

Running Title: Sperm capacitation and male fertility

Title: Multicentric, prospective observational data show sperm capacitation predicts male fertility, and cohort comparison reveals high prevalence of impaired capacitation in men questioning their fertility

Fady Sharara, M.D.,^a Eric Seaman, M.D.,^b Randy Morris, M.D.,^c Jay Schinfeld, M.D.,^d Jennifer Nichols, D.O.,^d Michael Sobel, D.O.,^d Annette Lee, M.D.,^d Stephen Somkuti, M.D., Ph.D.,^d Steven Hirshberg, M.D.,^d Tara Budinetz, D.O.,^d Larry Barmat, M.D.,^d Scott Smith, Ph.D.,^d Gianpiero Palermo, M.D.,^e Zev Rosenwaks, M.D.,^e Natan Bar-Chama, M.D.,^f Joshua Bodie, M.D.,^g John Nichols, M.D., Ph.D.,^h John Payne, M.D.,^h Travis McCoy, M.D.,^h Edward Tarnawa, M.D.,^h Gail Whitman-Elia, M.D.,^h Lauren Weissmann, M.D., M.S.C.E.,ⁱ Mark Leondires, M.D.,^j Cynthia Murdock, M.D.,^j Michael Butcher, D.O.,^k James Kashanian, M.D.,^l Peter Ahlering, M.D.,^m Mira Aubuchon, M.D.,^m G. Charles Ostermeier, Ph.D.,ⁿ Alexander J. Travis, V.M.D., Ph.D.^o

^aVirginia Center for Reproductive Medicine, Reston, VA; ^bNew Jersey Urology, Millburn, NJ; ^cIVF1, Naperville, IL; ^dAbington Reproductive Medicine, Abington, PA; ^eWeill Cornell Medicine, Ronald O. Perleman & Claudia Cohen Center for Reproductive Medicine, New York, NY; ^fRMA New York, New York, NY; ^gUniversity of Minnesota, Urology, Minneapolis, MN; ^hPiedmont Reproductive Endocrinology Group, Greenville, SC; ⁱSouth Jersey Fertility Center, Marlton, NJ; ^jRMA Connecticut, Norwalk, CT; ^kPark Nicollet Sexual Medicine & Male Infertility Clinic, St. Louis Park, MN; ^lWeill Cornell Medicine, Urology, New York, NY; ^mMCRM Fertility, Chesterfield, MO; ⁿAndrovia LifeSciences, Mountainside, NJ; and ^oMaster of Public Health Program, Cornell University, Ithaca, NY

Corresponding author:

Alexander J. Travis
Cornell University
Master of Public Health Program
S3-215, Schurman Hall
Ithaca, NY 14853 USA
ajt32@cornell.edu
607-256-5613

Capsule:

Sperm capacitation prospectively predicted male fertility/probability of generating pregnancy. Reduced capacitation was highly prevalent in men questioning their fertility, even if normozoospermic, relative to a fertile population.

Additional sections not for title page:

Conflict of interest statement: F. Sharara, E. Seaman, R. Morris, G.D. Palermo, and J. Schinfeld, have provided clinical advice to Androvia LifeSciences. G. C. Ostermeier is an employee of Androvia LifeSciences, LLC. G. Palermo is involved in intellectual property with Androvia that extends beyond the current publication. A. J. Travis' laboratory at Cornell identified the underlying technology, which was licensed by, and has been developed by, Androvia LifeSciences. He serves as a consultant to Androvia LifeSciences with duties of a Chief Scientific Officer, and holds equity interest. We know of nothing to disclose for: J. Nichols, M. Sobel, A. Lee, S. Somkuti, S. Hirshberg, T. Budinetz, L. Barmat, S. Smith, Z. Rosenwaks, N. Bar-Chama, J. Bodie, J. Nichols, J. Payne, T. McCoy, E. Tarnawa, G. Whitman-Elia, L. Weissmann, M. Leondires, C. Murdock, M. Butcher, J. Kashanian, P. Ahlering, and M. Aubuchon.

Objective: To assess relationships between sperm capacitation and male fertility.

Design: *Multicentric, prospective observational study:* Data (n=128; 6 clinics) were analyzed to test a published relationship between the percentage of fertilization competent, capacitated sperm (Cap-Score™) and probability of generating pregnancy (PGP) within three cycles of intrauterine insemination (IUI). Logistic regression of total pregnancy outcomes (n=252) assessed fit. *Cohort comparison:* Cap-Scores of Men Questioning their Fertility (MQF; n=2,155; 22 clinics) were compared to 76 fertile men.

Setting: Fertility and reproductive urology clinics; commercial laboratory.

Patients: All patients on whom Cap-Score was performed; all IUI outcomes.

Interventions: None.

Main Outcome Measures: Percent capacitated sperm (Cap-Score); pregnancy rate within 3 IUI cycles; semen analysis (SA; motility, concentration, volume).

Results: New outcomes (n=128) were rank-ordered by Cap-Score and divided into quintiles (25-26/group); chi-square revealed no difference between predicted and observed pregnancies (p=0.809). Total outcomes (n=252; 128 new + 124 previous) were pooled and model recalculated, yielding minor change in PGP (\bar{X} =2.4%), and improved fit (p<0.001). Akaike Information Criterion found the optimal model used Cap-Score alone. Cap-Scores were performed on 2,155 men (SA data available for 1,948). To compare fertilizing ability, men were binned by PGP (\leq 19, 20-29, 30-39, 40-49, 50-59, \geq 60). Distributions of PGP and corresponding Cap-Scores were significantly lower in MQF versus fertile men (p<0.001). Notably, 64% of normozoospermic MQF (757/1,183), had PGP \leq 39% (Cap-Scores \leq 31), versus 25% of fertile men.

Conclusion: Sperm capacitation prospectively predicted male fertility. Reduced capacitation affects many MQF, even if normozoospermic, informing diagnosis versus idiopathic infertility.

Keywords: Pregnancy, infertility, diagnostic, andrology, assisted reproduction

INTRODUCTION:

Infertility has often incorrectly been viewed as a “women’s health” problem, even though men contribute to 40-60% of the cases (1-3). Despite infertility affecting 10-15% of couples globally (4), the field of andrology lacks informative diagnostics (5). Men are often assumed fertile if they have enough morphologically normal, motile sperm to pass current WHO guidelines. This is despite the fact that it is well-known that most male fertility problems are a result of poor sperm function/fertilizing ability and are not detected by traditional semen analysis (SA; (6-8)). Lack of an appropriate diagnostic assessment of fertilizing ability has led to most male infertility cases being classified as “idiopathic,” and repeated calls in the literature for the development of tests capable of evaluating sperm fertilization competency (5, 9-11). New urgency is felt, as it is recognized that traditional SA metrics are falling precipitously in industrialized nations (12), and that male fertility can reflect and be prognostic for general male health (13).

Clinically, this gap between need and available diagnostics has resulted in four serious negative impacts. First, it has placed the onus for extensive and often invasive diagnostic testing almost exclusively on women, with men often going undiagnosed (14, 15). Second, the failure to correctly assess male fertility has resulted in innumerable cycles of intrauterine insemination (IUI) that had low chance of success; these repeated failures are then a basis for diagnosing idiopathic infertility (16, 17). Conversely, efforts to avoid IUI failure due to undiagnosed defects in sperm fertilizing ability have led to a third problem; namely, that couples are sometimes advised to go straight to invasive and expensive procedures such as intracytoplasmic sperm injection (ICSI), when in fact IUI might be effective (18, 19). Fourth, development and use of treatments to improve male infertility have been hampered by lack of an appropriate measure of sperm fertilizing ability that could not only identify which men need treatment, but also then gauge the impact of those interventions (e.g. lifestyle changes in diet, exercise, tobacco or alcohol exposure, surgical repair of varicocele, or treatment with various supplements, etc.) (20, 21). In short, a test that assesses sperm fertilizing ability could provide important benefits, enabling more personalized approaches to achieve pregnancy and to improve male fertility.

One quantifiable measure of sperm function is capacitation status. When sperm enter the female tract, they attain fertilization competence through the process of “capacitation,” in which the head acquires the ability to undergo acrosome exocytosis and the flagellum acquires hyperactivated motility (22-24). Capacitation is achieved in response to stimuli including removal of membrane sterols, and influx of calcium and bicarbonate (24). Over multiple studies, we identified the organization of membrane microdomains having varying compositions of sterols, the ganglioside G_{M1} , and proteins involved in capacitation and acrosome exocytosis (25-31). Using cell biological, pharmacological, and genetic approaches, we identified in murine sperm that G_{M1} regulates transient calcium influxes through R-type, voltage-gated channels that enable acrosome exocytosis (32). Of diagnostic relevance, we found that G_{M1} localization could quantify the percentage of sperm capable of fertilizing (33).

When tested in human sperm, G_{M1} localization indicated capacitation at the level of single cells that could undergo physiological acrosome exocytosis (34, 35). Use of the percentage of sperm in an ejaculate that capacitate (the Cap-Score™ Male Fertility Assay, Androvia LifeSciences, Mountainside, NJ), was validated in terms of precision, variance within samples, and variance between readers (34). Its relationship with male fertility was initially suggested at the level of ejaculates by the finding that higher percentages of capacitated sperm correlated strongly with history of success within 3 or fewer cycles of IUI (36). In repeated samples from the same individual, the percentage of capacitated sperm differed by an average of 6% points of the average for that individual, which is much lower than the variability observed with traditional SA parameters (36). The Cap-Scores of 76 men with known recent fertility had a normal distribution and were significantly greater than the Cap-Scores from 122 men questioning their fertility (MQF; (36)). In the same study, minimal to no relationship was detected between traditional SA parameters (morphology, motility and concentration) and Cap-Scores for those MQF (36).

A single threshold value was then tested for its ability to prospectively identify men predicted to have normal fertility (n=44) versus men predicted to have difficulty generating pregnancy (n=47). In that study, female partners had no factors that precluded their eligibility for IUI (37). Absolute and cumulative pregnancies differed significantly, with a 4.23-fold higher first cycle pregnancy success rate in men scoring above the cut-off (37). There were no differences in maternal or paternal age, or SA metrics, between the outcome groups (37). Because male fertility does not exist as a simple binary, “infertile” or “fertile” state, clinical outcomes data from a single clinic (n=57) were used to define a continuous relationship between the percentage of sperm that can capacitate and male fertility, in the form of the probability of generating pregnancy (PGP) in 3 cycles. The fit of this model was then tested by the addition of 67 outcomes from 5 clinics (total n=124), resulting in a small average change of 4% and improved fit (37). Further analysis revealed that Cap-Score alone, independent of traditional SA measures, provided the optimal model (37).

In the current report, we first performed a multicentric, prospective observational study to determine whether the relationship between the percentage of capacitated sperm and male fertility, as defined by the published model, would match observed clinical pregnancy outcomes under “real-world” conditions. We also compared all Cap-Scores and traditional SA metrics between the MQF cohort and the previously characterized fertile cohort.

MATERIALS AND METHODS:

Study Design

Methods and analyses are reported in accordance with the STROBE checklist for observational studies (38). Current analyses were approved by the Institutional Review Board of Cornell University and Western Institutional Review Board. Prior collection of research samples from 76 fertile men (187 samples) was approved by Western Institutional Review Board.

Quantification of sperm capacitation was performed by means of the Cap-Score, a Laboratory-Developed Test approved for clinical use throughout the United States (Clinical Laboratory Improvement Amendments certified, College of American Pathologists accredited, New Jersey Department of Health licensed, and both laboratory and assay permitted by the New York State Clinical Laboratory Evaluation Program). Although the Cap-Score has recently been approved for home collection and overnight shipment directly to the Androvia lab, all data included in this report were obtained from samples either produced at, or brought to, participating fertility clinics or urology offices. These clinics washed and prepared the samples as part of routine fertility examinations of MQF.

The participating physicians and clinics then shipped samples to Androvia's laboratory where the test was performed. Results were generated and reported to the physicians to inform decision-making, patient counseling, and design and implementation of treatment pathways. Clinics tracked pregnancy outcomes, which were later reported to Androvia. All data were de-identified for analysis. All methods were performed as described previously (37), and presented briefly below.

Settings

Multiple reproductive endocrinology/fertility clinics and reproductive urologists generated data. Clinics providing pregnancy outcomes included: Abington Reproductive Medicine, IVF1, New Jersey Urology, Piedmont Reproductive Endocrinology Group, Virginia Center for Reproductive Medicine, and Weill Cornell Medical College. Details on their IUI practices are provided as Supplemental Materials and Methods.

Participants

All clinical samples on which Cap-Scores were generated, and corresponding clinical IUI outcomes and SA metrics are included (11/2016—7/2019). The only pregnancy outcomes excluded were those using donor sperm on which Cap-Scores were not performed. Kit instructions require 10×10^6 motile sperm on initial count; however, 139 samples from men with fewer sperm were submitted. The Cap-Scores generated were included in our overall analysis, and were also analyzed separately. Selection criteria varied among physicians taking into consideration the details of the specific patient/couple.

Variables and Outcomes

Semen analyses were performed according to WHO guidelines (39). However, morphology assessment varied among clinics precluding its inclusion in overall analysis. Prior investigation

of Cap-Score and morphology in 122 MQF showed no relationship (36). Clinical pregnancies were identified and confirmed as described previously using blood hCG and ultrasonography (37).

Measurement of the Cap-Score

Cap-Scores were all performed by trained personnel at Androvia's laboratory (34). Sample processing and scoring were done as described previously (37). Briefly, semen samples were collected by masturbation and processed using kits provided by Androvia. After liquefaction and washing by density gradient centrifugation, sperm were incubated in mHTF (Irvine Scientific; catalogue # 90126), with/without 2-hydroxypropyl- β -cyclodextrin (CD; Sigma; St. Louis, MO; catalogue # C0926), a stimulus for capacitation. Following incubation, the samples were fixed and shipped overnight to Androvia's laboratory where the Cap-Score test was performed.

Upon receipt, samples were labeled with Alexa Fluor 488-conjugated CTB (Thermo Fisher; Waltham, MA; catalogue # C34775), placed on a slide, and moved to a fluorescence microscope where images were collected.

Readers were trained to identify G_{M1} localization patterns associated with both non-capacitated and capacitated human sperm (40). All readers passed proficiency testing and daily quality assurance testing as described (40). All samples were prepared and scored using these methods except an initial 37 samples provided by Weill Cornell, which were processed and scored prior to formation of Androvia (41).

Bias

Bias could result from inclusion of women with reduced fertility. In a prior study (37), a minimum suite of tests for female factor infertility was defined. The published relationship between Cap-Score and male fertility in the form of probability of generating pregnancy within 3 cycles (PGP) was therefore based on data from women without most identifiable forms of female factor infertility (e.g., tubal occlusion, hydrosalpinges; (37)). Although there is general agreement among clinics regarding tests that should be performed on women before pursuing IUI, we did not exclude data based on the female partner's fertility diagnosis; grounds for inclusion were only that IUI was attempted. Inclusion of infertile/subfertile women would make observed pregnancies fall below those predicted based solely on the male partner's fertility.

Sample preparation kits included instructions that the current version of the test is designed for men with 10 million or more total cells, and 3 million sperm required post-wash. Because of clinical interest, some samples from men with lower numbers (n=139 men) were prepared and submitted. These results were included in the overall count and were also broken out and analyzed separately. Men with moderate to severe oligozoospermia or azoospermia who were not considered eligible for IUI were typically not selected by their physicians to have their sperm's ability to capacitate quantified. Another potential source of bias would include

physicians preferentially selecting men for the assay because of reproductive or other medical history or disclosed behavior/lifestyle. To assess selection bias, we evaluated the data from the one practice performing the test as an initial screen on every man (n=423) versus the rest of the clinics which did not use it in their initial fertility examinations for every patient.

Study Size

The decision when to analyze/report data was determined by the first study, in which we tested the previously published model. That original relationship between Cap-Score and male fertility was based on 124 outcomes. Pregnancy data were collected monthly until that same number was reached with new outcomes (i.e., the dataset doubled). Because outcomes were reported in batches, in practice, 128 new outcomes of patients who completed treatment (achieved pregnancy or completed 3 cycles of IUI) were collected and all are included.

The second observational study evaluating how ability to capacitate is distributed in MQF versus fertile men, and how it compares to traditional SA metrics, was included at this time to provide more in-depth understanding of the prevalence of impaired capacitation in MQF. In this cohort comparison, all Cap-Score data (n=2,155) collected over the study period of about 2.7 years were included in the comparison of distributions. Androvia did not receive SA data for all men; therefore, those results were not included in comparisons of Cap-Scores and SA parameters (n=1,948 men for whom both Cap-Score and SA data were available).

Quantitative Variables

Cap-Score reflects the percentage of sperm having G_{M1} localization patterns consistent with capacitation, out of all sperm having G_{M1} localization patterns (34). Methodologies for traditional SA were established by the WHO (39).

Statistical methods

Statistical analyses (Logistic Regression, Akaike Information Criterion, ANOVA, Chi-Square and t-tests) were carried out in XLSTAT Version 2019.2.2.59398. For prospective comparison of the predicted PGP versus observed pregnancies, we rank ordered results by Cap-Score, and then divided the data into quintiles. The expected number of pregnancies was calculated by summing the PGP values in each quintile ($expected \# \text{ preg} = average \text{ PGP} * n = \frac{\sum_{i=1}^n x_i}{n} * n$), with PGP being predicted by the previously published logistic regression model (37). A goodness of fit chi-squared statistic was performed to determine whether predicted and observed outcomes differed.

Following best practice of having analyses confirmed/performed by independent statisticians, Singular Value Consulting (Houston, TX) was contracted and given Androvia's complete raw data set related to this study, to assess both appropriateness of analyses and determine their accuracy. Statistics and logistic regression analysis were carried out in R (42) and SciPy (43).

RESULTS:

The percentage of capacitated, fertilization competent sperm and traditional SA results were measured for men from 6 clinics (n=292), with pregnancy outcomes collected subsequently. Of these patients, 128 finished treatment (i.e., the couple became pregnant within, or completed 3 cycles of IUI) when data were analyzed. Three tests were employed to assess the predictive relationship between sperm capacitation and male fertility as defined previously for Cap-Score and PGP.

Prospective Test of the Predictive Relationship Between Capacitation and Male Fertility

First, to test whether the new data on Cap Scores and pregnancy outcomes were consistent with the previously published model,(37) we rank ordered the results by Cap-Score and divided them into quintiles (n=25 or 26 per group). The expected number of pregnancies for each quintile was calculated using PGPs that were predicted by that logistic regression model (37). The number of pregnancies observed and those predicted are presented in Table 1. In each quintile, the differences between observed and expected numbers of pregnancies are as expected due to the uncertainty in the model. To quantify this statement, we computed a chi squared statistic ($\chi^2 = 2.28$). We compared this value to a chi square random variable with 5 degrees of freedom. Such a random variable would have a mean of 5 and a 95% confidence interval of (0.83, 12.83). Our observed value of 2.28 is well within the confidence interval, indicating that our results are typical of what one would expect based on the logistic model. In short, the pregnancies prospectively predicted by the model are consistent with those observed ($p=0.809$, showing no difference between predicted and observed pregnancies).

Evaluation of Fit of the Logistic Model

Second, the new outcomes were added to the prior 124. Logistic regression models PGP as a function of Cap-Score as

$$\text{PGP} = 1/(1 + \exp(-(a + b*\text{Cap-Score})))$$

where the coefficients a and b are determined from data. Using the full data set (n=252) we obtained the estimates $a = -2.301$ and $b = 0.061$. The fact that b is positive shows that PGP increases with increasing Cap-Score. The p -values associated with both coefficients were less than 0.001.

The new logistic regression model was consistent with the previous model, which was demonstrated by overlapping confidence intervals for the logistic regression coefficients and by how similar the predictions were. The previous intercept term a was -2.863 with a 95% confidence interval of (-4.555, -1.331). The new estimate for a is -2.301 with a 95% confidence interval of (-3.316, -0.330). The previous linear term b was 0.078 with a 95% confidence interval of (0.029, 0.131). The new estimate for b is 0.061 with a 95% confidence interval of (-0.004, 0.095). In each case, the new coefficient estimates are within the confidence intervals of the previous model and vice versa. Overlapping 95% confidence intervals show that there is no

significant change in the logistic regression coefficients when the number of observations in the data set was doubled ($p > 0.05$).

Looked at more simplistically, only a slight average change in the predicted PGP ($\bar{X} = 2.4\%$), was noted from the original model when the new data were added, and fit improved. The greatest divergence from the original model occurred in the lower and higher range of Cap-Scores where there were not only fewer observations, but also (as discussed below), some differences in practice based on the Cap-Score results (Figure 1).

The third test of the relationship between capacitation and male fertility involved discerning whether the inclusion of one or more traditional SA parameters would improve fit. To test this, logistic regression models were fit on the combined data set using Cap-Score and SA measures alone and in every possible combination (Supplemental Table 1). The Akaike Information Criterion (AIC; (44)) was performed to test the relative quality of the models. In brief, the AIC penalizes increasing model complexity without a reciprocal increase in fit. Cap-Score alone was found to provide the optimal model, underscoring that capacitation served as the primary metric of male fertility.

Impact of Maternal Age

Use of IUI data enabled us to focus on male fertility, in that the clinics confirmed at least a minimum degree of female fertility. However, the impacts of advanced maternal age on multiple aspects of female fertility are well documented (45). To test whether maternal age impacted the relationship defined for male fertility, we combined the outcomes for which we had maternal age. When maternal age was added as a term in the logistic regression, the coefficient of age was not significant ($p = 0.42$).

Additionally, we disaggregated these data into the following maternal age groups: ≤ 29 , 30-34, 35-39, and ≥ 40 (Supplemental Table 2). No difference was observed between predicted and observed pregnancy outcomes across maternal age groups ($\chi^2 = 0.585$; $p = 0.965$; four degrees of freedom). Analysis of variance showed that Cap-Scores did not vary across maternal age stratifications ($p = 0.266$). Although female age and fertility are indisputably linked, if a woman was found eligible for IUI, then sperm capacitation accurately predicted pregnancy outcomes across maternal ages. Limitations in interpretation are discussed further below.

Although not necessarily related to age, other maternal effects might manifest themselves in failure to carry to term. As a preliminary investigation of whether pregnancies from IUI might be more likely to result in miscarriage, data were assessed from one clinic of 38 couples pregnant by IUI and 23 by natural conception (NC). There were no differences in couples that miscarried (34 and 35% with IUI and NC respectively) or delivered (66 and 65% with IUI and NC respectively).

Cohort Comparison of Men Questioning Their Fertility vs Fertile Men

To evaluate whether the percentage of capacitated sperm in a man's ejaculate differed between MQF and fertile men, we compared all data generated from the clinics (n=2,155 men, 22 clinics; 28.77 ± 7.53 ($\bar{X} \pm SD$)) against a cohort of men with known fertility (n=76 men, 187 samples, 35.34 ± 7.70 ($\bar{X} \pm SD$); (36)). The distribution of Cap-Scores in MQF was significantly different from that in fertile men (Fig. 2; $p < 0.001$), with 81% (1,741/2,155) falling below the fertile mean of 35.3 (36).

Of these 2,155 men, accompanying SA data were available for 1,948. Table 2 shows the distribution of data relating Cap-Scores, PGP, and traditional SA metrics. Because the relationship between Cap-Score and PGP is not linear, data are presented in bins by PGP (≤ 19 , 20-29, 30-39, 40-49, 50-59, ≥ 60). The lower distribution of Cap-Scores and associated PGPs is revealed in this presentation through several comparisons. For example, 67% of MQF (1,313/1,948) had PGPs ≤ 39 , in comparison to 25% of fertile men (19/76).

Consistent with multiple prior reports, traditional SA results did not correlate with sperm fertilizing ability or male fertility. Based on volume, concentration, and motility, 61% (1,183/1,948) of all MQF were normozoospermic based on WHO criteria. Of these normozoospermic men, 64% (757/1,183) had PGPs ≤ 39 . Failure to generate pregnancy in normozoospermic men would typically result in a diagnosis of idiopathic infertility; these data revealed that impaired sperm capacitation (relative to fertile men) was highly prevalent in MQF. Lastly, impaired sperm capacitation was equally prevalent regardless of an individual man's passing any single or multiple SA metric(s), or those having >10 million total motile cells, which is sometimes thought of as an indicator of minimally acceptable overall semen quality (TMC; $p=0.987$). The majority of MQF had >10 million TMC (93%, 1,809/1,948), but 66% of them had PGPs ≤ 39 (1,200/1,809).

One potential limitation or source of bias in interpreting these data would be if clinicians were successful at identifying men who would have "idiopathic infertility" based on habitus or history, and preferentially ordered Cap-Scores on these men. To evaluate the existence or impact of this potential confounder, we disaggregated the 423 Cap-Score data from the Virginia Center for Reproductive Medicine, which was the only clinic to perform the assay on all eligible men, and compared them against the remaining data from the other clinics. No difference was found when using the Mann-Whitney comparison of two samples ($p=0.107$).

Relationship of Cap-Score/PGP and Traditional SA Metrics

Previously, minimal to no relationship was found between Cap-Score and SA metrics (36). Here, we re-evaluated whether relationships might be revealed based on the considerably larger sample size (1,948 versus 122). Supplemental Figure 3 shows scatterplots and associated regressions exploring potential relationships between volume, motility and concentration with Cap-Score. No relationship was found between volume and Cap-Score ($r^2 < 0.001$, $p=0.65$).

Small, but statistically significant relationships were found for motility and concentration ($p < 0.001$ for each). Motility was found to contribute ~2% to the Cap-Score ($r^2 = 0.018$) and concentration was found to contribute ~1% to the Cap-Score ($r^2 = 0.013$). These data support prior reports that traditional SA parameters have little relationship with the fertilizing ability of sperm, or male fertility.

DISCUSSION:

These studies yielded several findings: 1) a measure of sperm capacitation, the Cap-Score, prospectively predicted male fertility across diverse clinical settings; 2) the previously defined mathematical relationship between Cap-Score and a metric of male fertility, the probability of generating pregnancy within 3 cycles, changed minimally with a doubling of the outcomes dataset; 3) impaired or reduced capacitation ability was highly prevalent in MQF; and 4) there was minimal to no relationship between sperm capacitation and traditional SA metrics.

Interpretation and Comparison with Other Studies

These data confirm that traditional SA metrics fail to identify impairments in fertilizing ability, which typically lead to diagnosis of idiopathic infertility (6-8). The predictive power of measuring capacitation confirms the important contribution of male factor in determining successful generation of pregnancy, and validates prior calls for development of tests of sperm function/fertilizing ability (5, 10, 11). Sperm capacitation involves a number of intracellular signaling and metabolic responses, presenting multiple alternative metrics such as protein tyrosine phosphorylation events, phospholipid scramblase activity, membrane potential, intracellular pH, etc. (46). Despite capacitation having first been identified close to 70 years ago (22, 23), clinical measurement of this essential component of male fertility is not commonly performed because its predictive relationship with fertility is just now being described, and a practical means of measurement has been lacking.

Strengths and Limitations

To test the relationship between sperm capacitation and male fertility, we utilized pregnancies within 3 cycles of IUI as the outcome measure. This design enabled more rigorous and focused evaluation of male fertility by providing some control regarding timing of inseminations relative to ovulation and a basic level of female fertility. Although they also control timing, classical in vitro fertilization (IVF) and ICSI bypass important physiological aspects of male fertility.

Multicentric observational data have the advantage of being generated under “real world” conditions reflecting diverse patient bases and clinical practices, and avoid potential unconscious bias with non-randomized, directed assignment to interventions. The prospective

nature of testing the predicted PGP, and inclusion of *all* non-donor pregnancy outcomes later observed were primary strengths of the first study. The primary strengths of the cohort comparison were the size of the pool of MQF and inclusion of *all* clinical data.

However, these studies investigating the relationship of sperm capacitation and male fertility do have several limitations worth noting. Of greatest importance, the logistic relationship between Cap-Score and male fertility in the form of PGP is predicated upon a fertile female partner. Inclusion of some women having female factor infertility would cause a systematic bias of lowering observed pregnancies relative to predicted. We did not see strong evidence for that here, although there was minor reduction in observed pregnancies for men with high Cap-Scores.

Another bias might have had the opposite effect and increased observed pregnancies; namely, several participating physicians reported modifying their clinical practices when receiving a result of a low Cap-Score. For example, several recommended to their patients with impaired capacitation ability that they make changes in lifestyle, take nutritional supplements, undergo varicocele repair, and/or have two inseminations performed in a single IUI cycle. The effects of these changes in practice might be reflected in the new outcomes, which were slightly elevated relative to those predicted for men with low Cap-Scores. Although the two logistic regression equations did not differ statistically, the potential impact of these changes in practice argues for the continued use of the original equation (37) in reporting of Cap-Scores.

Interpretation of outcomes data stratified by maternal age must be viewed with caution. The lack of difference across age ranges may result, in part, from the original relationship between Cap-Score and PGP being defined using clinical pregnancy outcomes generated from a variety of maternal ages (37). Although there was no difference between predicted and observed pregnancies for women ≥ 40 , it must be noted that the sample size of that group was the smallest of any age group tested.

A potential source of “noise” in the cohort comparison is the fact that the current SA data were generated by multiple andrologists at different clinics. While providing the advantage of a more diverse patient base, this approach undoubtedly introduced variations in technique and practice, such as those leading to our inability to compare morphology data across clinics. The recent development and regulatory approval of an “at home collection kit” could potentially help reduce this confounder in future studies by having various SA parameters as well as Cap-Score all performed at the same laboratory.

Implications for Clinical Practice

These results demonstrate that the percentage of capacitated sperm can provide important predictive information about male fertility, directly impacting a couple’s chances of conception. Tests of capacitation, such as the Cap-Score, can provide a functional complement to the

traditional SA. These can aid in identifying impairments in fertilizing ability that might otherwise only be found through repeated failed attempts at conceiving, resulting in diagnoses of “idiopathic infertility” and their associated physical, emotional and financial costs. Indeed, a successful measure of capacitation has been modeled to not only improve outcomes but also reduce cost per couple (47). A straightforward application for predictive information on male fertility is the personalized counseling and treatment of couples seeking assistance with fertility. When considered as part of the couple’s medical findings and personal context, this information will help clinicians and couples identify an approach that is optimal for them at that point, whether it be tailored expectant management, IUI, IVF or ICSI.

A finding of impaired capacitation could also identify those men who stand to benefit from seeing reproductive specialists and undergoing various treatments to improve male fertility, including change in lifestyle, taking of nutritional supplements, or undergoing varicocele repair as appropriate (20). A quantifiable metric of male fertility would also provide a way to assess response to such treatment. Measurement of impact on capacitation might also enable optimization of cryopreservation or semen handling practices (34).

Other applications with clinical relevance might include the testing of various drugs or nutritional supplements designed to promote male fertility or act as male contraceptives (whether intended or off-target). Whether sperm fertilizing ability can provide a window into the overall future health of a man, as is being discussed for other SA metrics (13), is an intriguing possibility that will require new research. This line of investigation could also be facilitated by collection of semen samples at home, since that would broaden geographic availability and overcome social and/or economic barriers such as concerns of privacy or conflicts with employment.

The present findings prospectively show a clear relationship between capacitation and male fertility and reveal a very high prevalence of impaired capacitation in men having difficulty conceiving. Together, these findings demonstrate that capacitation is a highly sensitive indicator of male fertility, and show both the need and ability to bring men back into the fertility equation, complementing the multiple assays performed on their female partners.

Acknowledgements: The authors are grateful to doctors and clinics who have utilized the assay but who did not meet the criteria for authorship, the patients and staff at all participating clinics, and the technical staff in Androvia's laboratory, particularly: C. Cardona, B. Kloos, R. Mendoza, M. Moody, A. Simpson.

Funding: Androvia LifeSciences supported costs associated with performing the Cap-Score and maintaining its database; individual clinics supported their own participation; Cornell University supported A. Travis.

Figures

Cap-Score™ Quintile	N	Observed pregnancies	Predicted pregnancies $\pm \sigma$
1 st	26	8	5.46 \pm 2.07
2 nd	25	7	6.98 \pm 2.24
3 rd	26	11	8.84 \pm 2.24
4 th	25	8	10.40 \pm 2.46
5 th	26	15	14.58 \pm 2.49

Table 1. Prospective test of predicted probability of generating pregnancy based on Cap-Score versus pregnancies observed within 3 cycles. There were no differences between predicted and observed pregnancies ($\chi^2=2.28$, with 5 degrees of freedom; $p=0.809$).

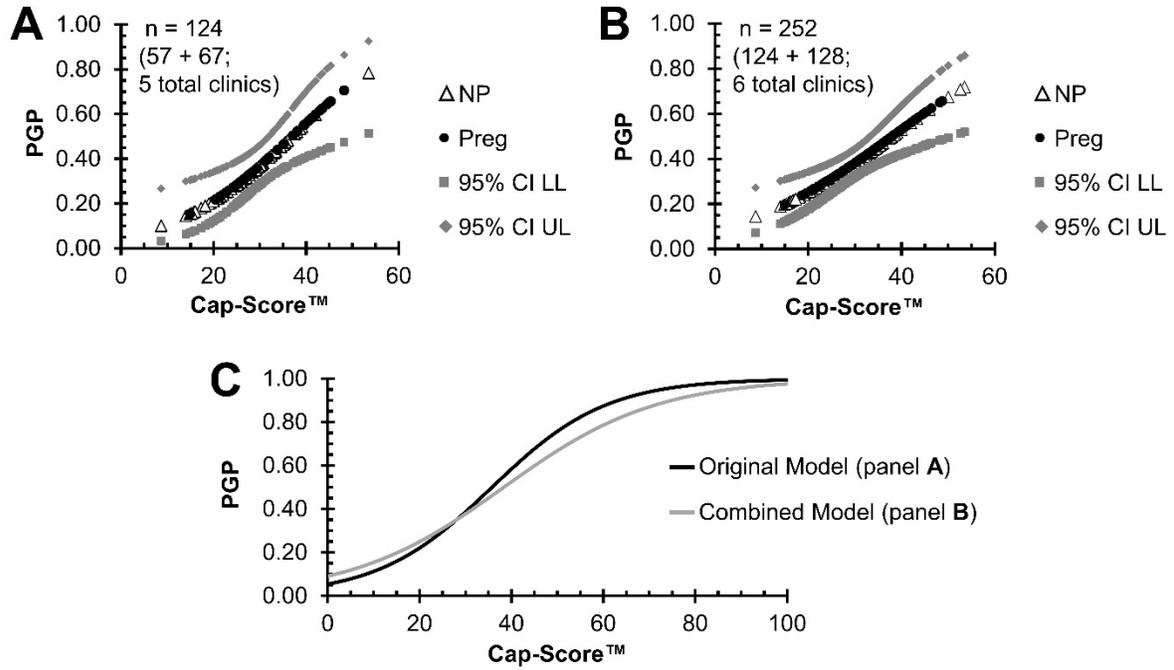


Figure 1. Original (A; 37) and combined (B) logistic regression models defining the relationship between Cap-Score and Probability of Generating Pregnancy (PGP) within 3 cycles. Overlay of original and combined models (C). Non-pregnant cycles (NP); Cycles resulting in pregnancy (Preg); lower limit confidence interval (CI LL); upper limit confidence interval (CI UL).

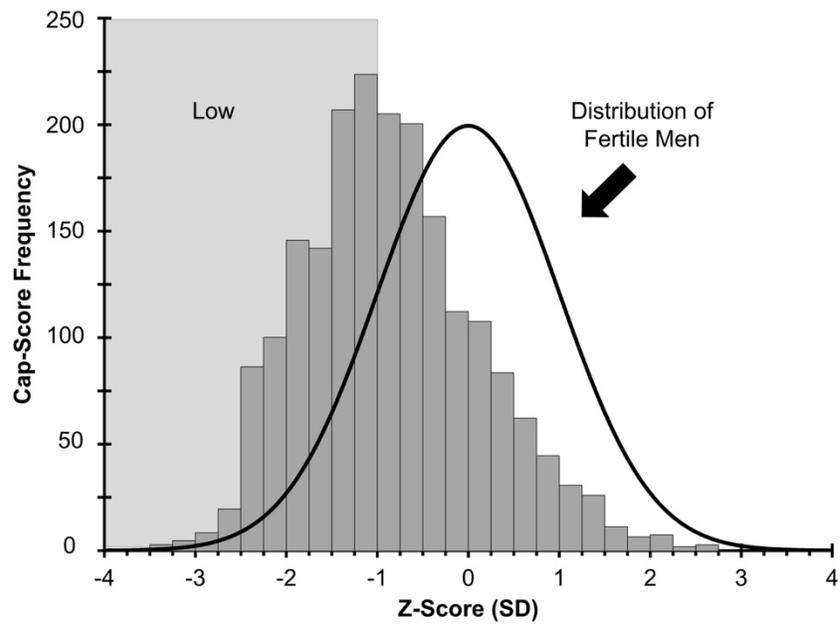


Figure 2. Cap-Scores from 2,155 men questioning their fertility (histogram) were significantly lower than the distribution of Cap-Scores previously defined for a cohort of fertile men (black curve approximates the normal distribution of a fertile cohort, $p < 0.001$). The x-axis shows Z-scores, with the mean of 35.3 set to 0, and every unit equal to one standard deviation of 7.7 (36).

Cap-Score (%)	PGP (%)	% of all men having fertility exams	% normozoospermic men having fertility exams	% men having fertility exams >10M TMC	% fertile men
≤ 18	≤ 19	8% (151/1,948)	6% (69/1,183)	7% (128/1,809)	1% (1/76)
19 - 25	20 - 29	28% (551/1,948)	27% (322/1,183)	28% (499/1,809)	9% (7/76)
26 - 31	30 - 39	31% (611/1,948)	31% (366/1,183)	32% (573/1,809)	14% (11/76)
32 - 36	40 - 49	17% (330/1,948)	19% (224/1,183)	18% (320/1,809)	36% (27/76)
37 - 42	50 - 59	10% (186/1,948)	10% (124/1,183)	10% (176/1,809)	24% (18/76)
> 42	≥ 60	6% (119/1,948)	7% (78/1,183)	6% (113/1,809)	16% (12/76)

Table 2. Distribution of data relating Cap-Scores, PGP, and traditional SA metrics. A nonlinear relationship exists between Cap-Score and PGP. Thus, the data bins presented were established using PGP. The lower distribution of Cap-Scores, and associated PGPs, in men having fertility exams is demonstrated through the highlighted comparisons.

Supplemental Figures

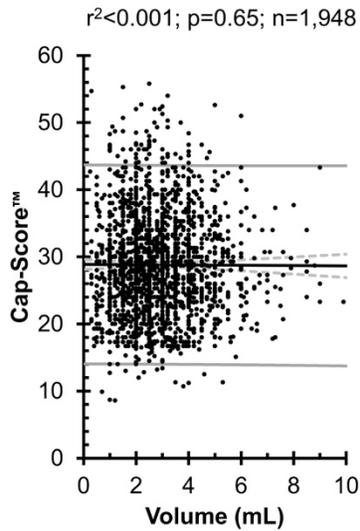
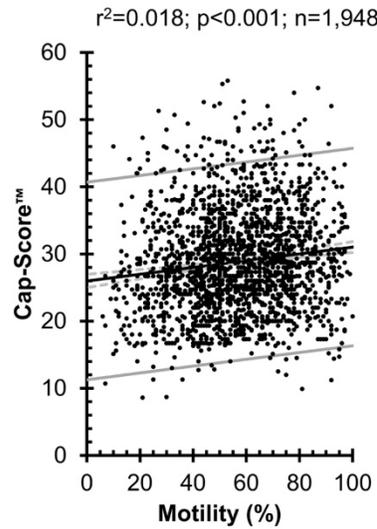
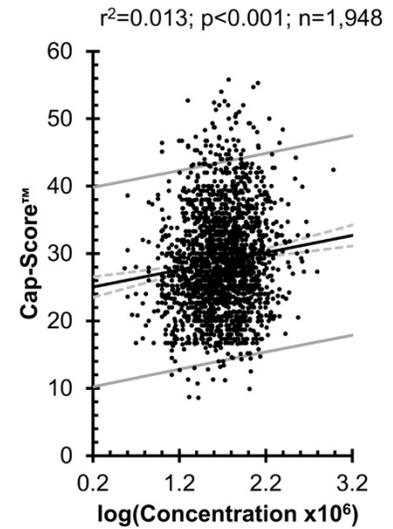
Model	AIC
Cap-Score	318.6
Motility	331.1
Concentration	333.7
Volume	333.1
Cap-Score + Motility	319.1
Cap-Score + Concentration	320.6
Cap-Score + Volume	319.1
Motility + Concentration	333.0
Motility + Volume	332.0
Concentration + Volume	334.7
Cap-Score + Motility + Concentration	320.6
Cap-Score + Motility + Volume	319.4
Cap-Score + Concentration + Volume	321.1
Concentration + Motility + Volume	333.9
Cap-Score + Concentration + Motility + Volume	320.8

AIC: Akaike Information Criterion. Model: shows the measures considered.

Supplemental Table 1. AIC was performed to test whether the inclusion of one or more traditional SA parameters would improve PGP fit. Briefly, the AIC penalizes increasing model complexity without a reciprocal increase in fit. Lower AIC values reflect the most appropriate models. Cap-Score alone was found to provide the optimal model, underscoring that capacitation served as the primary metric of male fertility.

Age range	N	Observed pregnancies	Predicted pregnancies $\pm \sigma$
≤ 29	34	12	11.85 \pm 2.68
30-34	115	46	42.55 \pm 4.99
35-39	66	22	22.64 \pm 3.73
≥ 40	16	4	5.22 \pm 1.78

Supplemental Table 2. Test of impact of maternal age on the relationship between Cap-Score and male fertility. All outcomes data were disaggregated into the defined age ranges. Predicted pregnancies were calculated by summing PGP values within an age range, with PGP being predicted by the original logistic regression model. No difference was observed between predicted and observed pregnancy outcomes, across maternal age groups (chi-square=0.585; $p=0.965$; four degrees of freedom).

A**B****C**

Supplemental Figure 1. Scatterplots showing no relationship between volume and Cap-Score, and minimal relationships between motility and Cap-Score and concentration and Cap-Score. Note that one outlier data point (volume = 15 ml; Cap-Score = 17.9%) was removed from plot A to facilitate visual discrimination of the majority of the data points. The outlier was included in the analysis of relationship between volume and Cap-Score.

SUPPLEMENTAL METHODS [PLEASE NOTE THAT THE WORDING IN THE FOLLOWING SECTIONS WILL NEED TO BE REVISED TO AVOID SELF-PLAGIARISM. PLEASE REVIEW THE INFORMATION FOR YOUR SPECIFIC CLINIC TO DETERMINE WHETHER THE PREVIOUSLY PUBLISHED PRACTICES ARE STILL ACCURATE. PLEASE MAKE ANY NECESSARY CHANGES USING THE TRACK CHANGES FUNCTION.]

Patients

Diagnosis of female factor that prevented a couple from pursuing IUI precluded their inclusion in the observational study. Similarly, outcomes from patients who did not pursue any form of assisted reproduction, or pursued IVF or ICSI without IUI first, were also excluded. Details of practices at clinics providing outcomes data, and their patient bases, were described previously (37). Typical practices are summarized here.

Intrauterine Insemination

Abington Reproductive Medicine: IUI was performed in stimulated cycles. Patients were stimulated either with clomiphene citrate (CC), letrozole (Let), or gonadotropins. All patients were inseminated 24-36 hours after human chorionic gonadotropin (hCG) injection. In rare cases, patients had a second insemination the following day, primarily due to low sperm numbers in the initial sample. Semen samples were produced by masturbation and allowed to liquefy. Semen analysis was performed to assess volume and concentration. Samples were washed as follows: First, 1ml of warmed lower medium was pipetted into the bottom of 15ml conical tube, then 1ml of warmed upper medium was slowly layered on top. The semen sample was carefully layered on top. After centrifugation at 400xg for 20 minutes, the supernatant was removed and the pellet was resuspended in 0.25-2.0 ml of wash medium with protein (volume dependent on the size of the sperm pellet). Post-wash count and motility were assessed. The sample was then centrifuged for 10 minutes at 400xg, and the supernatant was carefully removed. The pellet was resuspended in 0.5 ml wash media with protein and used for insemination.

IVF1: IUI was performed in stimulated cycles. Patients were stimulated either with CC, Let, or gonadotropins. For some of the patients stimulated with CC or Let, ovulation detection was performed by urine LH test. Patients tested their urine sample once or twice a day and were inseminated 20 to 24 hours after LH surge. Other patients stimulated with CC or Let were monitored at the fertility center and inseminated 30-36 hours after human chorionic gonadotropin (hCG) injection. Ovulation was triggered with hCG when a patient's follicles reached 20 mm or more in diameter. Patients stimulated with gonadotropins were inseminated 30-36 hours after hCG trigger. Ovulation was triggered with hCG when a patient's follicles reached at least 17 mm in diameter. Semen samples were produced by masturbation either at the fertility center or home. Samples underwent a simple wash and the pellet was resuspended in 0.3 ml of medium and used for insemination.

New Jersey Urology: No data related to female factor were obtained. Data from 4 patients who had a Cap-Score™ performed and responded to a follow-up questionnaire to self-report clinical outcomes were included in the test of the original model generated by logistic regression. Data were collected between 12/2016 and 7/2017.

Ronald O. Perelman & Claudia Cohen Center for Reproductive Medicine & Infertility, Weill Cornell Medicine: Ovarian stimulation was performed with CC at a dose of 50 or 100 mg daily for five days. The response to stimulation and endometrial thickness were monitored by serial transvaginal ultrasounds. Serum hormone assays were also used to measure estradiol and LH levels. In the absence of LH surge, ovulation was triggered with 10 000 IU human chorionic gonadotropin (hCG) when the dominant follicle(s) reached 20 mm. IUI was performed within 24 h after hCG injection. Semen samples were collected at the laboratory after 2–5 days of abstinence. Semen analysis was performed after 30 min of liquefaction. The samples were first diluted in HEPES-buffered human tubal fluid supplemented with human serum albumin for centrifugation at 600g for 10 min. For each sample, the pellet was then resuspended and layered on a density gradient (. It was then centrifuged for 10 min at 300g. The bottom layer containing motile spermatozoa was collected by aspiration with a glass Pasteur pipette and resuspended for a final 10 min centrifugation at 600g to remove silica gel particles. The final pellet was resuspended in 0.5 ml of medium and used for insemination after reassessing concentration and motility.

Virginia Center for Reproductive Medicine: IUI was performed in stimulated cycles. Depending on the patient's medical history, stimulations were done either with Let, Tamoxifen, gonadotropins or a combination of medications. Ovulation was triggered with hCG when a patient's follicles reached 18 to 20 mm in diameter. Time from hCG trigger and insemination depended on whether there was single or double insemination. If single insemination, IUI was performed 36 hours after the trigger. If double insemination, the first IUI was done 24 hours after the trigger and the second IUI was done 48 hours thereafter. Cycles were supplemented with progesterone, starting the night after the insemination. Patient stayed on progesterone until nine weeks of pregnancy.

Semen samples were kept in a 36°C warmer for 30 minutes for liquefaction. Semen analysis was then performed to assess volume, concentration, motility and morphology. The semen sample was then divided into two equal volumes between two 14-ml conical tubes. Two ml of pre-warmed Quinn's Sperm Wash was added to each tube and mixed by pipetting. After centrifuging at 1500 rpm for 5 minutes, the supernatant was removed from each tube, and both pellets were combined into one tube. Another 2 ml of warm Quinn's Sperm Wash was added to the combined pellet and mixed. After centrifuging at 1500 rpm for 5 minutes, the supernatant was removed until 0.3-0.5mL media was left covering the pellet. The medium-covered pellet was then kept in the warmer to allow the sperm to swim up for 2-4 hrs. About 30 minutes before the scheduled IUI time, the medium containing motile sperm was removed from the pellet and placed in a new, pre-warmed tube ready for IUI. The volume, concentration, and motility of the final media were assessed to calculate the percentage of motile sperm recovered before IUI.

Pregnancy outcome

Abington Reproductive Medicine: Pregnancies were confirmed by beta hCG blood tests starting 14 days after ovulation (confirmed by LH and progesterone levels). All blood tests were repeated every 48-72 hours and ultrasound typically scheduled at 5.5 weeks of gestational age.

IVF1: Clinical evidence of pregnancy was determined by beta hCG blood levels. If positive, the test was repeated two days later. If the hCG rise was deemed to be appropriate, then the patient was brought back to the office for a transvaginal ultrasound when it was predicted that the hCG level would be at least 2000 IU/mL. If the hCG level did not rise appropriately, then the patient would return for additional hCG levels. A clinical pregnancy was determined to be present if a fetal pole with evidence of heart motion was seen.

Ronald O. Perelman & Claudia Cohen Center for Reproductive Medicine & Infertility, Weill Cornell Medicine: Clinical pregnancies were identified by the presence of at least one fetal heartbeat using ultrasound.

Virginia Center for Reproductive Medicine: Two weeks after insemination a urine test was performed; if positive, hCG and progesterone blood levels were determined. Blood tests were repeated every two days, to make sure that hCG was doubling every 48 hours. Once hCG blood levels of 1000 mIU/mL were detected, ultrasonography was performed.

References:

1. Agarwal A, Mulgund A, Hamada A, Chyatte MR. A unique view on male infertility around the globe. *Reprod Biol Endocrinol* 2015;13:37.
2. Mehta A, Nangia AK, Dupree JM, Smith JF. Limitations and barriers in access to care for male factor infertility. *Fertil Steril* 2016;105:1128-37.
3. Petok WD. Infertility counseling (or the lack thereof) of the forgotten male partner. *Fertil Steril* 2015;104:260-6.
4. Sharma R, Biedenharn KR, Fedor JM, Agarwal A. Lifestyle factors and reproductive health: taking control of your fertility. *Reprod Biol Endocrinol* 2013;11:66.
5. Barratt CLR, De Jonge CJ, Sharpe RM. 'Man Up': the importance and strategy for placing male reproductive health centre stage in the political and research agenda. *Hum Reprod* 2018;33:541-5.
6. 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.
7. 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.
8. van der Steeg JW, Steures P, Eijkemans MJ, JD FH, Hompes PG, Kremer JA *et al.* Role of semen analysis in subfertile couples. *Fertil Steril* 2011;95:1013-9.
9. Lamb DJ. Semen analysis in 21st century medicine: the need for sperm function testing. *Asian J Androl* 2010;12:64-70.
10. Oehninger S, Franken DR, Ombelet W. Sperm functional tests. *Fertil Steril* 2014;102:1528-33.
11. Wang C, Swerdloff RS. Limitations of semen analysis as a test of male fertility and anticipated needs from newer tests. *Fertil Steril* 2014;102:1502-7.
12. Levine H, Jorgensen N, Martino-Andrade A, Mendiola J, Weksler-Derri D, Mindlis I *et al.* Temporal trends in sperm count: a systematic review and meta-regression analysis. *Hum Reprod Update* 2017;23:646-59.
13. De Jonge C, Barratt CLR. The present crisis in male reproductive health: an urgent need for a political, social, and research roadmap. *Andrology* 2019.
14. Steinkeler JA, Woodfield CA, Lazarus E, Hillstrom MM. Female infertility: a systematic approach to radiologic imaging and diagnosis. *Radiographics* 2009;29:1353-70.
15. Stevenson EL, Hershberger PE, Bergh PA. Evidence-Based Care for Couples With Infertility. *J Obstet Gynecol Neonatal Nurs* 2016;45:100-10; quiz e1-2.
16. Ruiz A, Remohí J, Minguez Y, Guanes PP, Simón C, Pellicer A. The role of in vitro fertilization and intracytoplasmic sperm injection in couples with unexplained infertility after failed intrauterine insemination. *Fertility and sterility* 1997;68:171-3.
17. Bungum L, Bungum M, Humaidan P, Andersen CY. A strategy for treatment of couples with unexplained infertility who failed to conceive after intrauterine insemination. *Reproductive biomedicine online* 2004;8:584-9.
18. Bhattacharya S, Hamilton MP, Shaaban M, Khalaf Y, Seddler M, Ghobara T *et al.* Conventional in-vitro fertilisation versus intracytoplasmic sperm injection for the treatment of non-male-factor infertility: a randomised controlled trial. *Lancet* 2001;357:2075-9.
19. Evers JL. Santa Claus in the fertility clinic. *Hum Reprod* 2016;31:1381-2.
20. Aly S, Seaman E. Improvement in sperm capacitation measured by Cap-Score(TM) demonstrated after varicocele repair. *J Reprod Med Gynecol Obstet* 2018;3.
21. Hayden RP, Flannigan R, Schlegel PN. The Role of Lifestyle in Male Infertility: Diet, Physical Activity, and Body Habitus. *Curr Urol Rep* 2018;19:56.
22. Austin CR. The capacitation of the mammalian sperm. *Nature* 1952;170:326.

23. Chang MC. Fertilizing capacity of spermatozoa deposited into the fallopian tubes. *Nature* 1951;168:697-8.
24. Travis AJ, Kopf GS. The role of cholesterol efflux in regulating the fertilization potential of mammalian spermatozoa. *J Clin Invest* 2002;110:731-6.
25. Asano A, Nelson JL, Zhang S, Travis AJ. Characterization of the proteomes associating with three distinct membrane raft sub-types in murine sperm. *Proteomics* 2010;10:3494-505.
26. Asano A, Nelson-Harrington JL, Travis AJ. Phospholipase B is activated in response to sterol removal and stimulates acrosome exocytosis in murine sperm. *J Biol Chem* 2013;288:28104-15.
27. Asano A, Selvaraj V, Buttke DE, Nelson JL, Green KM, Evans JE *et al.* Biochemical characterization of membrane fractions in murine sperm: identification of three distinct sub-types of membrane rafts. *J Cell Physiol* 2009;218:537-48.
28. Buttke DE, Nelson JL, Schlegel PN, Hunnicutt GR, Travis AJ. Visualization of GM1 with cholera toxin B in live epididymal versus ejaculated bull, mouse, and human spermatozoa. *Biol Reprod* 2006;74:889-95.
29. Selvaraj V, Asano A, Buttke DE, McElwee JL, Nelson JL, Wolff CA *et al.* Segregation of micron-scale membrane sub-domains in live murine sperm. *J Cell Physiol* 2006;206:636-46.
30. Selvaraj V, Asano A, Buttke DE, Sengupta P, Weiss RS, Travis AJ. Mechanisms underlying the micron-scale segregation of sterols and GM1 in live mammalian sperm. *J Cell Physiol* 2009;218:522-36.
31. Travis AJ, Merdiushev T, Vargas LA, Jones BH, Purdon MA, Nipper RW *et al.* Expression and localization of caveolin-1, and the presence of membrane rafts, in mouse and Guinea pig spermatozoa. *Dev Biol* 2001;240:599-610.
32. Cohen R, Buttke DE, Asano A, Mukai C, Nelson JL, Ren D *et al.* Lipid modulation of calcium flux through CaV2.3 regulates acrosome exocytosis and fertilization. *Dev Cell* 2014;28:310-21.
33. Selvaraj V, Buttke DE, Asano A, McElwee JL, Wolff CA, Nelson JL *et al.* GM1 dynamics as a marker for membrane changes associated with the process of capacitation in murine and bovine spermatozoa. *J Androl* 2007;28:588-99.
34. Moody MA, Cardona C, Simpson AJ, Smith TT, Travis AJ, Ostermeier GC. Validation of a laboratory-developed test of human sperm capacitation. *Mol Reprod Dev* 2017;84:408-22.
35. Ostermeier GC, Cardona C, Moody MA, Simpson AJ, Mendoza R, Seaman E *et al.* Timing of sperm capacitation varies reproducibly among men. *Mol Reprod Dev* 2018;85:387-96.
36. Cardona C, Neri QV, Simpson AJ, Moody MA, Ostermeier GC, Seaman EK *et al.* Localization patterns of the ganglioside GM1 in human sperm are indicative of male fertility and independent of traditional semen measures. *Mol Reprod Dev* 2017;84:423-35.
37. Schinfeld J, Sharara F, Morris R, Palermo GD, Rosenwaks Z, Seaman E *et al.* Cap-Score prospectively predicts probability of pregnancy. *Mol Reprod Dev* 2018;85:654-64.
38. STROBE statement--checklist of items that should be included in reports of observational studies (STROBE initiative). *Int J Public Health* 2008;53:3-4.
39. World Health Organization. WHO laboratory manual for the examination and processing of human semen. 5th ed: Cambridge University Press, 2010.
40. Moody MA, Cardona C, Simpson AJ, Smith TT, Travis AJ, Ostermeier GC. Validation of a laboratory-developed test of human sperm capacitation. *Molecular Reproduction and Development* 2017;84:408-22.
41. Cardona C, Neri QV, Simpson AJ, Moody MA, Ostermeier GC, Seaman EK *et al.* Localization patterns of the ganglioside GM1 in human sperm are indicative of male fertility and independent of traditional semen measures. *Molecular Reproduction and Development* 2017;84:423-35.
42. Team RC. R: A language and environment for statistical computing. In. Vienna, Austria: R Foundation for statistical computing, 2013.
43. Jones E, Oliphant E, Peterson P, et-al. SciPy: Open Source Scientific Tools for Python. In, 2001.

44. Akaike H. A new look at the statistical model identification. *IEEE transactions on automatic control* 1974;19:716-23.
45. Wyndham N, Marin Figueira PG, Patrizio P. A persistent misperception: assisted reproductive technology can reverse the "aged biological clock". *Fertil Steril* 2012;97:1044-7.
46. Puga Molina LC, Luque GM, Balestrini PA, Marin-Briggiler CI, Romarowski A, Buffone MG. Molecular Basis of Human Sperm Capacitation. *Front Cell Dev Biol* 2018;6:72.
47. Babigumira JB, Sharara FI, Garrison LP, Jr. Projecting the potential impact of the Cap-Score on clinical pregnancy, live births, and medical costs in couples with unexplained infertility. *J Assist Reprod Genet* 2018;35:99-106.