Facteurs influençant la classification morphologique des spermatozoïdes humains : Leçons d'une évaluation externe de la qualité auprès de 62 participants

Jacques AUGER¹, Florence EUSTACHE²

¹ Service d'Histologie-Embryologie/Biologie de la Reproduction/CECOS Paris Cochin, Hôpital Cochin, Paris

² CECOS Paris Necker, Hôpital Necker Enfants Malades, Paris

RESUME

L'évaluation morphologique des spermatozoïdes humains n'est pas exempte de problèmes méthodologiques rendant difficile les comparaisons entre laboratoires. Le but du présent travail était l'étude de la variabilité interindividuelle dans le classement des spermatozoïdes typiques et atypiques et des facteurs susceptibles de la moduler. Soixante deux techniciens et biologistes ont classé 100 spermatozoïdes dont les images vidéo captées sur un microscope ont été directement projetées sur un écran. La classification de David distinguant les spermatozoïdes normaux et anormaux et calculant l'index multiple d'anomalie, IAM, a été employée. Les mêmes spermatozoïdes ont aussi été classés ultérieurement par deux biologistes entraînés à partir d'une bande vidéo, les résultats de cette classification servant de valeurs de référence. L'analyse des données s'est concentrée principalement sur l'effet de l'expérience et sur l'impact possible du suivi des règles de bonne pratique.

Le coefficient de variation pour l'estimation du pourcentage de spermatozoïdes normaux était de 40%, il était inférieur pour différentes anomalies et égal à 12% pour l'IAM. Ce sont les participants les plus expérimentés qui avaient les résultats les plus proches des valeurs de référence. Enfin, il a été trouvé que les participants qui n'emploient pas en routine les méthodologies recommandées avaient les résultats différant le plus de ceux des experts, et ceci, quel que soit leur niveau d'expérience. La faible variabilité observée pour l'IAM suggérait l'intérêt pratique de cet index. Par ailleurs, les résultats réaffirmaient sans aucune ambiguïté la nécessité impérieuse d'une standardisation et de contrô-

les de qualités lors de l'évaluation morphologique des spermatozoïdes humains.

Mots clés : analyse du sperme, contrôle de qualité, index d'anomalies multiples, spermocytogramme

I. INTRODUCTION

L'évaluation morphologique des spermatozoïdes représente une étape essentielle dans l'analyse de routine du sperme. Malheureusement, cette évaluation n'est pas exempte de véritables difficultés compliquant les comparaisons d'un laboratoire à l'autre. Ces difficultés sont encore notamment dues à l'existence de nombreuses techniques de coloration, au recours à des systèmes de classification différents (plus de dix systèmes de classification ont été proposés depuis la fin des années 50) sans qu'il n'y ait encore eu un véritable consensus sur une méthode universelle. Cette situation est illustrée dans la troisième édition du manuel de l'OMS où il est dit que «les spermatozoïdes humains peuvent être évalués au grossissement final 1000x sur des préparations fixées et colorées ou bien, sur des préparations fraîches, en contraste de phase, au grossissement final 600x... » [27].

Correspondance :

Dr Jacques Auger - Biologie de la Reproduction, Pavillon Cassini, Hôpital Cochin, 123 Bd de Port-Royal, 75014 Paris, France. Tel 01.58.41.15.71 - Fax 01.58.41.15.65 - E-mail jacques.auger@cch.ap-hop-paris.fr

Conscient de ces problèmes, les experts de l'OMS dans la 4^{ème} édition du manuel ont recommandé l'utilisation d'une méthode unique et simple classant principalement les spermatozoïdes morphologiquement normaux en utilisant des critères stricts [28], se basant essentiellement sur la valeur pronostique de cette méthode en FIV [19]. Cependant, des travaux récents reposant sur l'utilisation d'autres méthodes, telle la méthode de David *et al.* [9] classant les diverses anomalies des spermatozoïdes, ont démontré leur intérêt clinique [3, 26]. La classification de David *et al.* permet le calcul d'un index d'anomalie multiple, IAM [16] approprié pour le diagnostic des infertilités masculines [3], le pronostic [16, 26] et, le calcul de cet index a été recommandé dans les 3^{ème} et 4^{ème} éditions du manuel de l'OMS [27, 28]. Cependant, quelle que soit la classification utilisée, la nature subjective de la perception visuelle est une autre source de variabilité intra- et interindividuelle. Au total, c'est un ensemble de facteurs qui contribue à limiter la fiabilité des résultats d'un point de vue diagnostique ou pronostique aussi bien que dans le cadre de recherches.

Pour évaluer la (les) source(s) et l'ampleur des erreurs au cours de l'analyse de routine du sperme, le contrôle de qualité interne (CQI) a été récemment introduit au laboratoire de biologie de la reproduction [18, 7, 6]. Un tel CQI est essentiel pour maintenir l'exactitude, la précision et la compétence et c'est un préalable à l'appréciation et à l'interprétation correctes dans le processus diagnostique de l'infertilité humaine. L'évaluation externe de qualité (EEQ) est l'évaluation des résultats pour les mêmes échantillons dans plusieurs laboratoires. En utilisant une telle approche, une variabilité dans les résultats de l'analyse de routine du sperme, d'assez importante à très importante, a récemment été rapportée [24, 20, 2]. Une grande variabilité inter-laboratoire des résultats de l'analyse du sperme n'est pas sans conséquence pour le patient : à partir d'un même prélèvement, il pourrait être classé hypofertile dans un laboratoire et normal dans un autre. Par conséquent, l'EEQ est également indispensable pour s'assurer que les différents laboratoires produisent des résultats comparables qui permettent à leur tour une évaluation diagnostique et pronostique similaire quel que soit le centre.

Quelques résultats d'EEQ pour l'analyse morphologique des spermatozoïdes, après envoi de lames colorées dans les différents centres participant, ont récemment été rapportés [24, 20, 8, 13]. Une telle pratique essentielle offre la possibilité que les différents laboratoires évaluent leurs propres résultats en comparaison à ceux des autres.

Dans l'étude rapportée ici, nous avons invité des techniciens et des biologistes à participer à un atelier d'EEQ pour évaluer la morphologie de spermatozoïdes dont les images directement captées sur un vidéo-microscope ont

été projetées sur un grand écran. Ceci a offert l'opportunité d'évaluer la variabilité interindividuelle dans l'analyse morphologique des spermatozoïdes humains par la classification de David *et al.* [9] - la grande majorité des laboratoires français utilisant cette méthode - et, d'étudier l'effet de l'expérience et du suivi des règles de bonne pratique.

II. MATERIEL ET METHODES

1. Participants à l'étude

62 participants se sont retrouvés à l'université de Nancy pendant deux jours au cours d'un atelier de contrôle de qualité principalement consacré à l'évaluation de la morphologie des spermatozoïdes humains. Cet atelier a été organisé par le Groupe Français d'Assurance de Qualité en Biologie de la Reproduction, la Société d'Andrologie de Langue Française (SALF), la Fédération Française des CECOS (Centres d'Etude et de Conservation des Oeufs et du Sperme humains), la Société Française de Biologie Cellulaire (SFBC) et l'Association Française des Biologistes de la FIV (BLEFCO). Les participants étaient des techniciens de laboratoire, des biologistes ou des médecins travaillant dans 40 laboratoires français et 2 laboratoires luxembourgeois qui étaient principalement des laboratoires universitaires ou hospitaliers, mais aussi des laboratoires privés d'analyses médicales. Chaque participant a rempli un questionnaire sur son statut professionnel, le statut de son laboratoire (privé ou public) et le secteur d'activité (spermologie, banque de sperme, aide médicale à la procréation, AMP, ou, biologie médicale). Sur ce même questionnaire, il devait préciser les méthodes employées pour l'évaluation morphologique des spermatozoïdes (techniques de réalisation du frottis et de coloration, objectif et grossissement utilisés, nombre de spermatozoïdes évalués, et méthode de classification utilisée). Les caractéristiques des participants et les méthodes employées sont récapitulées dans le Tableau 1. Les participants avaient des expériences diverses en matière d'analyse du sperme et d'évaluation de la morphologie. Le Tableau 2 résume les réponses au questionnaire concernant l'expérience.

2. Matériel

Un frottis préparé conformément aux recommandations de l'OMS [27] a été utilisé dans la présente étude. L'échantillon de sperme pour faire le frottis a été choisi en raison de son niveau élevé de diverses anomalies morphologiques. Brièvement, le frottis a été réalisé à partir d'un volume de 10 µl de l'échantillon de sperme, séché à l'air, fixé pendant une heure avec un mélange d'éthanol absolu (2/3) et d'acide acétique (1/3), puis, coloré au Shorr en utilisant un automate de coloration (Sakura DRS601, Bayer-diagnostic, Puteaux).

Tableau 1 : Caractéristiques des participants (A) et des procédures employées en routine pour l'évaluation morphologique des spermatozoïdes (B).

A - Caractéristiques étudiées		Nombre de participants a (%)
Statut professionnel	Pharmacien biologiste	9 (14,5)
	Médecin biologiste	5 (8,0)
	Technicien de laboratoire	43 (69,4)
	Scientifique	4 (6,5)
	Étudiant	1 (1,6)
Secteur d'activité	Laboratoire d'Analyses Médicales	25 (40,3)
	Hôpital privé/public	36 (58,1)
	Autre	1 (1,6)
Activité	Biologie générale incluant l'analyse du sperme	20 (32,3)
	Spécialisation en spermologie	24 (38,7)
	AMP	6 (9,6)
	Banque de sperme (CECOS)	12 (19,4)
B - Procédures utilisées en routine pour l'évaluation morphologique		Nombre de participants (%)
Volume fixe de l'échantillon pour faire le frottis	Oui b	30 (50,0)
	Non	30 (50,0)
Si oui, volume employé?	5µl b	8 (27,6)
	10µl b	20 (69,0)
	20µl b	2 (3,4)
Coloration	Papanicolaou b	8 (13,3)
	Shorr b	31 (51,7)
	May Grunwald Giemsa	5 (8,4)
	autre	16 (26,6)
Grossissement final	1000x b	53 (91,4)
	400x	4 (6,9)
	100x	1 (1,7)
Nombre de spermatozoïdes évalués	50	1 (1,7)
	100	54 (93,1)
	200 b	3 (5,2)
Méthode de classification	David	56 (93,2)
	autres	4 (6,8)
	critères stricts b	0

a un total < 62 signifie que quelques participants n'ont pas répondu; b recommandations de l'OMS [28]

Tableau 2 : Nombre d'années de pratique (a) et nombre moyen d'évaluations morphologiques par an (b) pour tous les participants à l'atelier.

a - Nombre d'années		Nombre de participants (%)
Moyenne ± écart-type :	6,2 ± 6,5 ans	
Médiane :	5,0 ans	
10 ^{ème} percentile :	0 ans	
90 ^{ème} percentile :	15 ans	
		0 ans : 9 (14,5)
		> 0-2 ans : 15 (24,2)
		> 2-5 ans : 13 (21,0)
		> 5-10 ans : 14 (22,6)
		> 10 ans : 11 (17,7)

b - Nombre d'évaluations morphologiques par an		Nombre de participants (%)
Moyenne ± écart-type :	587 ± 1397	
Médiane :	260 (soit 5 par semaine)	
10 ^{ème} percentile :	0	
90 ^{ème} percentile :	1040	
		0 : 9 (14,5)
		> 0-104 (> 0-2 par semaine) : 18 (29,0)
		> 104-520 (> 2-10 par semaine) : 18 (29,0)
		> 520-1040 (> 10-20 par semaine) : 11 (17,7)
		> 1040-3640 (> 20-70 par semaine) : 5 (8,1)
		> 3640 (> 70 par semaine) : 1 (1,7)

La préparation colorée a été disposée sur la platine d'un microscope BX50 (Olympus, Rungis, France) équipé d'un objectif x100 à immersion et d'une caméra couleur XC-003P 3-CCD RVB (Sony, Clichy, France). La caméra était reliée à un vidéoprojecteur VPL-FE110 (Sony, Clichy, France) donnant une très bonne définition de l'image projetée. L'atelier a eu lieu dans un amphithéâtre d'une capacité d'environ 200 places. Avant le début de l'atelier, la qualité de l'image projetée a été vérifiée et optimisée afin que les images des spermatozoïdes aient une définition permettant leur classification sans aucun biais visuel. Ce but était atteint si l'observateur était placé face à l'écran et dans les 6 premiers rangs de l'amphithéâtre.

3. Méthode de classification des spermatozoïdes normaux et anormaux

Pour la grande majorité des laboratoires français, l'évaluation de la morphologie des spermatozoïdes humains est faite selon la classification de David [9], modifiée après la publication de Jouannet *et al.* [16]. Aussi, c'est cette classification qui a été employée dans la présente étude. Le pourcentage des spermatozoïdes morphologiquement normaux et le nombre de chaque anomalie sont évalués à partir de 100 spermatozoïdes observés au grossissement final x1000. Cette classification distingue le spermatozoïde morphologiquement normal et sept anomalies de la tête, trois anomalies de la pièce intermédiaire et cinq anomalies du flagelle. Les spermatozoïdes normaux sont définis selon des critères stricts semblables aux critères proposés dans la 4ème édition du manuel de l'OMS [28], cependant, les for-

mes subnormales sont considérées comme normales. Toutes les anomalies sont également définies par des critères spécifiques [1]. Grâce à un système à entrées multiples, toutes les anomalies de chaque spermatozoïde sont enregistrées de telle sorte qu'aucune anomalie ne soit sous-estimée par rapport aux autres, ce qui offre la possibilité de calculer l'index multiple d'anomalies, IAM [16], index considéré comme un test additionnel dans la 4ème édition du manuel de l'OMS [28]. L'IAM correspond au nombre moyen d'anomalies par spermatozoïde anormal. D'autres éléments cellulaires (flagelles isolés, spermatozoïdes en lyse, cellules de la lignée blanche, cellules de la lignée germinale, autres cellules et débris cellulaires) sont enregistrés mais non inclus dans le compte et n'ont pas été analysés dans la présente étude.

4. Schéma organisationnel de l'évaluation externe de qualité (EEQ)

Deux heures avant l'EEQ, les définitions de la morphologie normale et anormale des spermatozoïdes ont été rappelées ainsi que les procédures standardisées pour l'évaluation selon la méthode de David [1]. Juste avant l'évaluation, une grille simplifiée (chaque spermatozoïde en ligne, chaque anomalie en colonne) a été distribuée à tous les participants. Les participants ont été invités à répondre sans l'aide de leurs voisins. Cent spermatozoïdes sélectionnés au microscope ont été projetés successivement et chaque spermatozoïde choisi était repéré à l'aide de la flèche lumineuse intégrée au microscope. Ainsi, tous les participants ont évalué le même spermatozoïde dans le même temps. Le temps d'observation et de réponse pour chaque spermatozoïde était de 45 secondes. De plus, tous les spermatozoïdes analysés ont été enregistrés à l'aide d'un magnétoscope professionnel (Panasonic). Dans les jours suivant l'EEQ, les 100 spermatozoïdes enregistrés sur la bande vidéo ont été classés en commun par deux biologistes expérimentés pour la lecture du spermocytogramme (FE et JA). Soit ils étaient d'emblée d'accord sur le diagnostic, ce qui était le cas pour la grande majorité des spermatozoïdes, soit, il y avait un désaccord entre eux (observé uniquement pour quelques anomalies de taille) et, dans ce cas, des mesures objectives ont été faites afin de parvenir à un consensus. Dans la suite du texte, les résultats de cette évaluation sont considérés comme les valeurs de référence.

5. Analyse des données

La variabilité globale de l'évaluation du pourcentage de spermatozoïdes normaux, des diverses anomalies et de l'IAM pour l'ensemble des participants (n = 62) a été étudiée. Puis nous avons estimé la variabilité globale pour l'évaluation des mêmes caractéristiques en se référant au niveau d'expérience des participants pour deux sous-groupes définis selon le nombre d'évaluations faites par an (<520 par an, c'est à dire, moins de 10 évaluations par semaine, n = 45 contre > 520 par an, n = 17).

Nous avons également recherché si les conditions habituelles de l'évaluation, notamment, le suivi en routine des recommandations de l'OMS [27] pour la réalisation des frottis, la coloration, le grossissement, le nombre de spermatozoïdes étudiés,... pouvaient avoir une influence sur les résultats de cette EEQ. Aussi, nous avons étudié les différences entre les résultats des participants et les valeurs de référence pour le pourcentage de spermatozoïdes normaux, les différentes catégories d'anomalies et l'IAM. Les participants ont été divisés en trois groupes. Le premier groupe était composé de participants expérimentés évaluant > 520 frottis par an, depuis au moins deux ans, et, suivant les recommandations de l'OMS de 1992 (frottis à partir d'une goutte calibrée, coloration de Shorr ou de Papanicolaou, classification d'au moins 100 spermatozoïdes au grossissement final x1000). Le deuxième groupe était composé de participants ayant une expérience plus récente (< 2 ans, quel que soit le nombre de frottis analysés par an) mais suivant également recommandations de l'OMS en routine. Le dernier groupe était composé des autres participants, c'est à dire ceux qui ne suivent pas en routine les recommandations de l'OMS, quel que soit leur expérience.

La variabilité inter-participant dans l'évaluation de la morphologie des spermatozoïdes a été exprimée sous la forme du coefficient de variation : $CV (\%) = \text{écart-type} / \text{moyenne} \times 100$. La comparaison de données quantitatives entre deux groupes a utilisé le test non paramétrique de Mann-Whitney. L'égalité des valeurs moyennes des caractéristiques étudiées parmi les trois sous-groupes choisis sur le niveau d'expérience et le suivi des recommandations de l'OMS a été examinée par une analyse de la variance, prenant si nécessaire en considération les variances inégales (test de Brown-Forsythe). Lorsque cette analyse donnait un résultat significatif, les comparaisons deux à deux ont été faites à l'aide du test de Bonferroni. Toutes les statistiques ont été faites avec le logiciel statistique BMDP [12].

Après analyse des données, un rapport détaillé individuel avec des recommandations a été envoyé à chaque participant. Ceci leur a permis d'évaluer leurs propres résultats par rapport aux résultats moyens du groupe et aux résultats des deux biologistes.

III. RESULTATS

1. Variabilité générale inter participant pour les différentes caractéristiques morphologiques

Les valeurs moyennes et médianes du pourcentage de spermatozoïdes normaux, de chaque anomalie enregistrée et de l'IAM, les coefficients de variation inter participant correspondant et les résultats de référence des deux biologistes sont présentées dans le Tableau 3. Par ordre croissant,

on a observé les plus bas coefficients de variation pour l'IAM (12%), les pourcentages de pièces intermédiaires angulées (23%), de flagelles absents (25%) et d'acrosomes anormaux ou absents (26%). Le coefficient de variation pour le pourcentage de spermatozoïdes normaux était égal à 40% (7ème position parmi les 17 caractéristiques étudiées). L'intervalle des valeurs trouvées pour cette dernière caractéristique était large, de 6% à 39% (avec une médiane égale à 15%), une différence marquée avec les intervalles de valeurs pour l'IAM. Les participants fournissant les évaluations les plus sévères du pourcentage des spermatozoïdes normaux ont trouvé les valeurs les plus élevées de l'IAM et réciproquement.

2. Influence de l'expérience

Les valeurs moyennes et médianes du pourcentage de spermatozoïdes normaux, de l'IAM, de chaque anomalie et les coefficients de variation inter participant globaux correspondant pour les deux groupes de participants choisis en fonction de leur expérience sont présentés dans le Tableau 4. Les participants expérimentés ont trouvé sensiblement moins de spermatozoïdes normaux, et plus de têtes multiples, d'acrosomes anormaux, de pièces intermédiaires grêles et de flagelles de calibre irrégulier. Cependant, il faut préciser que pour ces caractéristiques, la variabilité interindividuelle n'était que légèrement plus faible dans ce groupe en comparaison avec le groupe moins expérimenté. Une différence marquée de la variabilité (CV) inter participant comparant ces deux groupes n'a été trouvée que pour les acrosomes anormaux (19% contre 26%) et les flagelles absents (14% contre 27%).

3. Influence du suivi en routine des recommandations

Le nombre moyen de préparations évaluées par an dans les trois groupes de participants était de 1287 ± 876 dans le premier groupe ($n = 10$) contre 355 ± 41 dans le deuxième groupe ($n = 8$; $p < 0.01$) et 470 ± 1562 dans le troisième groupe ($n = 44$; différence non significative). Le nombre moyen d'années d'expérience pour l'évaluation de la morphologie des spermatozoïdes dans les trois groupes était de $10,0 \pm 5,3$ dans le premier groupe contre $1,5 \pm 1,4$ dans le deuxième groupe ($p < 0.05$) et $6,3 \pm 6,8$ dans le troisième groupe (différence non significative). Les caractéristiques morphologiques significativement différentes par ANOVA, comparant les résultats des trois groupes à ceux des deux biologistes entraînés sont présentées dans la Figure 1. Ce sont les participants expérimentés qui appliquent en routine les procédures recommandées qui avaient les résultats les plus proches des valeurs de référence. Pour la majorité des caractéristiques morphologiques étudiées, il n'y avait aucune différence significative entre les résultats de ce groupe et ceux du groupe de participants peu expé-

Tableau 3 : Distribution des caractéristiques morphologiques, coefficient de variation interparticipant correspondant et valeurs de référence.

Caractéristique	Moyenne	Ecart-type	Médiane	min - max	CV (%)	Référence
% spermatozoïdes normaux	15,6	6,3	15	6 - 39	40	7
<i>Anomalies de la tête</i>						
Allongée	12,5	9,0	11	0 - 42	72	10
Amincie	13,3	7,4	12	1 - 40	56	17
Microcéphale	8,0	4,5	7	2 - 24	56	9
Macrocéphale	3,6	1,7	3	1 - 8	47	2
Multiple	2,7	1,2	2	1 - 5	44	4
Postacrosome anormal	36,4	16,0	36	1 - 81	44	53
Acrosome anormal ou absent	52,1	13,5	55	23 - 80	26	62
Total des anomalies de la tête	129	29,9	130	56 - 183	23	157
<i>Anomalies de la pièce int. (PI)</i>						
Reste cytoplasmique	15,0	5,2	15	6 - 28	35	6
Grêle	2,1	2,4	1	0 - 10	114	2
Angulée	20,3	4,7	20	12 - 37	23	14
Total des anomalies de la PI	37	7,6	37	20 - 59	21	22
<i>Anomalies du flagelle</i>						
Absent	3,2	0,8	3	1 - 6	25	3
Court	3,1	4,5	2	0 - 28	145	7
Calibre irrégulier	6,3	5,2	5	0 - 32	83	8
Enroulé	9,6	3,1	9	5 - 18	32	8
Multiple	5,1	1,4	5	2 - 9	27	5
Total des anomalies du flagelle	27	9,7	26	16 - 77	36	31
Total des anomalies	193	33,8	189	110 - 266	18	210
Index d'anomalie multiple, IAM	2,28	0,27	2,25	1,80 – 2,83	12	2,26

Tableau 4 : Caractéristiques morphologiques et coefficient de variation (CV) correspondant en fonction de l'expérience

Caractéristique	< 520 (<10 par semaine) n=45	> 520 (>10 par semaine) n=17	Valeur de p*
	Moyenne (CV %)	Moyenne (CV %)	
% spermatozoïdes normaux	16,9 (40)	12,1 (32)	0,001
<i>Anomalies de la tête</i>			
Allongée	12,8 (64)	12,4 (90)	0,38
Amincie	12,9 (58)	13,8 (52)	0,68
Microcéphale	7,9 (53)	7,3 (52)	0,96
Macrocéphale	3,7 (46)	3,3 (45)	0,53
Multiple	2,4 (42)	3,3 (39)	0,03
Postacrosome anormal	35,5 (47)	40,1 (32)	0,19
Acrosome anormal ou absent	48,8 (26)	60,1 (19)	0,004
Total des anomalies de la tête	124 (25)	140 (18)	0,08
<i>Anomalies de la pièce int. (PI)</i>			
Reste cytoplasmique	15,8 (34)	13,1 (33)	0,08
Grêle	1,8 (128)	3,0 (89)	0,05
Angulée	19,9 (22)	21,2 (26)	0,50
Total des anomalies de la PI	38 (21)	37 (20)	0,99
<i>Anomalies du flagelle</i>			
Absent	3,3 (27)	2,9 (14)	0,13
Court	2,5 (100)	4,8 (160)	0,95
Calibre irrégulier	5,3 (74)	8,8 (83)	0,05
Enroulé	9,2 (32)	10,7 (44)	0,14
Multiple	5,0 (26)	5,7 (26)	0,13
Total des anomalies du flagelle	25 (26)	33 (43)	0,02
Total des anomalies	187 (18)	210 (15)	0,02
Index d'anomalie multiple, IAM	2,24 (12)	2,39 (11)	0,07

* test de Mann-Whitney

Figure 1 : Différences (%) par rapport aux valeurs de référence pour les trois groupes de participants étudiés.

mentés mais ayant de bonnes pratiques. En revanche, c'est le groupe composé de participants n'appliquant pas en routine la standardisation recommandée qui avaient le plus de résultats différant significativement de ceux du premier groupe (expérimenté + bonnes pratiques).

IV. DISCUSSION

Cette étude confirme la forte variabilité interindividuelle des résultats de l'analyse des spermatozoïdes morphologiquement normaux et anormaux déjà rapportée dans plusieurs publications. L'absence constatée de concordance entre techniciens, biologistes ou laboratoires rend difficile la comparaison des résultats au sein d'un même laboratoire ou d'un laboratoire à l'autre. Ceci n'est pas sans conséquence fâcheuse pour la prise en charge des patients d'un point de vue diagnostique et pronostique, notamment parce que la morphologie des spermatozoïdes est connue pour être un déterminant important de la fertilité humaine *in vivo* et *in vitro*.

Par conséquent, une première évaluation de l'étendue de la variabilité dans l'analyse de la morphologie des spermatozoïdes au sein du laboratoire d'andrologie (variabilité intra et inter technicien, CQI) et d'un laboratoire à l'autre (EEQ) est une étape nécessaire pour comprendre ses défauts et l'améliorer. À cet égard, les résultats du présent travail se rapportant à un groupe de biologistes et de techniciens de différentes régions françaises ayant des expériences et des pratiques diverses peuvent être considérés comme représentatifs de la variabilité de base pour cet examen dans un pays développé. Le coefficient de variation global pour le pourcentage de spermatozoïdes normaux était de 40%, se situant dans un intervalle de valeur similaire à ce qui a été précédemment trouvé dans diverses

EEQ sur la morphologie des spermatozoïdes [20, 25, 17], voir Tableau 5. Cependant, il faut préciser que dans cette EEQ, seule la variabilité relative à la classification (pour de mêmes cellules) a été étudiée (à la différence d'EEQ faite à partir de frottis envoyés par la poste).

En revanche, la variabilité pour différentes catégories d'anomalies et pour l'index d'anomalies multiples (MAI) était nettement inférieure. Par exemple, des coefficients de variation de 12%, de 23%, de 25% et de 26% ont été trouvés pour l'IAM, les pièces intermédiaires angulées, les flagelles absents et les acrosomes anormaux ou absents, respectivement. Cette meilleure concordance interindividuelle de l'évaluation en dépit d'expériences et de pratiques diverses ajoutée à la valeur clinique de certaines de ces caractéristiques, par exemple, la quantité de régions acrosomiques anormales [15, 21] ou l'IAM [16, 26], renforce la possible pertinence de l'évaluation en routine des anomalies morphologiques et du calcul de l'IAM qui en découle, à côté de l'évaluation des spermatozoïdes normaux. D'un autre côté, une variabilité assez importante à été trouvée pour d'autres catégories d'anomalies, comme par exemple, la pièce intermédiaire grêle ou le calibre irrégulier du flagelle qui nécessitent une très bonne qualité de coloration et une observation minutieuse.

En général et à l'exception de ces deux dernières catégories, la variabilité trouvée pour les anomalies impliquant la reconnaissance de forme (tête multiple, reste cytoplasmique, pièce intermédiaire angulée, flagelle absent, flagelle enroulé...) était inférieure à celle trouvée pour les catégories d'anomalies impliquant une évaluation de la taille (têtes allongées, amincies, microcéphales, macrocéphales, flagelle écourté... voir Tableau 3), un résultat précédemment rapporté [4]. Si la classification au microscope des catégories où il faut évaluer la taille n'est pas aisée d'emblée du fait des limites des mécanismes de la vision et de son intégration au niveau cérébral, les contrôles de qualité continus dans notre laboratoire ont indiqué que la formation à l'aide d'un oculaire à réticule gradué améliore fortement leur classification qui en retour diminue de manière indiscutable la variabilité intra et interindividuelle.

La variabilité globale dans l'analyse de la morphologie des spermatozoïdes étant mesurée, une seconde étape indispensable pour améliorer les résultats est l'identification des diverses sources de variabilité et l'évaluation de leur poids respectif. La présente étude s'est principalement focalisée sur cette question. Certaines de ces sources sont évidentes, par exemple, l'utilisation de divers systèmes de classification, le manque de définition précise des spermatozoïdes normaux et anormaux pour quelques classifications proposées ou encore, l'absence de règles pour la classification des spermatozoïdes présentant plus d'une anomalie morphologique.

Clairement, le choix de la méthode de classification devrait se fonder sur la preuve de sa validité clinique (relation avec la fertilité *in vivo* et *in vitro*), sa simplicité d'utilisation, la reproductibilité des résultats et l'observation des plus faibles variabilités intra et inter individuelles. Malheureusement, il n'existe à l'heure actuelle aucune méthode ayant démontré satisfaire l'ensemble de ces critères. Au total, des études comparatives soigneusement conçues sont nécessaires pour démontrer quelle méthode de classification est à la fois la plus exacte, la plus reproductible et la plus pertinente sur le plan clinique.

Notre étude a été conçue avec un schéma expérimental ayant quelques similitudes avec celui employé dans l'étude de Baker et Clarke [4] (questionnaire sur l'expérience, observation et la classification des mêmes cellules) et ressemblant dans les modalités techniques au schéma rapporté par Davis *et al.* [11], voir Tableau 5. Ces modalités techniques combinant en temps réel le microscope, la capture vidéo haute résolution et la vidéoprojection simultanée sont régulièrement utilisées dans notre laboratoire pour la formation et constituent un outil de choix pour le contrôle de qualité.

Puisqu'il n'existe aucune méthode établie de mesure objective de la morphologie des spermatozoïdes humains, la question du choix de la méthode fournissant les valeurs de référence est cruciale pour la discussion des résultats des contrôles de qualité interne ou externe. Dans certaines EEQ, les valeurs de référence correspondent à celles obtenues par un panel d'experts en matière d'analyse de sperme [8, 13] (valeur médiane ou moyenne en fonction de la distribution de la caractéristique). Une approche alternative pourrait être d'employer comme valeurs de référence les résultats obtenus par un seul technicien ou biologiste expérimenté dont la précision, la reproductibilité et l'absence de déviation avec le temps ont été montrées [18]. Dans ce travail, nous avons choisi de prendre comme valeurs de référence, celles fournies par les deux biologistes ayant soigneusement examiné la bande vidéo. Il est à noter que ces valeurs «de référence» étaient les plus strictes en comparaison avec celles des participants (Tableau 3).

Nous avons trouvé que les participants évaluant régulièrement un nombre élevé de préparations avaient des évaluations plus strictes que celles des participants plus récemment impliqués ou ayant une pratique plus épisodique. De plus, les participants expérimentés avaient des résultats plus proches des valeurs de référence, bien que parfois différents (le choix de cette approche pourrait avoir conditionné les écarts observés avec les participants les plus entraînés qui ne disposaient pas de moyen d'évaluation

objective de la taille des spermatozoïdes). Le rôle de la formation répétée sur la reproductibilité et la convergence avec les valeurs de référence a été rapporté [23, 13] et il a été montré que l'expérience avait clairement un impact sur les résultats de l'évaluation morphologique des spermatozoïdes humains [4]. Cependant, l'expérience seule n'est pas la garantie d'une variabilité inter individuelle acceptable même si plusieurs résultats d'EEQ indiquent que la variabilité inter individuelle est plus faible parmi des techniciens entraînés que parmi des techniciens novices [14, 4], voir Tableau 5. En effet, la variabilité interindividuelle parmi des observateurs censée être convenablement formés et entraînés peut être relativement élevée [24], voir Tableau 5 et, Tableau 4, notre étude. Ceci indique que d'autres facteurs que l'expérience peuvent moduler les résultats de l'analyse morphologique. Il a été suggéré que certaines étapes pré analytiques, telles que le volume de sperme utilisé, la qualité du frottis ou de la coloration pourraient avoir un impact sur les résultats [22, 5]. Il a aussi été montré qu'une standardisation de la préparation des lames améliorerait la qualité des résultats de l'analyse morphométrique automatisée des têtes de spermatozoïdes [10] (et sans aucun doute aussi, la caractérisation visuelle et la bonne classification des anomalies). A cet égard, dans le présent travail examinant les mêmes spermatozoïdes avec une classification unique, il était intéressant de noter que les participants n'employant pas en routine les procédures recommandées avaient des résultats souvent significativement différents des résultats de référence.

V. CONCLUSION

La présente étude se concentrant sur la variabilité dans la classification des spermatozoïdes morphologiquement normaux et anormaux a montré que l'expérience pratique aussi bien que la conformité aux méthodologies recommandées ont un impact sérieux sur les résultats. Ces données indiquent clairement le besoin fondamental de standardisation, de formation et de contrôle de qualité pour l'évaluation de la morphologie des spermatozoïdes humains.

Remerciements :

Nous tenons particulièrement à remercier tous les participants à cette Evaluation Externe de la Qualité ainsi que B. Foliguet et A. Clavert du Groupe Français d'Assurance de Qualité en Biologie de la Reproduction.

Tableau 5 : Evaluations externes de la qualité (EEQ) de l'évaluation morphologique des spermatozoïdes humains : similitudes et divergences.

	A	B	C	D	E	F	G
Participants, n	29	9	17	96	8	372	62
Participants expérimentés, n	3	9	?	?	7	?	17
Type d'EEQ :							
Atelier / Frottis postés	A (au microscope)	F	F	A (photo)	F	F	A (vidéo)
Test de la reproductibilité	Non	Oui	Oui	Oui	Oui	Non	Non
Qualité du sperme sur les préparations étudiées, n	Normale (homme fécond), 1	Variée, 3	Normale, 24	Variée	Térato. pour 16/20 tests	Normale et subnormale, 12	Térato., 1
Technique de coloration	?	?	? et	Papanicolaou	Papanicolaou	Papanicolaou,	Shorr
Classification utilisée	?	WHO87	Prépa. fraîche	WHO87 ^a , WHO92 ^b , Strict ^c et autre ^d	Strict ^a =6, WHO87=1, Düsseldorf =1	ASCP ^a , WHO87+ WHO92 ^b , Strict ^c	David
CV observés:							
spermatozoïdes normaux	ND*	ND	44% (total) (28% à 62%)	33% ^a , 45% ^b , 65% ^c , 32% ^d	ND (≈ 54% for ^a)	21% ^a , 33% ^b , 50% ^c	40% (tous), 32%(expérimentés)
spermatozoïdes anormaux	64% (tous) 20% (expérimentés)	ND	ND	ND	ND	ND	ND
ête anormale	ND	25%	ND	ND	ND	ND	ND
ête irrégulière	ND	63%	ND	ND	ND	ND	23%
èce intermédiaire anormale	ND	87%	ND	ND	ND	ND	20%
lagelle anormal	ND	63%	ND	ND	ND	ND	36%

A : Jequier et Ukombe, 1983 [14] ; B : Neuwinger *et al.*, 1990 [24] ; C : Matson *et al.*, 1995 [20] ; Davis *et al.*, 1995 [11] ; E : Ombelet *et al.*, 1998 [25] ; F : Keel *et al.*, 2000 [17] ; G : notre étude ; * non déterminé

REFERENCES

1. AUGER J., EUSTACHE F. : Standardisation de la classification morphologique des spermatozoïdes humains selon la méthode de David modifiée. *Andrologie*, 2000, 10 : 358-373.
2. AUGER J., EUSTACHE F., DUCOT B. et al. : Intra- and inter-individual variability in human sperm concentration, motility and vitality assessment during a workshop involving ten laboratories. *Hum. Reprod.*, 2000, 15 : 2360-2368.
3. AUGER J., EUSTACHE F., ANDERSEN A.G. et al. : Sperm morphological defects related to environment, lifestyle and medical history of 1001 male partners of pregnant women from four European cities. *Hum. Reprod.*, 2001, 16 : 2710-2717.
4. BAKER H.W.G., CLARKE G.N. : Sperm morphology : consistency of assessment of the same sperm by different observers. *Clin. Reprod. Fertil.*, 1987, 5 : 37-43.
5. BARROSO G., MERCAN R., OZGUR K. et al. : Intra- and inter-laboratory variability in the assessment of sperm morphology by strict criteria : impact of semen preparation, staining techniques and manual versus computerized analysis. *Hum. Reprod.*, 1999, 14 : 2036-2040.
6. CLEMENTS S., COOKE I.D., BARRATT C.L.R. : Implementing comprehensive quality control in the andrology laboratory. *Hum. Reprod.*, 1995, 10 : 2096-2106.
7. COOPER T.G., NEUWINGER J., BAHRS S., NIESCHLAG E. : Internal quality control of semen analysis. *Fertil. Steril.*, 1992, 58 : 172-178.
8. COOPER T.G., ATKINSON A.D., NIESCHLAG E. : Experience with external quality control in spermatology. *Hum. Reprod.*, 1999, 14 : 765-769.
9. DAVID G., BISSON J.P., CZYGLIK F. et al. : Anomalies morphologiques du spermatozoïde humain. 1) Propositions pour un système de classification. *J. Gynecol. Obstet. Biol. Reprod.*, 1975, suppl. 1 : 17-36.
10. DAVIS R.O., GRAVANCE C.G. : Standardization of specimen preparation, staining, and sampling methods improves automated sperm-head morphometry analysis. *Fertil. Steril.*, 1993, 59 : 1412-1417.
11. DAVIS R.O., GRAVANCE C.G., OVERSTREET J.W. : A standardized test for visual analysis of human sperm morphology. *Fertil. Steril.*, 1995, 63 : 1058-1063.
12. DIXON W.J. : *BMDP statistical software manual*. Berkeley, University of California Press, 1988.
13. FRANKEN D.R., SMITH M., MENKVELD R. et al. : The development of a continuous quality control programme for strict sperm morphology among sub-Saharan African laboratories. *Hum. Reprod.*, 2000, 15 : 667-671.
14. JEQUIER A.M., UKOMBE E.B. : Errors inherent in the performance of a routine semen analysis. *Br. J. Urol.*, 1983, 55 : 434-436.
15. JEULIN C., FENEUX D., SERRES C. et al. : Sperm factors related to failure of human in vitro fertilization. *J. Reprod. Fertil.*, 1986, 76 : 735-744.
16. 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 : 379-384.
17. KEEL B.A., QUINN P., SCHMIDT C.F. Jr et al. : Results of the American Association of Bioanalysts national proficiency programme in andrology. *Hum. Reprod.*, 2000, 15 : 680-686.
18. KNUTH U.A., NEUWINGER J., NIESCHLAG E. : Bias to routine semen analysis by uncontrolled changes in laboratory environment – detection by long term sampling of monthly means for quality control. *Int. J. Androl.*, 1989, 12 : 375-383.
19. KRUGER T.F., ACOSTA A.A., SIMMONS K.F. et al. : Predictive value of abnormal sperm morphology in in vitro fertilization. *Fertil. Steril.*, 1988, 49 : 112-117.
20. MATSON P.L. : External quality assessment for semen analysis and sperm antibody detection: results of a pilot scheme. *Hum. Reprod.*, 1995, 10 : 620-625.
21. MENKVELD R., RHEMREV J.P.T., FRANKEN D.R. et al. : Acrosomal morphology as a novel criterion for male fertility diagnosis : relation with acrosine activity, morphology (strict criteria), and fertilization in vitro. *Fertil. Steril.*, 1996, 65 : 637-644.
22. MESCHÉDE D., KECK C., ZANDER M. et al. : Influence of three different preparation techniques on the results of human sperm morphology analysis. *Int. J. Androl.*, 1993, 16 : 362-369.
23. MORTIMER D. : Technician training and quality control aspects. In : *Practical Laboratory Andrology*. Oxford, Oxford University Press, 1994 : 337-347.
24. NEUWINGER J., BEHRE H.M., NIESCHLAG E. : External quality control in the andrology laboratory: an experimental multicenter trial. *Fertil. Steril.*, 1990, 54 : 308-314.
25. OMBELET W., BOSMANS E., JANSSEN M. et al. : Multicenter study on reproducibility of sperm morphology assessments. *Arch. Androl.*, 1998, 41 : 103-114.
26. SLAMA R., EUSTACHE F., DUCOT B., et al. : Time to pregnancy and semen parameters : a cross-sectional study among fertile couples from four European cities. *Hum. Reprod.*, 2002, 17 : 503-515.
27. WORLD HEALTH ORGANIZATION : *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction*. 3rd edn. Cambridge, Cambridge University Press, 1992.
28. WORLD HEALTH ORGANIZATION : *WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction*. 4th edn. Cambridge, Cambridge University Press, 1992.

ABSTRACT

Factors Influencing the Morphological Classification of Normal and Abnormal Human Spermatozoa: Lessons from an External Quality Assessment Scheme Involving 62 Participants

Jacques AUGER, Florence EUSTACHE

The assessment of human sperm morphology raises a number of methodological problems, making inter-laboratory comparison difficult. Sixty-two technicians and biologists classified the same 100 spermatozoa in a blind study from images captured on a video microscope and projected onto a screen. David's morphology classification, which distinguishes normal and abnormal spermatozoa and calculates a Multiple Anomaly Index (MAI), was used. The same spermatozoa were also classified by two trained biologists using videotaped images to provide reference values. Data analysis mainly focussed on the effect of the level of experience and the possible impact of compliance with standard methods. The inter-observer coefficient of variation was 40% for normal sperm, lower for five defects and, 12% for MAI. Experienced participants had the closest results to the reference values. Moreover, participants not routinely using the recommended methodologies, regardless of their level of experience, consistently showed marked differences compared to reference values. The present data confirm the wide variability in the assessment of normal spermatozoa and show a lower variability for several anomalies and MAI. They underscore the role played by experience and the importance of following the recommended methodologies to obtain relevant and comparable results.

Key words: *routine semen analysis, quality control, multiple anomaly index, sperm morphology assessment*

Manuscrit reçu : mai 2003 ; accepté juin 2003.