semen analysis distribution values (Campbell et al., 2021). The authors reported several challenges in obtaining key information of the quality of the semen examination methods used across the studies being considered for inclusion. Standardization of semen examination has been a long-standing issue that the profession has collectively failed to address, despite the availability of proven accurate methods and robust training approaches (Björndahl et al., 2002, 2016; Barratt et al., 2011; Carrell and De Jonge, 2016; Cairo Consensus Workshop, 2020). Too many studies depending on semen analysis derived data continue to demonstrate poor robustness and rigour in semen analysis methodology (Serrano et al., 2014). When methods with a high degree of uncertainty are used, differences between normal and pathological conditions are likely to be impossible to discover since each observation, burdened by large variability due to measurement uncertainty, will have a more-or-less random result. This will cause considerable overlap in results from the different populations, making them practically inseparable.

The question for all of us working in Andrology, including Editors of journals publishing research in this field, is: What can be done to improve the situation? We believe there is currently a window of opportunity for action. The recent publication of ISO Standard 23162 for the basic examination of human semen (International Organization for Standardization, 2021) finally means that the field has de facto reference methods. These methods form the basis of those recommended in the new 6th edition of the WHO andrology laboratory manual (World Health Organization, 2021), which contains simple to follow and proven high-quality methods for semen examination. We propose that authors should be instructed to follow these laboratory methods in order to publish studies in peer-reviewed journals. To facilitate this, we present in the Appendix an author checklist, modified from Björndahl et al. (2016), which authors can complete and submit with their manuscript, making it simple for the journals, reviewers and readers alike to assess the quality of the semen assessment methods used,
and hence of the data being reported. We suggest that any deviation from the checklist, for purposes of testing a new reagent, different method or procedure for improving on the performance of a current recommendation, should be specified and measured against those in the checklist. If authors did not follow these methods, a separate section of the Materials and Methods should specify what differed and why, and how the variations might have impacted the accuracy of results. In other disciplines, checklists have assisted with improving the reporting of results (Nature, 2018; NPQIP Collaborative Group, 2019). This approach is consistent with the TOP Guidelines (Transparency and Openness Promotion; Centre for Open Science, https://www.cos.io/initiatives/top-guidelines; Nosek et al., 2015).

This is an important initiative. We suggest it be implemented by all journals in our discipline to help improve the transparency, openness and reproducibility of science.

Supplementary data
Supplementary data are available at Human Reproduction online.

Data availability
No new data were generated or included in the manuscript.

Authors’ roles
L.B., C.B. and D.M. outlined the first manuscript and contacted all other authors for comments. L.B., C.B. and D.M. summarized all suggestions and finalized the manuscript that all authors have received and confirmed their participation in.

Funding
No specific funding was used for this publication.

Conflict of interest
This is an Opinion article based on the necessity for improvement of standards in the field of Andrology and Reproductive Medicine, based on the two non-profit publications, by the World Health Organization and the International Organization for Standardization. No conflicts of interest are declared. CB, as an employee of the University of Dundee, serves on the Scientific Advisory board of ExSeed Health (from September 2021, financial compensation to the University of Dundee) and is a scientific consultant for Exscentia (from September 2021, financial compensation to the University of Dundee). CB has previously received a fee from Cooper Surgical for lectures on scientific research methods outside the submitted work (2020) and Ferring for a lecture on male reproductive health (2021). CB is Editor for Reproductive Bio Medicine Online. DL, as an employee Weill Cornell Medicine, declares: American Board of Bioanalysis (Secretary-Treasurer: Honorarium); Fellow Health (Equity); Roman Health (Consultant and Advisory Board) (Equity and compensation).

References
Amara SG. Empower with evidence. Science 2022;375:699.
Barratt CL, 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;26:3207–3212.
Begley CG, Ellis LM. Drug development: raise standards for preclinical cancer research. Nature 2012;483:531–533.
Björndahl L, Barratt CL, Fraser LR, Kvist U, Mortimer D. ESHRE basic semen analysis courses 1995–1999: immediate beneficial effects of standardized training. Hum Reprod 2002;17:1299–1305.
Björndahl L, Barratt CL, Mortimer D, Jouannet P. ‘How to count sperm properly’: checklist for acceptability of studies based on human semen analysis. Hum Reprod 2016;31:227–232.
Cairo Consensus Workshop. The current status and future of andrology: a consensus report from the Cairo workshop group. Andrology 2020;8:27–52.
Campbell MJ, Lotti F, Baldi E, Schlatt S, Festin MPR, Björndahl L, Toskin I, Barratt CLR. Distribution of semen examination results 2020—a follow up of data collated for the WHO semen analysis manual 2010. Andrology 2021;9:817–822.
Carrell DT, De Jonge CJ. The troubling state of the semen analysis. Andrology 2016;4:761–762.
Errington TM, Denis A, Perfito N, Iorns E, Nosek BA. Challenges for assessing replicability in preclinical cancer biology. Elife 2021;10:e67995.
International Organization for Standardization. ISO 23162:2021 Basic Semen Examination — Specification and Test Methods. Geneva: ISO, 2021.
Macleod M, Collings AM, Graf C, Kiermer V, Mellor D, Swaminathan S, Sweet D, Vinson V. The MDAR (Materials Design Analysis Reporting) Framework for transparent reporting in the life sciences. Proc Natl Acad Sci USA 2021;118:e2103238118.
Nature. Checklists work to improve science. Nature 2018;556:273–274.
Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, Buck S, Chambers CD, Chin G, Christensen G et al. SCIENTIFIC STANDARDS. Promoting an open research culture. Science 2015;348:1422–1425.
NPQIP Collaborative Group. Did a change in Nature journals’ editorial policy for life sciences research improve reporting? BMJ Open Sci 2019;3:e000035.
Rodgers P, Collings A. Reproducibility in Cancer Biology: What have we learned? Elife 2021;10:
Serrano M, Gonzalvo MC, Sanchez-Pozo MC, Clavero A, Fernandez MF, Lopez-Regalado ML, Mozas J, Martinez L, Castillo JA. Adherence to reporting guidelines in observational studies concerning exposure to persistent organic pollutants and effects on semen parameters. Hum Reprod 2014;29:1122–1133.
World Health Organization. WHO Laboratory Manual for the Examination and Processing of Human Sperm, 6th edn. Geneva: World Health Organization, 2021.
Appendix

Semen analysis methodology checklist for authors, reviewers and editors (modified from Bjornsdahl et al., 2016).

An accessible version of the checklist is available as a supplementary file. This checklist is based on the ISO Standard on basic semen examination and the current WHO recommendations ("WHO6") [1,2], and on general scientific standards. Full compliance requires that all boxes are ticked.

A deviation from this checklist does not necessarily mean that the study cannot be published, but all deviations must be declared in the Materials and Methods section of the manuscript, including their impact on accuracy and measurement uncertainty of the data, in order to allow the reader to evaluate the quality of the analyses performed. For studies not reporting all characteristics of a basic semen examination, the checklist includes the option 'Not applicable to the study'.

Investigations that would be subject to the requirements in this checklist can roughly be classified as clinical (evaluating patient treatment, diagnostic classification or predictive powers of certain assessments), experimental (e.g. exposure of spermatozoa to different compounds or in vitro treatments [3,4]) or epidemiological (evaluating variations in semen characteristics or effects of exposure populations to certain compounds or other circumstances).

Any scientific rationale for not complying with the guidelines, which is not included in the Materials and Methods section of the manuscript, must be substantiated to the Editor and Reviewers.

References

WHO Laboratory Manual for the Examination and Processing of Human Semen, 6th edn. Geneva: World Health Organization, 2021.
ISO 23162:2021 Basic Semen Examination – Specification and Test Methods. Geneva: International Standards Organization, 2021.
Mortimer D, Barratt CLR, Bjornsdahl L, de Jager C, Jequier AM, Muller CH. What should it take to describe a substance or product as 'sperm-safe'. Hum Reprod Update 2013;19 Suppl 1:i1–i45.
Sanchez-Pozo MC, Mendiola J, Serrano M, Mozas J, Bjornsdahl L, Menkveld R, Lewis SEM, Mortimer D, Jorgensen N, Barratt CLR et al.; on behalf of the Special Interest Group in Andrology (SIGA) of the European Society of Human Reproduction and Embryology. Proposal of guidelines for the appraisal of SEMen QUALity studies (SEMQUA). Hum Reprod 2013;28:10–21.

1. Patients

- Not applicable to the study
  - 1.1 The studied population (e.g. patients or volunteers) has been declared in the manuscript, together with the recruitment method and inclusion and exclusion criteria. In a study concerning couples being investigated for infertility, the following is specified in the manuscript: fertility status of female partner; and primary, secondary or other level of investigation of the man.
  - 1.2 If used in the manuscript, the term ‘male factor’ is completely defined.

- 1.3 Reference limits provided in WHO5 or 5th percentile of distribution of semen examination results in WHO6 have not been used to characterize subjects as infertile, subfertile or fully fertile.

2. General aspects

- Not applicable to the study
  - 2.1 Patients were instructed to maintain 2–7 days of sexual abstinence before collecting an ejaculate for investigation.
  - 2.2 Patients were informed about the importance of reporting any missed early ejaculate fractions, and their responses were noted on the laboratory record.
  - 2.3 For specimens not collected at the laboratory, patients were instructed to avoid cooling (under 20 °C) or heating (above 37 °C) of the semen specimen during transport to the laboratory.
  - 2.4 In the laboratory, specimens were kept at 37 °C before initiation of and during the analysis in case of sperm motility assessment.
  - 2.5 For specimens collected adjacent to the laboratory, analysis was initiated after completion of liquefaction and within 30 min after ejaculation. If some of the specimens were collected at the laboratory and others collected at home the influence on the data is declared and discussed in the manuscript.
  - 2.6 Liquefaction was first checked within 30 min after ejaculation.
  - 2.7 Volume was determined either by weighing or using a wide-bore volumetric pipette.
  - 2.8 Viscosity was measured using either a wide-bore pipette or a glass rod.
  - 2.9 All staff members who performed the analyses have been trained in basic semen analysis (ESHRE Basic Semen Examination Course—or equivalent—with further in-house training to establish competency), and participate regularly in internal quality control.
  - 2.10 When more than one method is recommended for a particular characteristic (e.g. to measure volume), only one was used in the study. For a multicentre study, all laboratories used the same method or variable methods are declared in the manuscript.

3. Sperm concentration assessment

- Not applicable to the study
  - 3.1 Semen aliquot to be diluted for sperm concentration assessment was taken with a positive displacement pipette (i.e. a 'PCR pipette') using a recommended diluent (state which diluent: ).
  - 3.2 Only standard dilutions were used (1:50, 1:20 or 1:10, i.e. 1 + 49, 1 + 19 or 1 + 9).
  - 3.3 Sperm concentration was assessed using haemocytometers with improved Neubauer ruling.
  - 3.4 Haemocytometers were allowed to rest for 10–15 min in a humid chamber to allow sedimentation of the suspended spermatozoa onto the counting grid before counting.
  - 3.5 Sperm counting was done using phase contrast microscope optics (200–400 ×).
  - 3.6 Comparisons were made between duplicate counts, and counts were re-done when the difference exceeded the acceptance limits.
  - 3.7 Typically at least 200 spermatozoa were counted in each of the duplicate assessments.
4. Sperm motility assessment
☐ Not applicable to the study
☐ 4.1 Motility assessments were performed at 37 °C ± 0.5 °C.
☐ 4.2 Motility assessments were initiated within 30–60 min after sample collection.
☐ 4.3 Motility assessments were performed using phase contrast microscope optics (200–400×).
☐ 4.4 Sperm motility was classified using a four-category scheme: rapid progressive, slow progressive, non-progressive, and immotile.
☐ 4.5 Motility assessments were done in duplicate and compared; counts were re-done on new preparations when the difference between duplicates exceeded the acceptance limits.
☐ 4.6 The wet preparation was made using a drop of ___ ml and a ___ × ___ mm coverslip to give a depth of ___ μm (must be at least 10 μm depth, but not too deep so as to allow spermatozoa to move freely in and out of focus; typically ca. 20 μm).
☐ 4.7 At least 200 spermatozoa were assessed in each duplicate motility count.
☐ 4.8 At least five microscope fields of view were examined in each duplicate count.

5. Sperm vitality assessment
☐ Not applicable to the study
☐ 5.1 A validated supravital stain, appropriate to the type of microscope optics employed, was used to assess sperm vitality.
☐ 5.2 At least 200 spermatozoa were evaluated.
☐ 5.3 Assessments were done under high magnification (×1000–1250) using a 100× high resolution oil immersion objective and bright field microscope optics (Köhler illumination).

6. Sperm morphology assessment
☐ Not applicable to the study
☐ 6.1 Tygerberg Strict Criteria were used for the evaluation of human sperm morphology. (Another classification could be used for scientific studies with specific aims if the classification is described or referenced. Depending on the aim of the study, the evaluation of particular abnormal forms might be useful.)
☐ 6.2 Abnormalities are recorded for the four defined regions of the spermatozoon (head, neck/midpiece, tail and cytoplasmic residue).
☐ 6.3 The Papanicolaou staining method adapted for the assessment of human sperm morphology was used. For specific aims, other staining methods could be used but must then be declared and explained.
☐ 6.4 At least 200 spermatozoa were assessed in each ejaculate.
☐ 6.5 Assessments were done under high magnification (×1000–1250) using a 100× high resolution oil immersion objective and bright field microscope optics (Köhler illumination).

7. External Quality Assessment (EQA)
☐ 7.1 The laboratory participated in EQA for the semen examination methods used to obtain data for the study.
☐ 7.2 Name of the EQA scheme: ____________________________

8. Other findings
☐ Not applicable to the study
☐ 8.1 The presence of abnormal clumping (aggregates and agglutinates) was recorded.
☐ 8.2 Abnormal viscosity was recorded.

9. Analysing data
☐ Not applicable to the study
☐ 9.1 The actual duration of sexual abstinence (in ‘hours’ or ‘days’) was recorded for each specimen and included in the data reported in the manuscript.
☐ 9.2 As a minimum in clinical studies, semen volume, sperm concentration, total number of spermatozoa per ejaculate and abstinence time are given to reflect sperm production and output; only samples identified as having been collected completely were included in the study.
☐ 9.3 Confounding factors have been considered for statistical analysis: e.g. abstinence time and age, to consider secular or geographical variations in sperm concentration or sperm count.
☐ 9.4 If appropriate, optional biochemical markers for prostatic, seminal vesicular and epididymal secretions were analysed and reported, both as concentration and total amount.
☐ 9.5 Signs of active infection/inflammation were noted and considered in the analysis of data in the study (e.g. presence of non-germ line round cells, inflammatory cells, impaired sperm motility, possibly also anti-sperm antibodies or reduction of secretory contributions).

10. Data Repository
☐ 10.1 For the sake of transparency, all data without identification of individual patients or research participants have been saved to a trusted online repository, and there is a statement of this in the Results section of the manuscript.

Declaration by the corresponding author:
The information provided in this checklist is solemnly declared to be true.
Signature: _______________________ Date: ________________
Name: ______________________________________________
Affiliation: ____________________________________________