Continued existence of significant disparities in the technical practices of sperm morphology assessment and the clinical implications: results of a French questionnaire

Nicolas Gatimel, Pharm.D., Ph.D.,[a,b] Lucile Mansoux, M.Sc.,[b] Jessika Moreau, M.D.,[a,b] Jean Parinaud, M.D., Ph.D.,[a,b] and Roger D. Léandri, M.D., Ph.D.[a,b]

[a] Department of Reproductive Medicine; and [b] EA 3694 Human Fertility Research Group, Paule de Viguier Hospital, Toulouse Teaching Hospital Group, Toulouse, France

Objective: To investigate current practices for sperm morphology assessment.
Design: E-mail survey questionnaire.
Setting: Diagnostic and clinical institutions/laboratories.
Patient(s): French biologists and clinicians (n = 225).
Intervention(s): None.
Main Outcome Measure(s): Answers to 24 questions.
Result(s): The survey shows that even now in France there is great inconsistency in the use of sperm morphology assessments. For example, the survey revealed that no fewer than six different staining techniques were in use. Automatic reading is hardly used (used by 2.5% of the biologists replying to the survey). More than 33.6% of biologists use a threshold of normal forms unsuited to their classification, and 20% do not perform any internal quality control in this area. Prescribing doctors seldom trust the tests, likely due to their lack of analytic reliability. Among the biologists surveyed 26% said the percentage of normal forms is either unreliable or not very reliable in analytic terms, and 24% of clinicians stated that it has little clinical relevance.
Conclusion(s): The survey reveals a marked lack of uniformity in French laboratories for performing sperm morphology assessment and in the use of the results by physicians. Regular quality control procedures and well-trained personnel, up to date with their training and conversant with the latest techniques as well as harmonized practices, are clearly indispensable. It is time for a consensus on the practice and interpretation of this particular test. (Fertil Steril® 2017;107:365–72. ©2016 by American Society for Reproductive Medicine.)
Key Words: Analytical reliability, clinical relevance, quality control, sperm morphology

Discuss: You can discuss this article with its authors and with other ASRM members at https://www.fertstertdialog.com/users/16110-fertility-and-sterility/posts/12873-22965

Assessing the morphology of spermatozoa using stained slides and standard microscopic magnification (×100) has been routine for more than 30 years in many countries. A change in the morphology of spermatozoa may be associated with such sperm malfunctions as modifications in chromatin conformations, defects in the acrosome reaction, disorders in flagellar motility, or even an increase in the rate of apoptosis and necrosis (1, 2). The definition of a morphologically normal spermatozoon is based on observations of spermatozoa found in cervical mucous and is supported by observations of spermatozoa bound to the zona pellucida (3). Menkveld et al. (1) proposed a strict definition of a morphologically normal spermatozoon. In many studies, the percentage of abnormal sperm was found to be a strong indicator for male fertility (4, 5) and for fertilization rates in an in vitro fertilization (IVF) attempt (6). Its

Received August 22, 2016; revised September 25, 2016; accepted October 24, 2016; published online November 22, 2016.
N.G. has nothing to disclose. L.M. has nothing to disclose. J.M. has nothing to disclose. J.P. has nothing to disclose. R.D.L. has nothing to disclose.
Reprint requests: Nicolas Gatimel, Pharm.D., Ph.D., 330 avenue de Grande Bretagne, Toulouse Cedex 9 31059, France (E-mail: gatimel.n@chu-toulouse.fr).

Fertility and Sterility® Vol. 107, No. 2, February 2017 0015-0282/$36.00

http://dx.doi.org/10.1016/j.fertnstert.2016.10.038

predictive value for successful intrauterine insemination (IUI) (7–9) is very controversial, and in the majority of assisted reproductive technology (ART) centers, sperm morphology is no longer an important factor in clinical decisions.

Since this test has become commonplace in laboratories, many investigators have regularly alerted the scientific community to the widespread variability, both intralaboratory and interlaboratory, in its practical implementation. They stress the need for harmonizing techniques for preparation and interpretation of the slides and the classifications used, the importance of performing quality controls, and the implementation of program for training and skill maintenance for the technicians (10–22).

The manner in which sperm morphology should be assessed remains a subject of debate. The lack of standardization of the method and the subjective nature of sperm morphology assessment make it very difficult to compare results with the World Health Organization (WHO) reference values (23) and also between laboratories. The debate on differences in the clinical values among the different classifications (traditional or liberal approach, strict criteria, David classification, etc.) has raged for years and is certainly not over (23–25). Prescribing clinicians appear have less and less faith in the test, and the lack of analytic reliability may be one of the reasons.

French diagnostic laboratories are now obliged to be ISO/EN 15189 accredited, so it is an appropriate time to assess practices concerning the use of the sperm morphology assessment by surveying the clinicians and biologists using this test in France. The purpose of the questionnaire was twofold. First, we wanted to learn the various methods being used to prepare and read the slides and recording the results in French laboratories. Second, we evaluated how clinicians (gynecologists and andrologists) have used the results of this test in their diagnoses and ART decisions.

MATERIALS AND METHODS

The anonymous questionnaire used Google Form and was distributed by e-mail in April 2016 to biologists and clinicians in France who are operating in private or academic practice and are involved in infertility treatment. The exhaustive list of addresses for doctors performing ART was obtained from the French Agency for Biomedicine (ABM for Agence de la biomédecine). The survey was also disseminated using distribution lists for appropriate learned societies, specifically for gynecologists and andrologists, including the Fédération des biologistes de la fécondation et de la conservation de l'œuf (BLEFCO), Société d'andrologie de la langue française (SALF), and Collège des gynécologues du Midi. The questionnaire had 24 questions of which three were common to all addressees; the others were specific to biologists or to practicing clinicians (Supplemental Fig. 1, available online). Data analyses were performed using descriptive statistics in Excel (Microsoft) on the database produced using Google Form. Institutional review board approval was not necessary because the study concerned a survey to biologists and clinicians to assess their medical practices without any impact on patients during the study.

RESULTS

Questionnaire Participant Characteristics

We received 225 responses from a total of 535 questionnaires distributed (42%): 106 clinicians (85 gynecologists and 21 andrologists) and 119 biologists. In addition to infertility assessment (diagnosis), 75.5% of the clinicians who responded practice an ART activity: 20% do only intrauterine insemination (IUI) and 55% do both IUI and IVF within an ART center. Among the biologists, 95% are practicing ART: 29% do only sperm preparation for IUI in their laboratory, and 65% do both IUI and IVF within an ART center.

Specific questions for biologists. Table 1 shows the great disparity of laboratory practices. The majority (73%) of French laboratories use the David classification. It is noteworthy that for a given classification the threshold for normalcy may vary between 4% and 30% depending on the laboratory (Fig. 1), even though 92% claim that their limit value was found in the literature. Although a majority perform intern quality control procedures at least twice a year, close to 20% never perform such activities.

Specific questions for clinicians (gynecologists and andrologists). Ninety-five percent of the clinicians surveyed prescribe a sperm morphology evaluation systematically for an infertility assessment; 5% prescribe it on an exceptional basis for an andrological evaluation, and 22% before performing ART. Thirty-five percent of clinicians do not always use the same laboratory, and 18% are not aware of the classification used. Fifty-two percent take into account specific anomalies, but the type of anomaly is extremely variable: for 64%, a defect of the head, for 21% an anomaly of the tail, and for 12% discrepancies in the acrosome. Supplemental Table 1 (available online) shows that only 30% of clinicians regularly use an index of multiple sperm defects: teratozoospermia index (TZI), multiple anomalies index (MAI), or sperm deformity index (SDI). French practitioners mainly use the multiple anomalies index (MAI) which is used with David's classification.

Perception of the Value of a Sperm Morphology Assessment and its Thresholds

Among the biologists surveyed, 26% believe that the percentage of normal forms is either unreliable or not very reliable in analytic terms, and 24% of clinicians state that it has very little clinical relevance (Fig. 2).

Using the Percentage of Normal Forms before ART

As shown in Supplemental Table 2 (available online), 55% of clinicians and 40% of biologists regularly use the percentage of normal spermatozoa to choose the ART technique. Moreover, when used, the threshold varies widely (Fig. 3).

DISCUSSION

In France, huge disparity exists in the techniques used for preparing, staining, and interpreting slides in this area. Indeed, 23% of the biologists in France who completed the survey

TABLE 1

Description of the different methods used to prepare and read the slides and interpret the results.

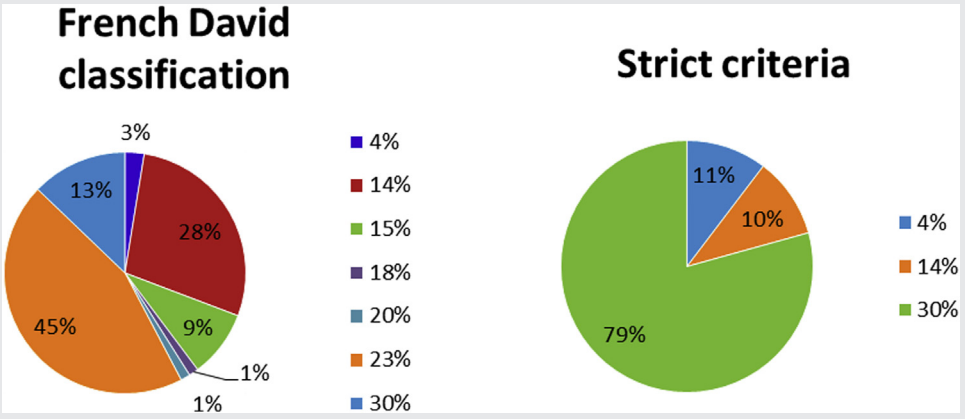
Sperm washing before smears		Air drying		Staining method		Type of reading		Classification		No. of examined spermatozoa	
Never	71.4% (85)	Yes	91.6% (109)	Papanicolaou	11.1% (13)	Visual	97.5% (116)	David's	73.1% (87)	100	68.6% (35)
				Shorr	59.8% (70)						
In case of severe oligospermia or increased viscosity	21.8% (26)			Diff-Quik	6% (7)					200	29.4% (15)
		No	8.4% (10)	RAL	14.5% (15)	CASA	2.5% (3)	Strict criteria (WHO)	26.9% (32)		
Always	6.7% (8)			Other	8.6% (10)					>200	2% (1)

Reference values (threshold) used for normal forms		Determination of the threshold		Details of anomalies		No. of people performing evaluation of sperm morphology in laboratory		Frequency of internal quality controls		Frequency of external quality controls	
30%	13.4% (16)	Published data	91.5% (108)	Always	91% (108)	Mean	3.7	0/y	20.3% (24)	0/y	9.5% (11)
23%	31.1% (37)					Median	3	1/y	13.6% (16)	1/y	3.4% (4)
14%	25.2% (30)			On request	1% (1)	Minimum	1	2/y	18.6% (22)	2/y	61.2% (71)
4%	21% (25)	Own study	8.5% (10)			Maximum	12	3/y	8.5% (10)	3/y	5.2% (6)
Other	9.2% (11)			Never	8% (10)			3–5/y	15.3% (18)	3–5/y	15.5% (18)
								>5/y	23.7% (28)	>5/y	5.2% (6)

Note: CASA = computer-assisted sperm analysis.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.

FIGURE 1

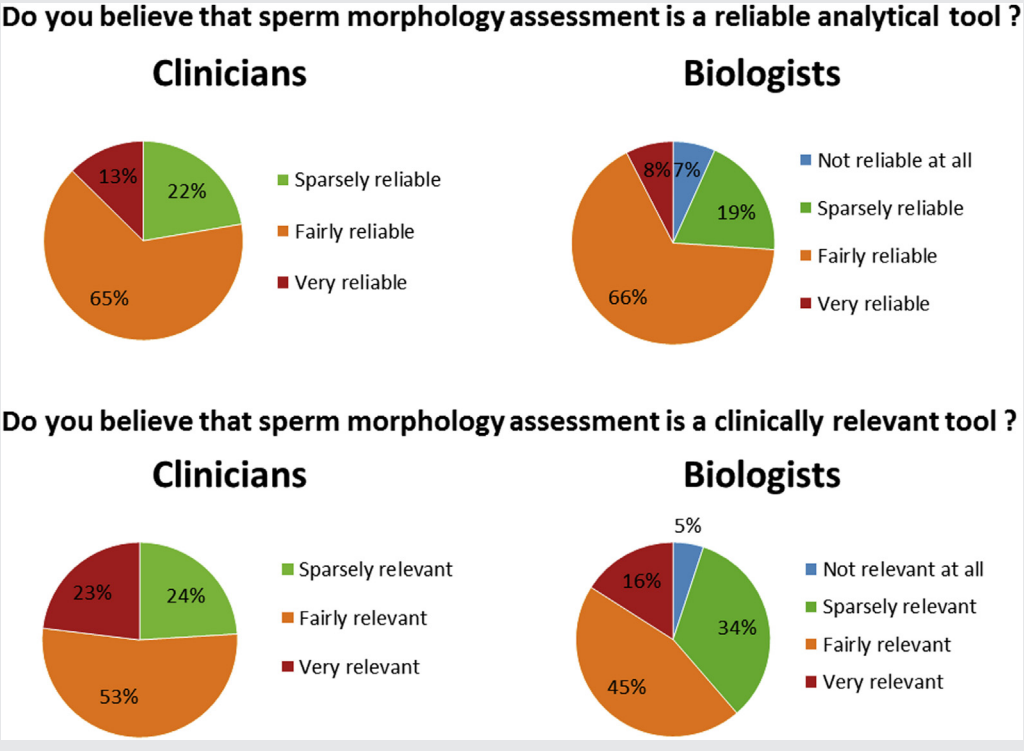


Reference thresholds for the percentage of normal forms used by the French diagnostic laboratories questioned as a function of the classifications used.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.

do not use the staining techniques recommended by the WHO (Papanicolaou, Diff-Quik, and Shorr). Most stains modify the size of the cell but in different ways depending on the stain used, largely as a result of the different osmolarities (Van der Horst 2015). According to Mortimer and Menkveld (26), Papanicolaou staining is the recommended method for the best morphologic assessment in routine applications. Although no statistically significant difference has been observed when comparing the Papanicolaou and the Diff-Quik staining methods (27, 28), Natali et al. (29) revealed

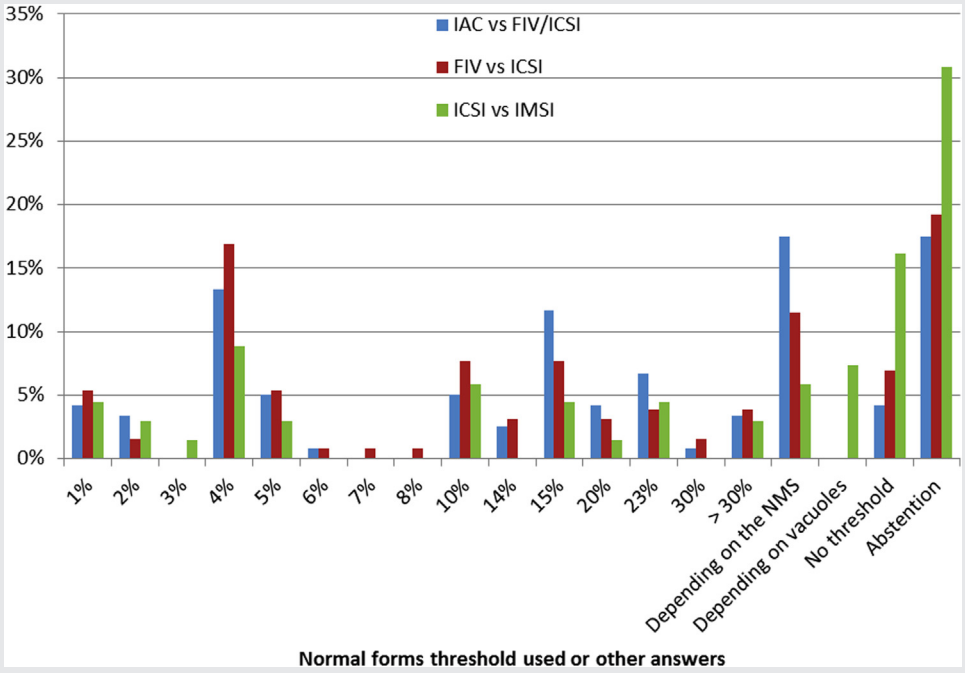
FIGURE 2



Perception of clinical relevance and analytic reliability as seen by physicians and biologists.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.

FIGURE 3



The split of the various normal forms thresholds used by clinicians and biologists when choosing an assisted reproductive technique. NMS = number of motile spermatozoa after sperm preparation.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.

very different results when comparing Diff-Quik and Testsimplets TS. This underlines the importance of using the same staining when making interlaboratory comparisons or when using threshold decisions in an ART situation, and also the need to use the reference values published in the literature correctly.

Automatic evaluation is hardly used in France (used by just 2.5% of the biologists replying to the survey). When using automatic reading, the effects of any operator influence on the assessment of the cell are reduced, although certain computer-assisted sperm analysis (CASA) systems require making choices about the zones to be studied and where to focus on the cell. In addition, automatic systems allow archiving for pictures and videos, thus providing excellent material for personnel training sessions. Such systems would seem to offer better reproducibility both intraoperator and interindividual (30, 31) even though intraoperator and interoperator variability problems were reported with some of the earlier automatic systems (28). In this latest study, the investigators emphasized that the coefficients of variation were very high (from 40% to 60%) for both manual and automated evaluations. Nevertheless, training in automatic systems should not be neglected, even though it is simpler than training for manual sperm analysis. In France, the number of spermatozoa analyzed per patient is most often 100 (69% of biologists surveyed), which is contrary to the latest guidelines from the WHO that recommend using at least 200 over two separate evaluations. The lack of compliance with this recommendation by French laboratories is probably related to the time-consuming nature of manually interpreting and evaluating 200 items. In this area there has been no change in 20 years compared with the practices recorded in 1997 (22).

Our survey shows that in France most biologists (73%) use the David classification and that there exists an enormous disparity in the various reference thresholds used for the percentage of normal forms. David's French classification considers all borderline aspects to be normal whereas according to strict criteria all borderline forms should be considered abnormal (1). The World Health Organization (2010) has recommended the use of strict sperm morphology to identify normal spermatozoa.

In the 2010 WHO recommendations (32), the lower reference limit is 4%, based on the fifth percentile of the combined data from recent publications according to known and standardized methodologies (33). This threshold value is in agreement with cutoff values in studies that compared the semen parameters of a fertile and a subfertile population and determined the threshold values with a receiver operating characteristic curve analysis (23, 34, 35). In the modified David's classification, the threshold values were reassessed by Auger et al. (36) as 23% in a study that reported the distribution of various head, midpiece, and principal piece defects of spermatozoa in a large group of fertile men (male partners of pregnant women).

Surprising data from this survey are that, among laboratories that cite using values of normal forms based on the published data in the scientific literature, 33.6% use

a threshold of normal forms unsuited to their classification (e.g., 14% with David's classification) (see Fig. 1). Some have suggested that each laboratory should ideally have its own standards (19), in view of the wide variations in techniques and operator interpretation and the significant diversity of methods used from one laboratory to another. However, this practice is rare among biologists in France (8.5%), probably due to the difficulty of obtaining a reference population.

For the vast majority (91%) of biologists, the count of the various anomalies is handled systematically, whereas 48% of the clinicians surveyed take no account of this. The clinicians' lack of interest in the details of anomalies is probably linked to the limited clinical relevance of the various anomalies, with the exception of the very rare monoamorphous anomalies, which are often linked to genetic disorders (globozoospermia, macrocephalia, decapitated sperm syndrome, and fibrous sheath dysplasia). The frequency of each morphologic abnormality is statistically significantly higher in infertile men than fertile men (36). The 2010 WHO guidelines do not recommend systematic recording of the details of anomalies. In addition, the main problem in evaluating such abnormalities is the high degree of variability between operators (15), which probably explains the lack of faith in this detailed evaluation among clinicians. In another study (37) concerning the various defects, the coefficients of variation ranged between 4.80% (dubious tail) and 132.97% (thin midpiece), and the coefficients of agreement for specific defects fell into the fair or slight level, which is not acceptable in terms of quality requirements.

In France, the ISO/EN 15189 standard means that it is mandatory to perform regular internal quality controls as well as an external quality control exercise between different laboratories. It is surprising to find that 20% of biologists surveyed do not perform any internal quality control procedures. In a Chinese team (38), quality control training for standardizing sperm morphology criteria reduced the individual interpreting differences: the mean percentage difference reported among three technicians varied from 4.57% $\pm$ 3.69% to 1.96% $\pm$ 1.19% after training. Franken et al. (39) have shown that good sperm morphology reading skills can be achieved and maintained by initial training sessions followed by a continuous external quality control program and annual refresher courses. An important point highlighted by these investigators is that only those technicians who attended refresher courses were able to maintain their morphology reading skills over an extended period of 40 months. The other participants showed a decline in sperm assessment skills 6 to 9 months after their initial training course.

The most used index of multiple defects in France is the MAI in the David's classification. The MAI records the mean number of abnormalities per abnormal spermatozoon. In the literature, the data on the clinical pertinence of these sperm malformation indexes are very poor. There exist one or two publications for each index, which would appear insufficient for a routine parameter used by most laboratories. These indexes are correlated with fertility in vivo for MAI (5, 40) and for TZI (35), and correlated with IVF for SDI (41). For in vivo and IVF outcomes TZI has little predictive value (42). Our survey reveals that clinicians place little faith in this index because 45% of them never or very rarely take it into account for diagnosing infertility, and 50% never or very rarely take it into account when making decisions about ART (see Supplemental Table 1). Given the data found in the literature, one can justifiably question the relevance of a systematic recording of such indexes.

In France, 50% of the clinicians and biologists surveyed who perform ART frequently use the percentage of normal forms as an aid in choosing the type of technique (see Supplemental Table 2). The thresholds used in such choices are extremely variable (see Fig. 3). The impact of sperm morphologic characteristics on pregnancy rates in IUI cycles is still debated. In a meta-analysis (43), the threshold values showing an impact of sperm morphology on pregnancy rates were 4% in most cases when strict criteria were used, and varied from 8% (44) to 50% (45) when the WHO (1987, 1992) criteria were used. Recent publications have failed to demonstrate a statistically significant difference in pregnancy rates in IUI cycles between groups with or without teratozoospermia (9, 46).

Concerning IVF procedures, in 1998 a structured literature review (47) and a review of the majority of surveys (more than 80%) among 18 such surveys performed between 1976 and 1996 showed that the percentage of normal forms was positively associated with a conventional IVF outcome, the clinical pregnancy rate and fertilization rate using the threshold of 5% (strict criteria) or 14% according to WHO guidelines. In ICSI situations, most publications found no statistically significant difference in either fertilization rate or pregnancy rate as a function of the percentage of normal forms (48–51).

There are insufficient arguments in the literature to choose between IVF and ICSI in cases of severe teratozoospermia (52). The decision to assign a couple to either an IVF or an ICSI treatment must be based mainly on the number of total motile spermatozoa after selection. In the present survey, it was very surprising to find that, among those clinicians and biologists who use the percentage of normal forms (in some cases, always) for deciding on the ART technique, more than 65% of them use the David classification, whereas all the publications that demonstrate a morphologic impact on the outcome of induced pregnancies have used the former WHO classification or the more recent strict criteria.

It would be interesting to compare the predictive value of each technique, taking into account the staining method, the classification system, and the type of reading (visual or CASA) to know the best technique and methods to work with more homogeneity. Data in the literature are poor on this topic. Blanchard et al. (53) showed a lower discriminative power for the David classification on fertilization rates in conventional IVF compared with the strict criteria assessed by CASA. We must note that in this publication the techniques were Shorr staining, David classification, and visual reading for the first group and sperm washing before smears, Diff-Quik staining, and CASA assessment for the other. We must consider all the steps of preparation of the smears and reading in view of their impact. In another study (31), the investigators compared the predictive value of two different sperm

morphology assessment techniques on fertilization rate in IVF: CASA assessment with strict criteria after washing and Diff-Quik staining versus conventional visual reading by David classification after Shorr staining. The fertilization rate was poorly correlated with sperm morphology using both methods, but in this study morphology assessed with CASA was more reproducible and a little more discriminative.

A possible source of bias in this type of survey-based study is that more responses are likely to be received from people who are dissatisfied by the morphology assessment than from those who are indifferent (without opinion) or those who are satisfied. From another point of view, another bias could be to have overcoverage of well-informed, interested clinicians and biologists. However, for a survey-based study sent by e-mail, the participation rate was quite good, indicating that clinicians and biologists seemed interested in this topic.

In conclusion, we have shown that in France, although sperm morphology assessment is almost always systematically performed in the search for infertility etiologies, there is a great disparity in the practices used by laboratories and a lack of compliance with recommendations. Twenty years ago, the same kind of study concluded, "Lack of standardization of sperm morphology assessments remains the main reason for the debatable usefulness of this parameter in the laboratory evaluation of semen" (22). In 2016, our conclusion is the same, at least in France. One result of this is the lack of interest shown by clinicians and biologists in France because 22% of clinicians and 26% of biologists believe that this test is totally or largely unreliable in an analytic sense. Twenty-four percent of clinicians and 39% of biologists believe that the test is totally or largely irrelevant as a clinical aid. The majority of the individuals surveyed who think sperm morphology is not clinically relevant think the test is unreliable, too (data not shown). This lack of analytic reliability of sperm morphology assessment could explain the little faith that clinicians and biologists place in this test.

A change in clinical and biological practices for assessing sperm morphology should be initiated. These results lead us to formulate some proposals directed toward two main goals.

First, we recommend the following changes to increase the analytic reliability of the test. [1] Standardize the staining methods and decrease their number. [2] Perform internal and external quality control programs and regular refresher courses to maintain the skills of technicians and the knowledge of biologists. [3] Use only one classification system in France. We believe strict criteria should be preferred because [a] the concept of borderline aspects is responsible for more interobserver and intraobserver variability, [b] the vast majority of publications studying the impact of teratozoospermia on ART procedures concern only strict criteria, and finally [c] the test has broad international use. [4] Avoid systematic determination of the frequency of each morphologic abnormality as they are sources of even higher variability as recommended by WHO guidelines. Such analyses should only be performed on request during an andrologic investigation.

Second, to increase the trust of clinicians in the test we propose to [1] stop the systematic recording of any index with multiple defects (MAI, TZI, SDI), given the poor data in the literature relative to its clinical interest; [2] sensitize clinicians to the real relevance of teratozoospermia and importance of monomorphous abnormalities. Such relevance could only be admitted once the analytic quality of sperm morphology is optimized and once clinical studies, ideally performed in each ART center, prove its benefit for patient care. Once such best practices are implemented for this test, the prescribing physicians should have more faith in the results, and the patients will be better served.

Acknowledgments: The authors thank Myriam Daudin, Nathalie Moinard, François Isus, and Carole Teston from the Service de Médecine de la reproduction, CHU Toulouse, for reading the questionnaire; and the Agence de Biomédecine (ABM), Société d'Andrologie de Langue Française (SALF), Biologistes de Laboratoires d'Etude de la Fécondation et de Conservation des Œufs (BLEFCO), and Collège de Gynécologie du Midi for providing access to the addresses of physicians and biologists.

REFERENCES

1. Menkveld R, Stander FS, Kotze TJ, Kruger TF, van Zyl JA. The evaluation of morphological characteristics of human spermatozoa according to stricter criteria. Hum Reprod 1990;5:586–92.
2. Franken DR. How accurate is sperm morphology as an indicator of sperm function? Andrologia 2015;47:720–3.
3. Liu DY, Baker HW. Sperm nuclear chromatin normality: relationship with sperm morphology, sperm-zona pellucida binding, and fertilization rates in vitro. Fertil Steril 1992;58:1178–84.
4. Guzick DS, Overstreet JW, Factor-Litvak P, Brazil CK, Nakajima ST, Coutifaris C, et al. Sperm morphology, motility, and concentration in fertile and infertile men. N Engl J Med 2001;345:1388–93.
5. Jouannet P, Ducot B, Feneux D, Spira A. Male factors and the likelihood of pregnancy in infertile couples. I. Study of sperm characteristics. Int J Androl 1988;11:379–94.
6. Kruger TF, Acosta AA, Simmons KF, Swanson RJ, Matta JF, Veeck LL, et al. New method of evaluating sperm morphology with predictive value for human in vitro fertilization. Urology 1987;30:248–51.
7. Lemmens L, Kos S, Beijer C, Brinkman JW, van der Horst FA, van den Hoven L, et al. Predictive value of sperm morphology and progressively motile sperm count for pregnancy outcomes in intrauterine insemination. Fertil Steril 2016;105:1462–8.
8. Oehninger S, Franken DR, Ombelet W. Sperm functional tests. Fertil Steril 2014;102:1528–33.
9. Lockwood GM, Deveneau NE, Shridharani AN, Strawn EY, Sandlow JI. Isolated abnormal strict morphology is not a contraindication for intrauterine insemination. Andrology 2015;3:1088–93.
10. Comhaire F, Schoonjans F, Vermeulen L, De Clercq N. Methodological aspects of sperm morphology evaluation: comparison between strict and liberal criteria. Fertil Steril 1994;62:857–61.
11. Menkveld R. Sperm morphology assessment using strict (Tygerberg) criteria. Methods Mol Biol 2013;927:39–50.
12. Henkel R, Schreiber G, Sturmhoefel A, Hipler UC, Zermann DH, Menkveld R. Comparison of three staining methods for the morphological evaluation of human spermatozoa. Fertil Steril 2008;89:449–55.
13. Franken DR, Menkveld R, Kruger TF, Sekadde-Kigondu C, Lombard C. Monitoring technologist reading skills in a sperm morphology quality control program. Fertil Steril 2003;79(Suppl 3):1637–43.
14. Mallidis C, Cooper TG, Hellenkemper B, Lablans M, Uckert F, Nieschlag E. Ten years' experience with an external quality control program for semen analysis. Fertil Steril 2012;98:611–6.e4.
15. Eustache F, Auger J. Inter-individual variability in the morphological assessment of human sperm: effect of the level of experience and the use of standard methods. Hum Reprod 2003;18:1018–22.

16. Leushuis E, van der Steeg JW, Steures P, Repping S, Bossuyt PM, Blankenstein MA, et al. Reproducibility and reliability of repeated semen analyses in male partners of subfertile couples. Fertil Steril 2010;94:2631–5.
17. Bonde JP, Ernst E, Jensen TK, Hjollund NH, Kolstad H, Henriksen TB, et al. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. Lancet 1998;352:1172–7.
18. Keel BA, Quinn P, Schmidt CF Jr, Serafy NT Jr, Serafy NT Sr, Schalue TK. Results of the American Association of Bioanalysts national proficiency testing programme in andrology. Hum Reprod 2000;15:680–6.
19. Matson PL. External quality assessment for semen analysis and sperm antibody detection: results of a pilot scheme. Hum Reprod 1995;10:620–5.
20. Ombelet W, Bosmans E, Janssen M, Cox A, Maes M, Punjabi U, et al. Multicenter study on reproducibility of sperm morphology assessments. Arch Androl 1998;41:103–14.
21. Davis RO, Gravance CG. Standardization of specimen preparation, staining, and sampling methods improves automated sperm-head morphometry analysis. Fertil Steril 1993;59:412–7.
22. Ombelet W, Pollet H, Bosmans E, Vereecken A. Results of a questionnaire on sperm morphology assessment. Hum Reprod 1997;12:1015–20.
23. Menkveld R, Holleboom CA, Rhemrev JP. Measurement and significance of sperm morphology. Asian J Androl 2011;13:59–68.
24. Eliasson R. Semen analysis with regard to sperm number, sperm morphology and functional aspects. Asian J Androl 2010;12:26–32.
25. Auger J. Assessing human sperm morphology: top models, underdogs or biometrics? Asian J Androl 2010;12:36–46.
26. Mortimer D, Menkveld R. Sperm morphology assessment—historical perspectives and current opinions. J Androl 2001;22:192–205.
27. Kruger TF, Ackerman SB, Simmons KF, Swanson RJ, Brugo SS, Acosta AA. A quick, reliable staining technique for human sperm morphology. Arch Androl 1987;18:275–7.
28. Menkveld R, Lacquet FA, Kruger TF, Lombard CJ, Sanchez Sarmiento CA, de Villiers A. Effects of different staining and washing procedures on the results of human sperm morphology evaluation by manual and computerised methods. Andrologia 1997;29:1–7.
29. Natali I, Muratori M, Sarli V, Vannuccini M, Cipriani S, Niccoli L, et al. Scoring human sperm morphology using Testsimplets and Diff-Quik slides. Fertil Steril 2013;99:1227–1232 e2.
30. Menkveld R, Kruger TF. Advantages of strict (Tygerberg) criteria for evaluation of sperm morphology. Int J Androl 1995;18(Suppl 2):36–42.
31. Marnet B, Vieitez G, Milhet P, Richoilley G, Lesourd F, Parinaud J. Computer-assisted assessment of sperm morphology: comparison with conventional techniques. Int J Androl 2000;23:22–8.
32. World Health Organization. Laboratory manual for the examination and processing of human semen. 5th ed. Geneva: WHO; 2010.
33. Cooper TG, Noonan E, von Eckardstein S, Auger J, Baker HW, Behre HM, et al. World Health Organization reference values for human semen characteristics. Hum Reprod Update 2010;16:231–45.
34. Ombelet W, Bosmans E, Janssen M, Cox A, Vlasselaer J, Gyselaers W, et al. Semen parameters in a fertile versus subfertile population: a need for change in the interpretation of semen testing. Hum Reprod 1997;12:987–93.
35. Menkveld R, Wong WY, Lombard CJ, Wetzels AM, Thomas CM, Merkus HM, et al. Semen parameters, including WHO and strict criteria morphology, in a fertile and subfertile population: an effort towards standardization of in-vivo thresholds. Hum Reprod 2001;16:1165–71.
36. Auger J, Jouannet P, Eustache F. Another look at human sperm morphology. Hum Reprod 2016;31:10–23.
37. Wang Y, Yang J, Jia Y, Xiong C, Meng T, Guan H, et al. Variability in the morphologic assessment of human sperm: use of the strict criteria recommended by the World Health Organization in 2010. Fertil Steril 2014;101:945–9.
38. Yao KS, Zhang XZ, Wu Y. Assessment of sperm morphology without quality control may be meaningless for clinicians. Asian J Androl 2010;12:607–8.
39. Franken DR, Kruger TF. Lessons learned from a sperm morphology quality control programme. Andrologia 2006;38:225–9.
40. Slama R, Eustache F, Ducot B, Jensen TK, Jorgensen N, Horte A, et al. Time to pregnancy and semen parameters: a cross-sectional study among fertile couples from four European cities. Hum Reprod 2002;17:503–15.
41. Aziz N, Sharma RK, Mahfouz R, Jha R, Agarwal A. Association of sperm morphology and the sperm deformity index (SDI) with poly (ADP-ribose) polymerase (PARP) cleavage inhibition. Fertil Steril 2011;95:2481–4.
42. van Zyl JA, Menkveld R. Oligozoospermia: recent prognosis and the outcome of 73 pregnancies in oligozoospermic couples. Andrologia 2006;38:87–91.
43. Van Waart J, Kruger TF, Lombard CJ, Ombelet W. Predictive value of normal sperm morphology in intrauterine insemination (IUI): a structured literature review. Hum Reprod Update 2001;7:495–500.
44. Comhaire F, Zalata A, Mahmoud A, Depoorter B, Huysse L, Christophe A, et al. Diagnostic and therapeutic approach to moderate and severe male subfertility in 1995. Hum Reprod 1995;10(Suppl 1):144–50.
45. Francavilla F, Romano R, Santucci R, Poccia G. Effect of sperm morphology and motile sperm count on outcome of intrauterine insemination in oligozoospermia and/or asthenozoospermia. Fertil Steril 1990;53:892–7.
46. Deveneau NE, Sinno O, Krause M, Eastwood D, Sandlow JI, Robb P, et al. Impact of sperm morphology on the likelihood of pregnancy after intrauterine insemination. Fertil Steril 2014;102:1584–90.e2.
47. Coetzee K, Kruge TF, Lombard CJ. Predictive value of normal sperm morphology: a structured literature review. Hum Reprod Update 1998;4:73–82.
48. Keegan BR, Barton S, Sanchez X, Berkeley AS, Krey LC, Grifo J. Isolated teratozoospermia does not affect in vitro fertilization outcome and is not an indication for intracytoplasmic sperm injection. Fertil Steril 2007;88:1583–8.
49. French DB, Sabanegh ES Jr, Goldfarb J, Desai N. Does severe teratozoospermia affect blastocyst formation, live birth rate, and other clinical outcome parameters in ICSI cycles? Fertil Steril 2010;93:1097–103.
50. Osawa Y, Sueoka K, Iwata S, Shinohara M, Kobayashi N, Kuji N, et al. Assessment of the dominant abnormal form is useful for predicting the outcome of intracytoplasmic sperm injection in the case of severe teratozoospermia. J Assist Reprod Genet 1999;16:436–42.
51. Host E, Lindenberg S, Ernst E, Christensen F. Sperm morphology and IVF: embryo quality in relation to sperm morphology following the WHO and Kruger's strict criteria. Acta Obstet Gynecol Scand 1999;78:526–9.
52. Hotaling JM, Smith JF, Rosen M, Muller CH, Walsh TJ. The relationship between isolated teratozoospermia and clinical pregnancy after in vitro fertilization with or without intracytoplasmic sperm injection: a systematic review and meta-analysis. Fertil Steril 2011;95:1141–5.
53. Blanchard M, Haguenoer K, Apert A, Poret H, Barthelemy C, Royere D, et al. Sperm morphology assessment using David's classification: time to switch to strict criteria? Prospective comparative analysis in a selected IVF population. Int J Androl 2011;34:145–52.

SUPPLEMENTAL TABLE 1

Usefulness of index of multiple sperm defects (MAI, TZI, or SDI) for physicians.

Frequency	For diagnosis	For the choice of ART technique
Never	17.9% (12)	25.8% (17)
Very rarely	26.9% (18)	24.2% (16)
Sometimes	25.4% (17)	21.2% (14)
Often/frequently	19.4% (13)	24.2% (16)
Always	10.4% (7)	4.5% (3)

Note: ART = assisted reproductive technology; MAI = multiple anomalies index; SDI = sperm deformity index; TZI = teratozoospermia index.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.

SUPPLEMENTAL TABLE 2

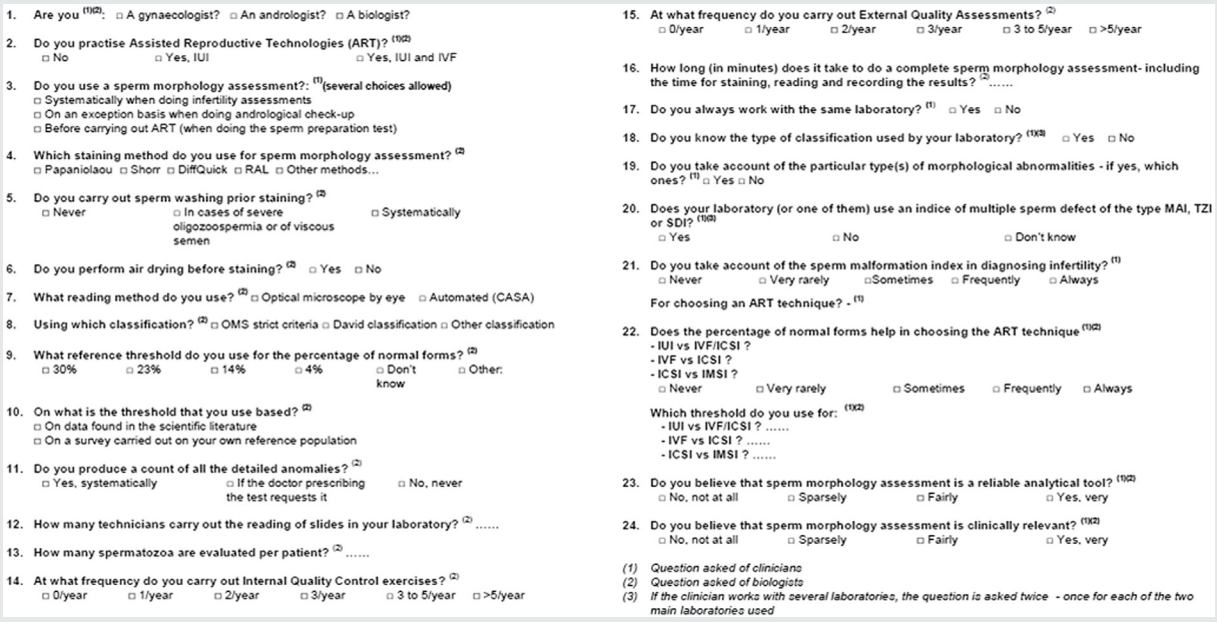
Usefulness of the results of sperm morphology for choice of ART procedure.

	IUI vs. IVF–ICSI		IVF vs. ICSI		ICSI vs. IMSI	
Frequency	**Physicians (%)**	**Biologists (%)**	**Physicians (%)**	**Biologists (%)**	**Physicians (%)**	**Biologists (%)**
Never	5.1	14.3	8.5	3.9	46.4	44.9
Very rarely	15.3	18.2	13.6	23.4	12.5	14.5
Sometimes	23.7	29.9	23.7	31.2	23.2	15.9
Often	27.1	27.3	28.8	27.3	14.3	17.4
Always	28.8	10.4	25.4	14.3	3.6	7.2

Note: ICSI = intracytoplasmic sperm injection; IMSI = intracytoplasmic morphologically selected sperm injection; IUI = intrauterine insemination; IVF = in vitro fertilization.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.

SUPPLEMENTAL FIGURE 1



1. Are you [1][2]: □ A gynaecologist? □ An andrologist? □ A biologist?
2. Do you practise Assisted Reproductive Technologies (ART)? [1][2]
   □ No □ Yes, IUI □ Yes, IUI and IVF
3. Do you use a sperm morphology assessment?: [1](several choices allowed)
   □ Systematically when doing infertility assessments
   □ On an exception basis when doing andrological check-up
   □ Before carrying out ART (when doing the sperm preparation test)
4. Which staining method do you use for sperm morphology assessment? [2]
   □ Papaniolaou □ Shorr □ DiffQuick □ RAL □ Other methods...
5. Do you carry out sperm washing prior staining? [2]
   □ Never □ In cases of severe oligozoospermia or of viscous semen □ Systematically
6. Do you perform air drying before staining? [2] □ Yes □ No
7. What reading method do you use? [2] □ Optical microscope by eye □ Automated (CASA)
8. Using which classification? [2] □ OMS strict criteria □ David classification □ Other classification
9. What reference threshold do you use for the percentage of normal forms? [2]
   □ 30% □ 23% □ 14% □ 4% □ Don't know □ Other:
10. On what is the threshold that you use based? [2]
    □ On data found in the scientific literature
    □ On a survey carried out on your own reference population
11. Do you produce a count of all the detailed anomalies? [2]
    □ Yes, systematically □ If the doctor prescribing the test requests it □ No, never
12. How many technicians carry out the reading of slides in your laboratory? [2] .....
13. How many spermatozoa are evaluated per patient? [2] .....
14. At what frequency do you carry out Internal Quality Control exercises? [2]
    □ 0/year □ 1/year □ 2/year □ 3/year □ 3 to 5/year □ >5/year
15. At what frequency do you carry out External Quality Assessments? [2]
    □ 0/year □ 1/year □ 2/year □ 3/year □ 3 to 5/year □ >5/year
16. How long (in minutes) does it take to do a complete sperm morphology assessment- including the time for staining, reading and recording the results? [2] .....
17. Do you always work with the same laboratory? [1] □ Yes □ No
18. Do you know the type of classification used by your laboratory? [1][3] □ Yes □ No
19. Do you take account of the particular type(s) of morphological abnormalities - if yes, which ones? [1] □ Yes □ No
20. Does your laboratory (or one of them) use an indice of multiple sperm defect of the type MAI, TZI or SDI? [1][3]
    □ Yes □ No □ Don't know
21. Do you take account of the sperm malformation index in diagnosing infertility? [1]
    □ Never □ Very rarely □ Sometimes □ Frequently □ Always
    For choosing an ART technique? - [1]
22. Does the percentage of normal forms help in choosing the ART technique [1][2]
    - IUI vs IVF/ICSI ?
    - IVF vs ICSI ?
    - ICSI vs IMSI ?
    □ Never □ Very rarely □ Sometimes □ Frequently □ Always
    Which threshold do you use for: [1][2]
    - IUI vs IVF/ICSI ? .....
    - IVF vs ICSI ? .....
    - ICSI vs IMSI ? .....
23. Do you believe that sperm morphology assessment is a reliable analytical tool? [1][2]
    □ No, not at all □ Sparsely □ Fairly □ Yes, very
24. Do you believe that sperm morphology assessment is clinically relevant? [1][2]
    □ No, not at all □ Sparsely □ Fairly □ Yes, very

(1) Question asked of clinicians
(2) Question asked of biologists
(3) If the clinician works with several laboratories, the question is asked twice - once for each of the two main laboratories used

Questionnaire contents. IUI = intrauterine insemination; IVF = in vitro fertilization; CASA = computer-assisted sperm analysis; MAI = multiple anomalies index; TZI = teratozoospermia index; SDI = sperm deformity index; ICSI = intracytoplasmic sperm injection; IMSI = intracytoplasmic morphologically selected sperm injection.

Gatimel. Survey on sperm morphology assessment. Fertil Steril 2016.