Human Reproduction vol.12 no.12 pp.2710–2712, 1997

A simple monthly means chart system for monitoring sperm concentration

Gary N.Clarke

Andrology Unit, Division of Pathology, The Royal Women's Hospital, Carlton, Victoria, Australia, 3053

A simple, robust monthly means chart system for monitoring sperm concentration was developed and used in our andrology laboratory. The system relied on the central limit theorem to allow estimation of chart limits without resorting to data transformation procedures. Once the chart was drawn up, the arithmetic mean of sperm concentration was determined monthly and plotted. Monthly means which fell outside 2 SE from the baseline mean were investigated. Initial experience with this chart system suggests that it provides a simple and useful quality assurance method which should be applicable to other semen variables such as motility and sperm morphology.

Key words: quality assurance/quality control/semen analysis/sperm concentration/sperm count

Introduction

Quality control in the andrology laboratory, as in other clinical laboratories, should involve a multi-faceted approach including both internal and external quality assurance programmes (QAP; Matson, 1995). Although several early studies (Freund and Carol, 1964; Jequier and Ukombe, 1983) identified large intra-and inter-technician variations in semen analysis results, implementation of rigorous QAP has been a slow process. Dr P.Matson has recently (1995–96) co-ordinated an external QAP pilot study in Australia under the auspices of the Fertility Society of Australia (FSA). The FSA intends to establish a full QAP in 1997. However, it is essential that laboratories participating in an external QAP also implement rigorous and ongoing internal QAP. An important component of internal quality assurance involves the maintenance of charts portraying mean values for laboratory results calculated at regular intervals (Knuth *et al.*, 1989; Cooper *et al.*, 1992). Such charts provide the laboratory manager with a potential means of detecting consistent changes or drift in laboratory results, but they are unable to detect pre-existing systematic bias.

The aim of this study was to develop a simple, robust system for charting monthly means for sperm counts, which could be readily implemented in most laboratories. The chart system developed during this study relies on the central limit theorem, thus obviating the need for cumbersome transformation procedures and complex statistical manipulations.

Materials and methods

Semen analysis

Semen analyses were performed manually as described previously (Johnston *et al.*, 1995).

Monthly means charts

The development of the charts was based on the central limit theorem of statistics (Sokal and Rohlf, 1981). The main tenet of the theorem is that as sample size (n) increases, the means of samples taken from a population of *any* distribution will approach the normal distribution. Thus, in the case of sperm counts, which are strongly skewed to the right, the distribution of monthly means will approximate a normal distribution, assuming that the monthly sample is of reasonable size. This laboratory averages 160 semen analyses per month; the procedure, however, would be applicable to any laboratory performing at least 80 analyses per month. Smaller laboratories may need to take bi-monthly means.

In order to determine the limits for the charts, it was necessary to calculate the mean and SD for each variable from a previous year's tests (baseline). The SEM for the monthly sample was then estimated by dividing the SD of the baseline data by the square root of the average number of tests per month (assuming a stable workload from year to year):

$$\text{SEM} = \text{SD}/\sqrt{n}$$

The appropriate chart was then drawn up with control lines indicating the position of the mean and limits defined as ± 2 or 3 SEM. For each monthly point, the arithmetic mean sperm count was plotted on the chart. Monthly means falling outside the chosen outer limits were investigated. In addition, the chart was examined for consistent trends or runs. For example, six consecutive points on the same side of the mean would suggest that the population mean had changed ($P < 0.02$); similarly, two consecutive points outside 2 SEM on the same side of the mean would be highly improbable ($P < 0.001$) if the test procedure was still in control. In monitoring these charts, it is important to recall that the limits are based on the average monthly sample size from a previous year's data – a much smaller actual sample size for a particular current month would mean that the real limits for that month are wider than those drawn on the chart. The actual limits for that month can then be accurately calculated ($\text{SD}/\sqrt{n}$) and drawn in. However, for most purposes the limits calculated using the average monthly sample sizes are quite adequate. Those laboratories with large fluctuations in workload may need to draw in stepped confidence limits on their charts.

Results

Sperm counts ($n = 1915$) from 1994 were used as the baseline data for drawing up the monthly means charts. The heavily skewed distribution of the original data is evident from Figure 1. The mean sperm count for the 1994 analyses was 79.18×10^6/ml, with an SD of 86.72×10^6/ml. Thus, the SEM

© European Society for Human Reproduction and Embryology

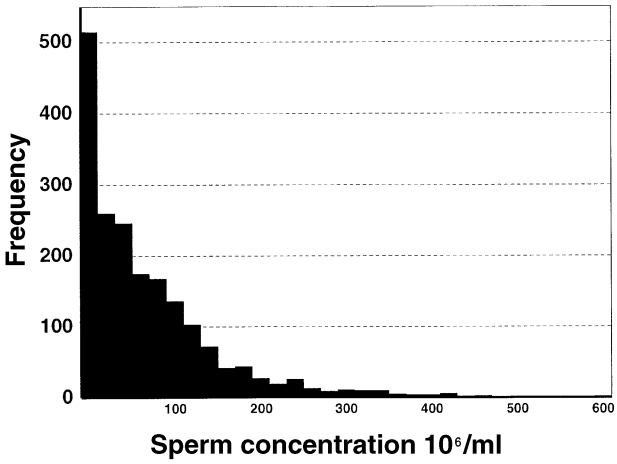

Figure 1. Frequency distribution of baseline sperm concentration results for 1994.

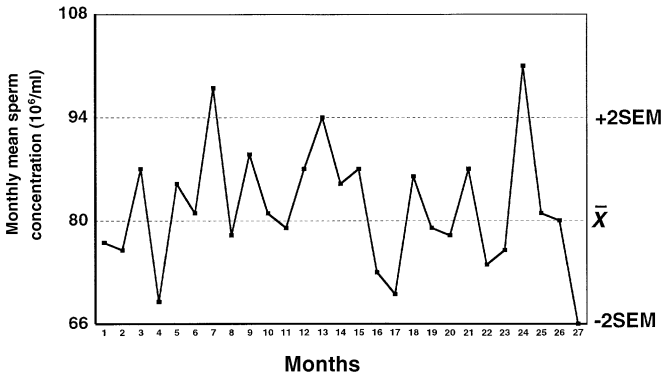

Figure 2. Monthly means control chart for sperm concentration.

for monthly samples of 160 (i.e. 1915/12) will be 6.9 (round to 7). The monthly means chart was then constructed using ±2 or 3 SEM around the mean (Figure 2). The arithmetic monthly means for sperm counts were then plotted on the chart for 1995 to April 1996 data. Two high monthly means were identified (July 1995, December 1996) and investigated. The July 1995 mean included three very high counts (722, 512, 528), whilst the December 1996 mean included two high counts (970, 670). In both months, these few very high counts were sufficient to increase the monthly mean by ~10, thus explaining the unusually high means. Further investigation indicated that these counts had been calculated and recorded correctly. There was therefore no evidence that these high means were due to any errors or changes in laboratory procedures.

The utility of the central limit theorem for monthly means of sperm counts was evident from the relatively normal distribution of the monthly means (grand mean = 81.6, median = 81, skewness = 0.4) versus the original individual counts for 1994 (mean = 79.2, median = 55, skewness = 2.6).

Discussion

The results of this investigation suggest that simple charts of the arithmetic means of semen variables can provide a useful method for the long term monitoring of laboratory output,

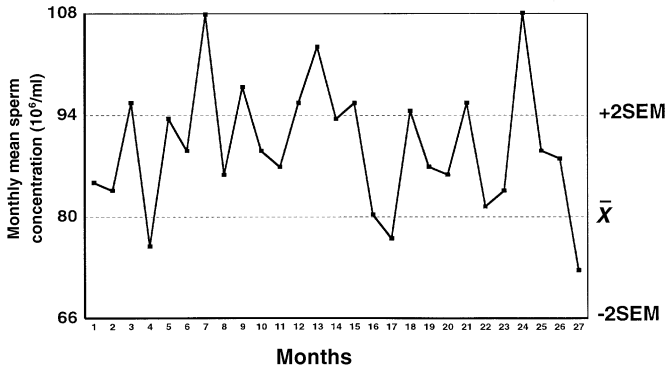

Figure 3. Monthly means from Figure 2 replotted with 10% bias.

with the potential to detect consistent errors or drift in mean values. Knuth *et al.* (1989) suggested the use of monthly means charts for semen variables; however, their system required relatively complex statistical procedures to detect significant differences in monthly means. In contrast, the system developed during the present study requires mainly visual appraisal coupled with simple probability calculations and further investigations if indicated. Thus, the present system would be more amenable to routine use in many laboratories. If the mean or SD of the patient population alters due to an identifiable change in the patient mix, or to changes in laboratory procedure, then it may be necessary to analyse a new set of baseline data to determine the new mean and chart limits.

In order to demonstrate the sensitivity of the monthly means QAP, Figure 3 illustrates the effect on the chart of a 10% increment in the original monthly means shown in Figure 2. It is apparent that a consistent 10% shift in sperm concentration would be readily detectable by the laboratory manager, without necessarily performing any further statistical analysis. Errors of this magnitude can result from counting chambers having incorrect dimensions, or dilution errors due to poorly serviced pipettes. The practical benefits of monthly means monitoring for sperm motility and morphology have been demonstrated previously by Knuth *et al.* (1989) and Cooper *et al.* (1992). However, in order to allow a more rapid response to changes in laboratory output, larger laboratories may wish to plot weekly or fortnightly means.

Other important components of internal QAP include the training of new technicians, regular 'blind' comparisons between trained technicians using pooled semen, and regular calibration of pipettes, counting chambers and weighing balances. However, it is important that the laboratory manager has realistic expectations concerning the accuracy and precision of the procedures performed in the andrology laboratory (see Barratt, 1995). For example, because cells in a haemocytometer approximate a Poisson distribution (Barnett, 1979; Sokal and Rohlf, 1981), the minimum SD of a sperm count (using one side of the haemocytometer) will be equal to the square-root of the number of sperm counted (e.g. n = 200, SD = 14, confidence limits of ±28). Thus, the 95% confidence limits for a sample with a mean concentration of 100×10^6/ml, would range from 86 to 114×10^6/ml. Greater precision may be

obtained by counting more squares in the chamber and/or by counting both sides of the haemocytometer.

In conclusion, the monthly means chart described in this paper provides a simple, robust quality control procedure for monitoring laboratory output for skewed semen variables such as sperm concentration. An optional refinement to the present procedure could involve the substitution of the 5% trimmed mean, which would minimize the potential bias created by a few very high counts. Initial investigation suggests that the same monthly means procedure is applicable to sperm motility and morphology.

Acknowledgements

The authors thank andrology scientists, Mr Peter Elliott, Mr Henry Oh, and Mrs Jing An, and also Kay Kerrison for typing the manuscript.

References

Barratt, C.L.R. (1995) On the accuracy and clinical value of semen laboratory tests. *Hum. Reprod.*, **10,** 250–252.

Barnett, R.N. (1979) *Clinical Laboratory Statistics*. Little, Brown and Company, Boston, pp.31–33.

Cooper, T.G., Neuwinger, J., Bahrs, S. and Nieschlag, E. (1992) Internal quality control of semen analysis. *Fertil. Steril.*, **58,** 172–178.

Freund, M. and Carol, B. (1964) Factors affecting haemocytometer counts of sperm concentration in human semen. *J. Reprod. Fertil.*, **8,** 149–155.

Jequier, A.M. and Ukombe, E.B. (1983) Errors inherent in the performance of a routine semen analysis. *Br. J. Urol.*, **55,** 434–436.

Johnston, R.C., Clarke, G.N., Liu, D.Y. and Baker, H.W.G. (1995) Assessment of the sperm quality analyzer. *Fertil. Steril.*, **63,** 1071–1076.

Knuth, U.A., Neuwinger, J. and Nieschlag, E. (1989) 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.*, **12,** 375–383.

Matson, P.L. (1995) External quality assessment for semen analysis and sperm antibody detection: results of a pilot scheme. *Hum. Reprod.*, **10,** 620–625.

Sokal, R.R. and Rohlf, E.J. (1981) *Biometry.* Freeman and Company, San Francisco.

Received on July 8, 1997; accepted on September 16, 1997