Human Reproduction vol.15 no.8 pp.1667–1671, 2000

Fourier transform infrared spectroscopy of follicular fluids from large and small antral follicles

Non Thomas[1], Royston Goodacre[2], Éadaoin M.Timmins[2], Marco Gaudoin[1], and Richard Fleming[1,3]

[1,3]Assisted Conception Service, Glasgow Royal Infirmary University NHS Trust, Glasgow, G31 2ER and [2]Institute of Biological Sciences, Cledwyn Building, University of Wales, Aberystwyth, Ceredigion, Wales, SY23 3DD, UK

[3]To whom correspondence should be addressed at: Assisted Conception Service, Glasgow Royal Infirmary University NHS Trust, Glasgow G31 2ER, UK. E-mail: gqta13@udcf.gla.ac.uk

Fourier transform infrared spectroscopy (FTIR) was used to obtain 'biochemical fingerprints' for the constitution of follicular fluids from large and small antral luteinized follicles (n = 54 pairs). All samples gave reproducible characteristic biological infrared absorption spectra, with recognizable amide I protein vibrations and acyl vibrations from fatty acids. Discriminant function analysis of the first derivative FTIR spectra, together with hierarchical cluster analysis used to construct a dendrogram, showed fluid from large follicles formed a homogeneous closely related cluster, whilst that from small follicles was distinct from the large, and heterogeneous in nature. The large follicle fluids showed closer biochemical similarity to each other than to the corresponding fluid taken from small matched follicles. An artificial neural network was trained and following validation with an independent test set, successfully distinguished follicular fluids from large and small follicles. The sex steroid concentrations in the fluids from large and small follicles were significantly different. These results show that fluid from large follicles is distinct in biochemical nature from that from small follicles, but the degree of homogeneity implies size-specific changes take place. These may have consequences for the developmental potential of the oocyte.

Key words: artificial neural networks/follicle fluid/follicle size/fourier transform infrared spectroscopy/steroid analysis

Introduction

In the normal menstrual cycle, follicular diameter (FD), oestradiol and inhibin are the main indicators of the maturational stage of the follicle and enclosed oocyte. FD has become the marker of choice in the treatment of assisted conception patients, to whom human chorionic gonadotrophin (HCG) is administered when a number of follicles achieve the desired size of between 17–20mm in diameter. During retrieval, oocytes are removed from follicles of a wide range of sizes but it is often presumed that follicles >16mm in diameter yield oocytes with a higher developmental capacity than those from smaller follicles. Previous studies have shown that larger follicles yield oocytes with increased developmental, fertilization and embryo developmental rates than those yielded by smaller follicles in IVF programmes (Ectors *et al*., 1997; Bergh *et al*., 1998). Conflicting results were achieved in the same studies using intracytoplasmic sperm injection (ICSI), suggesting that any developmental deficiency in small oocytes may (Bergh *et al*., 1998) or may not (Ectors *et al*., 1997) be overcome using sperm injection. However, despite the recorded differences, there are still notable rates of fertilization and pregnancy achieved with oocytes taken from smaller follicles. Indeed, although previous studies have shown fertilization rates to be lower in oocytes taken from small follicles (<16mm diameter or <1ml volume) compared with large (>16mm diameter or >1ml volume) no difference in embryo quality (Salha *et al.*, 1998) or pregnancy rate could be demonstrated between the two groups (Nagai *et al.*, 1997). These clinical studies indicate that there may be biochemical differences related to these later stages of follicular development.

As an ovarian follicle grows, follicular fluid, a mixture of follicular secretions and plasma exudate, accumulates between the granulosa cells. The importance of this fluid cannot be underestimated since it is closely associated with the oocyte and its composition may influence the latter part of oocyte development. Many follicular fluid factors will have been secreted by cells in close association with the oocyte and the fluid contains factors responsible for ensuring the correct maturational status of the oocyte. Constituents of the fluid include steroids, lipids, glycosaminoglycans and numerous proteins and peptides.

It has long been a goal to identify factors in follicles which may indicate improved chances of embryo development and conception. The aim of this study was to establish if a definitive difference exists between follicular fluid derived from different sized follicles. Paired samples were subjected to the analytical technique of Fourier transform infrared (FTIR) spectroscopy.

FTIR is a physico-chemical method which measures the vibrations of bonds within functional groups (Griffiths and de Haseth, 1986; Stuart, 1997). In FTIR analysis a particular bond absorbs light electromagnetic (EM) radiation at a specific wavelength; for example, the infrared (IR) spectra of proteins exhibit strong amide I absorption bands at 1653 cm^{-1} associated with the characteristic stretching of C=O and C-N and the bending of the N-H bonds (Stuart, 1997). Therefore by interrogating a biological sample with EM radiation of many wavelengths in the mid-IR range (usually defined as 4000–600 cm^{-1}) one can construct an IR 'fingerprint' of the original biological sample under investigation. Since different bonds

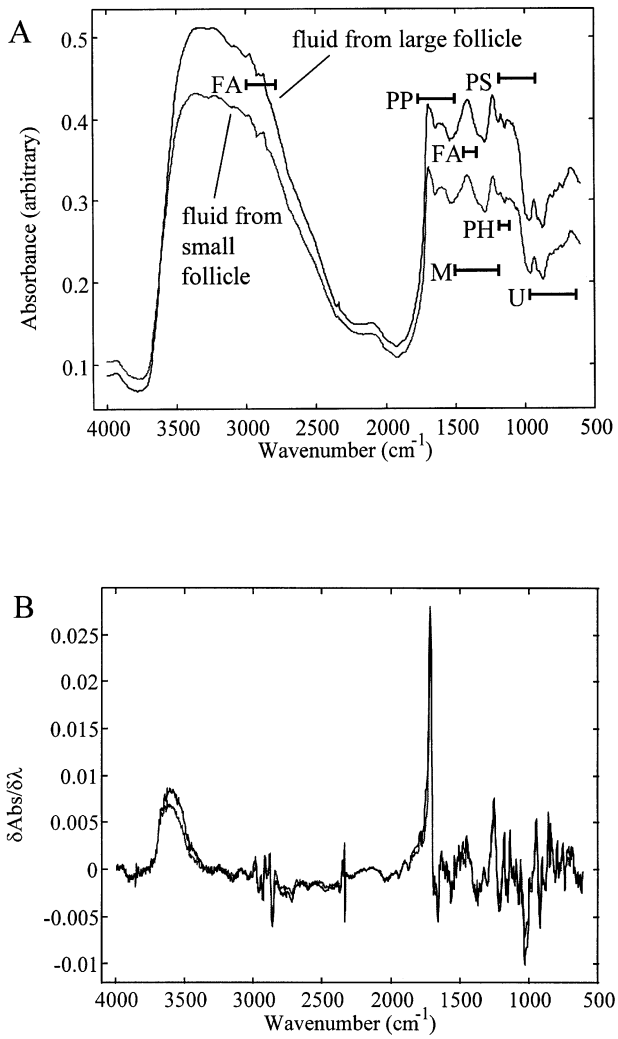

Figure 1. FTIR spectra of follicular fluids from large and small antral follicles from patient 1. Spectra shown are either raw (**A**) or the smoothed first derivatives (**B**). In the derivative spectrum (**B**) the original spectrum (**A**) is differentiated. The resolution is enhanced in the first derivative because it is a measure of the change in gradient, therefore the *y* axis is a measure of the difference in relative absorbance divided by the difference in wavenumber.FA = fatty acids; PH = phosphates; PP = peptides; PS = polysaccharides; U = unassigned; M = mixed region of proteins, fatty acids, phosphate rich.

absorb different wavelengths of EM radiation, these 'fingerprints' are made up of the vibrational features of all chemical components in the sample analysed. Thus this method gives quantitative information about the total biochemical composition of a follicular fluid, without its destruction, and produces 'fingerprints' which are reproducible and distinct for different biological materials. Within medicine there are precedents for using IR measurement techniques for the analysis of human biofluids such as urine, blood and synovial fluid (Wang *et al.*, 1996; Jackson *et al.*, 1997; Diem *et al.*, 1999).

Materials and methods

Patients and treatments

Follicular fluid was collected at oocyte retrieval from patients undergoing routine IVF and ICSI at the Glasgow Royal Infirmary. Paired

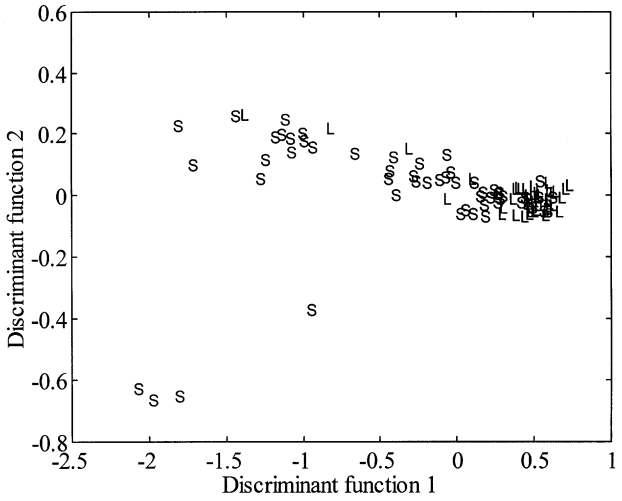

Figure 2. Discriminant function analysis on 54 pairs of follicular fluids. Data points plotted are the discriminant function score averages and the identifiers are either S (small antral follicles) or L (large).

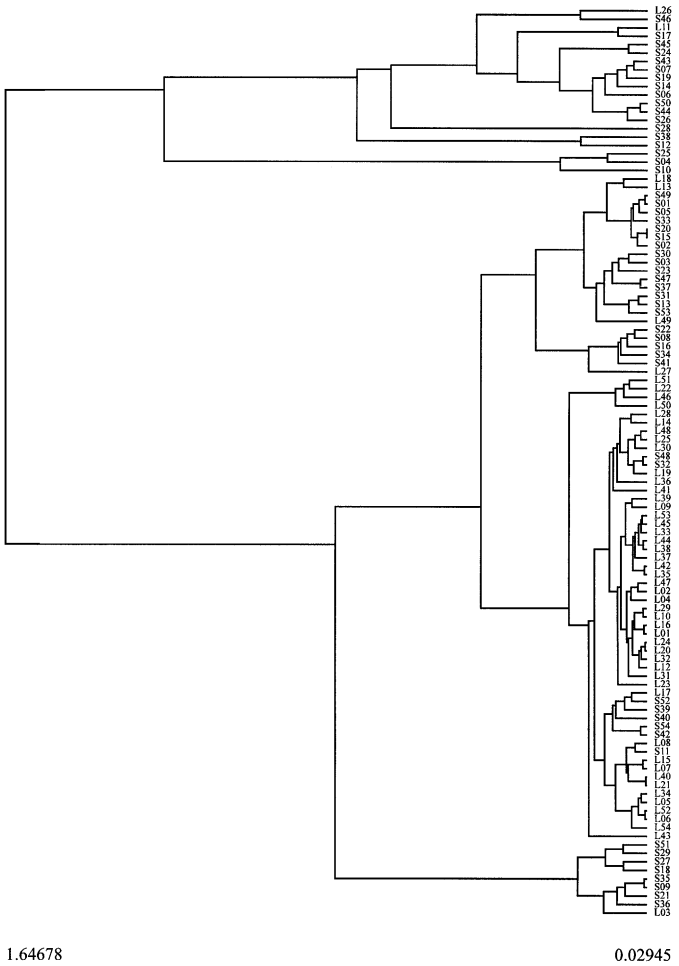

Figure 3. Hierarchical cluster analysis on 54 pairs of follicular fluids.

follicular fluid samples were obtained from 54 patients (mean age 33, range 26–42 years). Diagnoses were as follows:polycystic ovarian syndrome (PCOS) (n = 1); endometriosis (n = 6); endometriosis plus PCOS (n = 1); male factor (n = 9); tubal (n = 25); tubal plus endometriosis (n = 3); tubal plus male (n = 1) and unexplained

Table I. Identity of the size of the small and large antral follicles from the FTIR spectra of follicular fluids in the training and test sets as judged by artificial neural networks (ANN).

Training set				Test set			
Identifier	Prediction	Identifier	Prediction	Identifier	Prediction	Identifier	Prediction
S01	–0.04	L01	1.02	S28	–0.04	L28	0.98
S02	0.01	L02	1.05	S29	–0.07	L29	0.93
S03	0.00	L03	1.01	S30	0.23	L30	0.81
S04	–0.04	L04	1.03	S31	0.03	L31	1.02
S05	0.01	L05	0.99	S32	0.33	L32	1.06
S06	–0.03	L06	1.02	S33	0.13	L33	0.93
S07	–0.02	L07	1.00	S34	0.41	L34	1.07
S08	–0.02	L08	1.03	S35	0.19	L35	1.05
S09	0.00	L09	0.99	S36	0.28	L36	1.05
S10	–0.05	L10	1.04	S37	0.16	L37	1.06
S11	0.03	L11	0.98	S38	0.13	L38	0.95
S12	–0.02	L12	1.04	S39	0.26	L39	0.99
S13	–0.03	L13	1.03	S40	0.10	L40	1.04
S14	–0.03	L14	1.06	S41	0.16	L41	0.86
S15	–0.03	L15	1.06	S42	0.46	L42	1.06
S16	–0.02	L16	1.06	S43	0.11	L43	1.05
S17	–0.02	L17	0.96	S44	0.14	L44	1.06
S18	–0.02	L18	1.04	S45	–0.02	L45	1.01
S19	–0.01	L19	1.01	S46	0.34	L46	1.09
S20	–0.04	L20	1.04	S47	0.08	L47	1.06
S21	–0.02	L21	0.98	S48	0.28	L48	1.04
S22	–0.04	L22	1.04	S49	0.14	L49	0.96
S23	0.00	L23	1.00	S50	0.32	L50	1.05
S24	–0.05	L24	1.04	S51	0.10	L51	1.00
S25	–0.02	L25	1.07	S52	0.08	L52	1.05
S26	–0.03	L26	0.93	S53	0.28	L53	1.10
S27	–0.02	L27	1.03	S54	0.41	L54	0.99

The predictions given are the average of five separate ANN. The values shown are the averages for the three replicate samples from the same follicular fluid.

(n = 8). All patients were initially given a long course of GnRH analogue (buserelin) as a method of down-regulation. On day 3 of the cycle, gonadotrophin stimulation was commenced and the follicular phase was monitored by transvaginal ultrasonography and serum oestradiol measurements. When the presence of three or more follicles, >17 mm diameter was detected, the patients received HCG administration and the follicular fluids and oocytes were collected ~36 h later.

Local ethical committee approval was granted for this project by the Glasgow Royal Infirmary University NHS Trust.

Follicular fluid collection

Two follicular fluid samples were collected from each patient. One from a follicle with a diameter >17mm (large) and one from a follicle <15mm (small). Follicle size was determined by using the mean of two perpendicular measurements. In order to minimize contamination of fluids, samples were collected following a needle flush so that there would be no dilution in the catheter with fluids from follicles not included in the study. In practice the small follicle was the first to be aspirated at retrieval. The large follicle used in the study was the first approach to the contralateral ovary. Only mid-stream samples, free from blood contamination, were collected and aliquots were stored at –20°C. All the samples used in this study were paired.

Fourier transformed IR spectroscopy

Aliquots (20µl) of the 108 follicular fluids were evenly applied onto a sand-blasted aluminium plate. Prior to analysis the samples were oven-dried at 50°C for 30 min. Samples were run in triplicate. The FTIR instrument used was the Bruker IFS28 FTIR spectrometer (Bruker Spectrospin Ltd., Banner Lane, Coventry, UK) equipped with an MCT (mercury-cadmium-telluride) detector cooled with liquid N_2. The aluminium plate was then loaded onto the motorized stage of an adapted reflectance thin-layer chromatography accessory (Timmins *et al*., 1998). The personal computer used to control the IFS28 was also programmed (using OPUS version 2.1 software running under IBM O/S2 Warp provided by the manufacturers) to collect spectra over the wavenumber range 4000 cm^{-1} to 600 cm^{-1}. Spectra were acquired at a rate of 20 s^{-1}. The spectral resolution used was 4 cm^{-1}. To improve the signal-to-noise ratio, 256 spectra were co-added and averaged. Each sample was thus represented by a spectrum containing 882 points and spectra were displayed in terms of absorbance as calculated from the reflectance-absorbance spectra using the Opus software [which is based on the Kubelka-Munk theory (Griffiths and deHaseth, 1986)].

To minimize problems arising from baseline shifts the following procedure was implemented: (i) the spectra were first normalized so that the smallest absorbance was set to 0 and the highest to +1 for each spectrum, (ii) next the first derivatives of the original FTIR spectra were smoothed using the Savitzky-Golay algorithm (Savitzky and Golay, 1964) using 5-point smoothing. Typical spectra, pre- and post-processing, are shown in Figure 1.

Cluster analysis

The initial stage involved the reduction of the multidimensional FTIR data by principal components analysis (PCA; Jolliffe, 1986). PCA is a well known technique for reducing the dimensionality of multivariate data whilst preserving most of the variance; PCA was performed

according to the NIPALS algorithm (Wold, 1966). Discriminant function analysis [DFA; also known as canonical variates analysis (CVA)] then discriminated between groups on the basis of the retained principal components and the knowledge *a priori* of which spectra were replicates, and thus this process did not bias the analysis in any way (Manly, 1994). Finally, the Euclidean distance between *a priori* group centres in DFA space was used to construct a similarity measure, with the Gower general similarity coefficient S_G (Gower, 1966), and these distance measures were then processed by an agglomerative clustering algorithm to construct a dendrogram (Manly, 1994). These cluster analysis methods were implemented using Matlab version 5 (The Math Works, Inc., Natick, MA, USA), which runs under Microsoft Windows NT on an IBM-compatible PC.

Artificial neural network analysis

When the desired responses (targets) associated with each of the inputs (spectra) are known then the system may be 'supervized'. The goal of supervized learning is to find a model that will correctly associate the inputs with the targets; this is usually achieved by minimizing the error between the target and the model's response (output) (Massart *et al.*, 1997).

The artificial neural network (ANN) method of standard back-propagation multi-layer perceptrons (MLP) (Rumelhart *et al.*, 1986; Bishop, 1995) was used, and was carried out with a neural network simulation program, NeuFrame version 3,0,0,0 (Neural Computer Sciences, Southampton, Hants, UK), which runs under Microsoft Windows NT on an IBM-compatible PC. In-depth descriptions of the *modus operandi* of this type of ANN analysis are given elsewhere (Goodacre *et al.*, 1998). The architecture of the ANNs was 878 input nodes, 10 nodes in the hidden layer, and a single output node (this topology can be represented as 878–10–1). The ANNs were trained with the first 27 samples from the fluids from both small and large follicles [therefore there were 162 spectra (27*3*2) in the training set]. These ANN were trained for 1.5×10^3 epochs (calculations), when the RMS (root mean squared) error between the observed and desired outputs was typically 0.046 ± 0.004; on a Pentium 133, with 128 MB RAM, this typically took ~90 min. The ANN was then interrogated with both the training and test sets and a correct identity for a fluid from a small antral follicle was taken as <0.5 and for a fluid from a large follicle as $\geqslant 0.5$; note that after training, the interrogation of these ANN takes only a few milliseconds.

Steroid hormone assays

Oestradiol and progesterone were analysed using a solid-phase chemiluminescent enzyme immunoassay (DPC, Los Angeles, CA, USA) on an Immulite automated immunoassay system. Matching paired samples were assayed in two batches with intra-assay coefficients of variations of 4.5% (oestradiol) and 6% (progesterone).

Results

Fourier transformed IR spectroscopy

All samples gave reproducible characteristic biological IR absorption spectra with recognizable amide I protein vibrations and acyl vibrations from fatty acids (see Figure 1A). However, these FTIR spectra (and all the others collected) showed broad and complex contours, highlighting the fact that there was very little qualitative difference between the spectra, although at least some complex quantitative differences between them were observed. Such spectra, essentially uninterpretable by the naked eye, readily illustrated the need to employ multivariate statistical techniques for their analyses.

Table II. Steroid concentrations of follicular fluids

	Small	Large	*P*-value
Progesterone (nmol/l)	4449 ± 2999 (67)	7796 ± 3272 (42)	< 0.001
Oestradiol (pmol/l)	169 ± 124 (73)	261 ± 137 (52)	< 0.007

Values in parentheses are SD as percentage of mean.

DFA were used to observe the relationships between the follicular fluids as judged from their derived FTIR spectra, and DFA was performed as detailed above. The resulting DFA plot is shown in Figure 2 where it can be seen that the follicular fluids from large follicles formed a homogeneous, closely related cluster whilst those fluids from small follicles were very heterogeneous and distinct from the fluids from large antral follicles.

An alternative method of viewing the relationship between these follicular fluids is by using hierarchical cluster analyses and the resulting dendrogram is shown in Figure 3. This plot also showed clearly that the fluids from large follicles were more closely related to each other than to their corresponding paired follicular fluids from small antral follicles.

To distinguish between spectra from fluids from large and small antral follicles, the supervized learning method of ANNs was implemented. ANN were trained on the first 27 fluids from both the small and large antral follicles, with the output node encoded as either '0′ for fluids from small follicles or '1' for fluids from large follicles. Once calibrated, the ANN was then interrogated with derivative spectra from both the training and test sets, and its predictions are shown in Table I. It can be seen that both the training set and independent test set were correctly identified for whether the follicular fluids had been taken from large or small antral follicles. It is therefore evident that there was reproducible (bio)chemical information in the FTIR spectra of these follicular fluids that could distinguish unequivocally between large and small antral follicles.

Steroid analysis

The sex steroid concentrations of the follicular fluids also showed that fluid from large follicles was significantly different from small (Table II). Furthermore, the standard deviations of the steroid hormone concentration values were significantly greater in the small follicles (progesterone, 67%; oestradiol, 73%) than in the large follicles (42% and 52% respectively), again demonstrating greater heterogeneity in the smaller follicles.

Discussion

DFA showed the FTIR spectra of fluids collected from large antral follicles to be tightly clustered, indicating a similar biochemical profile, whereas fluids from small follicles were heterogeneous. This heterogeneity in the FTIR 'biochemical fingerprint' of fluids from small follicles may reflect differences in the maturational stage of the antral follicles. Dendrogram analysis of the same fluids also showed similarities in the biochemistry of the fluids from large follicles which were

more closely related to each other than to fluids from their corresponding small antral follicles.

FTIR spectroscopy coupled with ANN analysis distinguished successfully follicular fluid from large and small antral ovarian follicles indicating that biochemical differences do indeed exist between the fluids in different size follicles.

The differences in distribution of steroid concentrations do not reflect those of the FTIR, suggesting that other aspects of biochemical analysis are not under the same control mechanism as the steroids. These differences between factors in the fluids from large and follicles, which may be quantitative or qualitative, may be related to differences in oocyte quality. These results highlight a difference in the biochemical nature of fluids from large and small follicles which we suggest reflect the developmental stage of the follicle. Further work will investigate this premise and relate the FTIR spectra of a follicular fluid to the developmental capacity of the oocyte that the follicle contained.

In conclusion, this is the first time that FTIR spectroscopy has been used to analyse follicular fluids and it has shown that gross biochemical differences do exist between the follicular fluids from large and small antral ovarian follicles.

Acknowledgements

R.G. and E.M.T. are indebted to the Wellcome Trust for financial support (grant number 042615/Z/94/Z).

References

Bergh, C., Broden, H., Lundin, K. *et al.* (1998) Comparison of fertilization, cleavage and pregnancy rates of oocytes from large and small follicles. *Hum. Reprod.,* **13,** 1912–1915.

Bishop, C.M. (1995) *Neural networks for pattern recognition*. Clarendon Press, Oxford.

Diem, M., Boydston White, S. and Chiriboga, L. (1999) Infrared spectroscopy of cells and tissues: Shining light onto a novel subject. *Appl. Spectrosc*. **53,** A148-A161.

Ectors, E., Vanderzwalmen, P., Van Hoeck, J. *et al*. (1997) Relationship of human follicular diameter with oocyte fertilization and development after *in vitro* fertilization or intracytoplasmic sperm injection. *Hum. Reprod.,* **12,** 2002–2005.

Goodacre, R., Timmins, É.M., Burton, R. *et al.* (1998) Rapid identification of urinary tract infection bacteria using hyperspectral, whole organism fingerprinting and artificial neural networks. *Microbiology,* **144,** 1157–1170.

Gower, J.C. (1966) Some Distance Properties of Latent Root and Vector Methods used in Multivariate Analysis. *Biometrika,* **53,** 325–338.

Griffiths, P.R. and de Haseth, J.A. (1986) Fourier transformed infrared spectrometry. John Wiley, New York.

Jackson, M., Sowa, M.G. and Mantsch, H.H. (1997) Infrared spectroscopy: a new frontier in medicine. *Biophysical Chemistry,* **68,** 109–125.

Jolliffe, I.T. (1986) *Principal Component Analysis*. Springer-Verlag: New York.

Manly, B.F.J. (ed.) (1994) *Multivariate Statistical Methods: A Primer.* Chapman & Hall, London.

Massart, D.L., Vandeginste, B.G.M., Buydens, L.M.C. *et al.* (1997) *Handbook of Chemometrics and Qualimetrics: Part A*. Elsevier, Amsterdam.

Nagai, S., Yasumizu, T., Kasai, T. *et al.* (1997) Effect of oocyte retrieval from a small leading follicle in fixed-schedule *in vitro* fertilization program. *J. Obs.Gynaecol. Res.,* **23,** 165–169.

Rumelhart, D.E., McClelland, J.L. and The PDP Research Group (1986) *Parallel Distributed Processing, Experiments in the Microstructure of Cognition*, Vol. I and II, MIT Press, Cambridge, MA.

Salha, O., Nugent, D., Dada, T. *et al.* (1998) The relationship between follicular fluid aspirate volume and oocyte maturity in *in vitro* fertilization cycles. *Hum. Reprod.,* **13,** 1901–1906.

Savitzky, A. and Golay, M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem*., **36,** 1627–1623.

Stuart, B. (1997) *Biological Applications of Infrared Spectrocopy.* John Wiley & Sons, Chichester.

Timmins, É. M., Howell, S. A., Alsberg, B. K. *et al.* (1998) Rapid differentiation of closely related candida species and strains by pyrolysis-mass spectrometry and Fourier transform-infrared spectroscopy. *J. Clin. Microbiol.,* **36**,367–374.

Wang, J., Sowa, M., Mantsch, H.H. *et al*. (1996) Comparison of different IR measurement techniques in the clinical analysis of biofluids. *TRAC – Trends in Analytical Chemistry* **15,** 286–296.

Wold, H. (1966) Estimation of principal components and related models by iterative least squares. In Krishnaiah, K.R. (Ed) *Multivariate Analysis*, Academic Press, New York pp. 391–420.

Received November 5, 1999; accepted April 25, 2000