Dear Editor,

Reproductive problems affect 8%–12% of human populations worldwide; 40%–50% of all infertility cases are due to the male, and up to 2% of men have suboptimal sperm quality.¹ Semen analysis is the first test in male infertility clinics, but interpretation of the results is not straightforward, with the significance of results for predicting fertility after intrauterine insemination (IUI), intracytoplasmic sperm injection (ICSI), and in vitro fertilization (IVF) differing between studies.²

Classical semen evaluation treats seminal variables individually, so that whereas poor semen quality is an indicator of subfertility,³ good semen quality is no guarantee of fertility.⁴

Considering that all the data and their distribution provide more information, computing variables by multivariate statistics offers a new understanding of the relationship of semen quality to fertility.¹ Different datasets corresponding to several variables were standardized from human semen samples and combined here to examine subpopulations of seminal variables with the eventual aim of providing a conceptual mathematical approach to fertility studies.

Thirteen volunteers (aged 25–59 years) signed informed consent forms to participate in the study and provide samples for the study. Semen samples were collected by masturbation after 3–5 days of abstinence. The mean of this age range was 36.9 years. Seminal volume and pH using strips (Macherey-Nagel, Düren, Germany) were assessed within 30 s of collection. The resting parameters were assessed after 30 min for liquefaction at 37°C, counting almost 250 cells per sample. Sperm motility and concentration (total sperm count) were determined in a reusable liquefaction at 37°C, counting almost 250 cells per sample. Sperm motility and negatively to TB. The third was positively related to sperm immobility and AB.

The first component was related positively to seminal pH, progressive sperm motility, AB, and TB and negatively to sperm immobility and AB. The second was related positively to total count and nonprogressive motility and negatively to TB. The third was positively related to vitality, nonprogressive motility, and TB. Although F, AB, and TB each indicated an aspect of DNA status, there was a greater correlation between AB and TB (r = 0.712) than that between F and AB (r = 0.242) or TB (r = 0.126).

Cluster analyses were performed on the data from the entire dataset. Principal component analysis (PCA) included the feasibility of factorial analysis, verified by the Bartlett’s test of sphericity, to confirm that the correlation matrix was an identity matrix, and the Kaiser–Meyer–Olkin index, which determines the correlations between two variables once the influence of other variables is eliminated. Only components with an eigenvalue >1 were used for the next, two-step cluster procedure with the sperm-derived indices obtained by the PCA. All the data were assessed in a nonhierarchical clustering procedure (k-means model and Euclidean distance), which classifies the spermatozoa into subpopulations according to joint characteristics and allows the detection of outliers. The effects of clusters within and between the semen measurements were analyzed by a generalized linear model. Statistical significance was considered as P < 0.05 with data analyzed by InfoStat Software (version 2017; InfoStat, Córdoba, Argentina) for Windows.

PCA rendered three components, explaining 77% of the variation. The first component was related positively to seminal pH, progressive sperm motility, AB, and TB and negatively to sperm immobility and F. The second was related positively to total count and nonprogressive motility and negatively to TB. The third was positively related to vitality, nonprogressive motility, and TB. Although F, AB, and TB each indicated an aspect of DNA status, there was a greater correlation between AB and TB (r = 0.712) than that between F and AB (r = 0.242) or TB (r = 0.126). Cluster analysis revealed two subpopulations, SP1, characterized by progressive motility and high levels of TB and AB,
and SP2 with total count and vitality, poor motility, and F (Table 1). Eight (61.5%) samples were included in SP1.

For decades, clinicians have sought a subjectively evaluated semen parameter that defines the fertility of a sample. This approach implies a great limitation in the value of the results, particularly for that parameters not completely standardized, as DNA fragmentation. Objectively, computer-assisted sperm analysis (CASA)-generated data were initially considered independently, although this approach has limited power.

Table 1: Subpopulations analysis of the seminogram data

| Variables                  | SP1 (n=8, 61.5%) | SP2 (n=5, 38.5%) |
|----------------------------|------------------|------------------|
|                            | MD    | Q1    | Q3    | MD    | Q1    | Q3    |
| Semen volume (ml)          | 3.15  | 2.50  | 5.30  | 4.00  | 3.80  | 4.50  |
| Semen pH                   | 7.65  | 7.30  | 8.00  | 7.50  | 7.30  | 7.50  |
| Total count (10⁶)          | 415   | 188   | 543   | 801   | 659   | 947   |
| Sperm motility (%)         | c     | 1.30  | 0.90  | 1.40  | 3.00  | 2.70  | 3.20  |
|                            | d     | 38.00 | 25.30 | 46.90 | 49.10 | 39.30 | 54.00 |
|                            | a + b | 60.65 | 46.20 | 69.30 | 48.40 | 42.20 | 57.5  |
| Sperm vitality (%)         | 60.55 | 49.36 | 64.85 | 64.19 | 62.27 | 77.78 |
| Sperm DNA fragmentation (%)| 8.30  | 4.75  | 11.20 | 14.10 | 12.82 | 16.61 |
| Sperm chromatin stability (%)| 85.58 | 82.30 | 87.20 | 73.95 | 64.50 | 80.68 |
| Sperm DNA stability (%)    | 77.63 | 52.81 | 81.02 | 62.50 | 30.00 | 69.00 |

*p < 0.05. SP1 and SP2 are represented with eigenvalues. a: fast and progressive; b: rapid nonprogressive and means; c: slow; MD: median; Q1: first quartile; Q3: third quartile

For this study on adult human semen samples is promising for improving the diagnostic potential of semen analysis. For developing a mathematical model integrating all semen variables, future studies must include more men, and the concept must be extended to subjects of variable age, fertility, residence, socioeconomic status, ethnicity, and with pharmacological and toxicological conditions. This will make possible definition of universal criteria that improve the prediction of fertility and optimize the diagnosis for assisted reproduction techniques.

AUTHOR CONTRIBUTIONS
AGM conceived and designed the study, and drafted and revised the manuscript. A Valverde conducted statistical design of the study and data analysis. DB, A Vendrell, and CC participated in data acquisition. CS participated in designing the study and critical review of the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS
All authors declared no competing interests.

ACKNOWLEDGMENTS
A Valverde was funded by the CONICIT and MICITT, Costa Rica. CC received funding from the EU Horizon 2020 research and innovation program under the Marie Skłodowska-Curie project IMPRESS (Grant No. 642893). Although AGM, DB, CC, MS, and CS work for Proiser R+D S.L., no commercial interest has had any influence on the content of this paper.

REFERENCES
1. Kumar N, Singh AK. Trends of male factor infertility, an important cause of infertility: a review of literature. J Hum Reprod Sci 2015; 8: 191–6.
2. Barratt CL, Björndahl L, Menkveld R, Mortimer D. ESHRE special interest group for andrology basic semen analysis course: a continued focus on accuracy, quality, efficiency and clinical relevance. Hum Reprod 2011; 26: 3207–12.
3. Bonde JP, Ernst E, Jensen TK, Hjollund NH, Kolstad H, et al. Relation between semen quality and fertility: a population-based study of 430 first-pregnancy planners. Lancet 1998; 352: 1172–7.
4. Colenbrander B, Gadella BM, Stout TA. The predictive value of semen analysis in the evaluation of stallion fertility. Reprod Dom Anim 2003; 38: 305–11.
5. Santolaria P, Soler C, Recreo P, Carretero T, Bono A, et al. Morphometric and kinematic sperm subpopulations in split ejaculates of normozoospermic men. Asian J Androl 2016; 18: 831–4.
6. World Health Organization. WHO Laboratory Manual for the Examination and Processing of Human Semen. Geneva: World Health Organization; 2010.
7. Yoshida M, Kawano N, Yoshida K. Control of sperm mobility and fertility: diverse factors and common mechanisms. Cell Mol Life Sci 2008; 65: 3446–57.
8. Santolaria P, Vicente-Fiel S, Palacin I, Fantova E, Blasco ME, et al. Predictive capacity of sperm quality parameters and sperm subpopulations on field fertility after artificial insemination in sheep. Anim Reprod Sci 2015; 163: 82–8.
9. Vásquez F, Soler C, Camps P, Valverde A, García-Molina A. Spermogram and sperm head morphometry assessed by multivariate cluster analysis results during adolescence (12-18 years) and the effect of varicocele. Asian J Androl 2016; 18: 824–30.
10. Yániz JL, Vicente-Fiel S, Soler C, Recreo P, Carretero T, et al. Comparison of different statistical approaches to evaluate morphometric sperm subpopulations in men. Asian J Androl 2016; 18: 819–23.

This is an open access journal, and articles are distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 License, which allows others to remix, tweak, and build upon the work non-commercially, as long as appropriate credit is given and the new creations are licensed under the identical terms.

©The Author(s)(2019)