Noninvasive metabolomic profiling of human embryo culture media using Raman spectroscopy predicts embryonic reproductive potential: a prospective blinded pilot study

Richard Scott, M.D., a Emre Seli, M.D., b Kathy Miller, B.S., a Denny Sakkas, Ph.D., b Katherine Scott, B.S., a and David H. Burns, Ph.D. c

a Reproductive Medicine Associates of New Jersey, Morristown, New Jersey; Department of Obstetrics, Gynecology, and Reproductive Sciences, UMDNJ–Robert Wood Johnson Medical School, New Brunswick, New Jersey; b Department of Obstetrics, Gynecology, and Reproductive Sciences, Yale University School of Medicine, New Haven, Connecticut; and c Department of Chemistry, McGill University, Montreal, Quebec, Canada

Objective: To determine if metabolomic profiling of embryonic development was associated with implantation rates in IVF.

Design: Prospective blinded.

Setting: University-affiliated assisted reproductive technology program.

Patient(s): Unselected assisted reproductive technology population.

Intervention(s): None.

Main Outcome Measure(s): Raman-based biospectroscopic metabolomic profiling of spent culture media and delivery rates.

Results: Forty-one spent media samples from 19 patients with known reproductive potential (0 or 100% delivery rates of each embryo that implanted) were evaluated. Raman-based metabolomic profiling was used to calculate a viability index for each sample. On day 3, the spent media of embryos with proven reproductive potential (n = 33) demonstrated higher viability indices (0.875 ± 0.12) than those that failed to implant (0.56 ± 0.09). Similar findings were present in spent media from embryos transferred on day 5 (n = 8) (0.40 ± 0.21 vs. −0.81 ± −0.08). Receiver operating characteristic curve analyses were used to select thresholds with the greatest ability to discriminate outcomes. Overall diagnostic accuracy for predicting delivery or a failed implantation was 80.5%.

Conclusions: There is a clear relationship between the reproductive potential of human embryos and their modification of their culture media as detected by Raman biospectroscopy-based metabolomic profiling. This technology offers great potential for development as a tool to allow rapid noninvasive assessment of embryonic reproductive potential before transfer. (Fertil Steril 2008;90:77–83. ©2008 by American Society for Reproductive Medicine.)

Key Words: Implantation, Raman, biospectroscopy, metabolomics, embryo quality, reproductive potential, culture media, IVF, delivery rate

Use of the assisted reproductive technologies (ART) continues to increase, with 127,977 treatment cycles being initiated in the United States in 2004. This trend is driven by the steady improvement in ART delivery rates, improving access to care in many areas, and the relative ineffectiveness of other treatment options. At the current time, >1% of all children born in the United States are from ART-related conceptions (1).

The high success rates enjoyed through IVF are attained in many cases only through the simultaneous transfer of multiple embryos. In 2004, a mean number of 2.8 embryos were transferred, leading to a 27.7% delivery rate per initiated IVF cycle and an overall multiple birth rate of 33.5%. Even in oocyte donation cycles where the embryos should possess the highest reproductive potential, the transfer of more than one embryo led to a multiple gestation rate of 42% (1).

The risks related to multiple gestations are well known, and include preterm delivery with all of the sequelae, low birth weight, and a dramatic increase in the relative risk for cerebral palsy. Decreasing the prevalence of multiple gestations while maintaining or improving overall pregnancy rates remains the most significant contemporary goal of infertility research.

The sentinel issue surrounding multiple gestations and IVF is the inability to precisely estimate the reproductive potential of an individual embryo. Almost from the inception of IVF, clinicians and scientists have sought to develop techniques that would allow them to accurately make this assessment. Initial efforts focused principally on morphologic assessment of the embryo to estimate reproductive potential. Grading
systems based on embryo morphology and cleavage rates were developed by several investigators and have continued to evolve (2–8). Significant improvements in implantation and pregnancy rates and reductions in multiple gestation rates were attained through careful application of these scoring systems (9). Unfortunately, their precision is still insufficient to compel most patients and clinicians to reduce the number of embryos transferred to a point where twins are uncommon and high-order multiple gestations are rare or eliminated. In 2004, only 7.2% of day 3 embryo transfer and 8.9% of day 5 transfers involved only a single embryo (1). No doubt that in many of these cases, the decision does not reflect a desire to proceed with a conservative transfer, but rather the availability of only a single embryo.

The limits of cleavage rate and morphologic evaluation of embryos has led many investigators to pursue adjunctive technologies that might aid in assessing the reproductive potential of a given embryo. Gardner and colleagues (10) have reported that nutrient use is correlated with the highest morphologic grade in human blastocysts. Conaghan et al. (11) showed an inverse relationship between pyruvate uptake and human embryo development. Other investigators have recently demonstrated that changes in amino acid metabolism correlates with embryonic implantation potential (12, 13). These insightful data provide strong evidence that there are metabolic differences between embryos that appear healthy and those whose development is impaired.

Despite promise, the application of these technologies to the clinical setting has been limited. These techniques may be extremely expensive, require dedicated equipment and staff, or be invasive to the embryo. Many cannot produce results rapidly enough to allow the information to be used to clinically in the limited window of time acceptable for embryo transfer. Most importantly, none of these technologies has been shown in a blinded fashion to correlate with subsequent implantation potential of embryos that have been transferred. The need to identify a technology that predicts embryo viability through a rapid, noninvasive, consistent, and clinically applicable platform remains.

One approach to addressing the problem of evaluating embryonic reproductive competence would be metabolomic profiling of spent culture media. It seems intuitive that the metabolism of a healthy embryo might alter the surrounding environment differently from one that is less healthy, and thus possessing less reproductive potential. Metabolomic profiling should offer the advantage of simultaneous evaluation of a large number of analytes within the media, and may produce a profile that might more clearly be associated with embryo viability. We have recently shown that metabolomic profiling of human embryo culture media using Raman or near-infrared spectroscopy, combined with bioinformatics, correlates with pregnancy outcome (14). The purpose of the current study was to determine if the metabolomic profile of spent culture media could predict embryonic reproductive potential based on a previously described algorithm (14).

MATERIALS AND METHODS
Patient Population
All patients attempting conception through IVF and who gave informed consent for evaluation of their discarded material were considered for participation in the study. All treatment cycles were performed at one clinical site (Reproductive Medicine Associates of New Jersey). This study was conducted under an institutional review board-approved protocol.

The patients underwent a variety of stimulation and monitoring protocols as previously described (15). Follicular aspiration was completed, the oocyte cumulus complexes isolated, the majority (but not all) of the cumulus mechanically stripped, and the oocytes placed into individual 50-μL droplets of media. Conventional insemination or intracytoplasmic sperm injection was use as indicated. Tri-gas incubators, which provide a 5% oxygen environment, were used throughout the study.

After confirming fertilization on day 1 of development, embryos with two pronuclei were placed into individual 50-μL droplets for culture to the cleavage stage (day 1 through day 3). Quinn’s Advantage Cleavage Medium (SAGE In-Vitro Fertilization, Inc., Trumbull, CT) supplemented with 5% Serum Substitute Supplement (Irvine Scientific, Santa Ana, CA) was used for culture from day 1 to day 3. All embryos were cultured in individual droplets at all times.

Patients who did not have a sufficient number of high-quality embryos to justify extended culture had their embryos selected for transfer on day 3. Following removal of the embryos in preparation for actual transfer, the dishes containing the spent media were placed into a research incubator. The spent media samples were then placed individually into labeled cryovials, snap frozen in liquid nitrogen, and stored at −80°C.

In those cases where six or more high-quality embryos were present on day 3, the embryos were placed into extended culture media and continued until day 5. BlastAssist media supplemented with human serum albumin (MediCult, Jyllinge, Denmark) was used for extended culture.

In cases where pregnancies occurred, serial hCG levels and appropriately timed ultrasonography was done as necessary to document fetal cardiac activity and confirm viability. The patients then returned to their obstetricians. Confirmation of viable delivery was confirmed in each case.

Experimental Design
The patients who had attained 100% implantation rate progressing to delivery represent the group whose embryos had proven reproductive potential and were identified first. In each case, the next consecutive patient of the same age with a basal FSH level within 1 IU/L with a 0 implantation rate was selected to represent those embryos with the lowest implantation potential.
The spent media samples from embryos of the selected patients were deidentified, labeled with a randomly assigned accession number, and shipped on dry ice to the metabolomics laboratory at McGill University. The specimens were maintained frozen at all times including during the relabeling process. The individuals relabeling the specimens and the spectral analysis team at McGill University were blinded to the implantation rates at all times.

Raman biospectroscopy was completed, and a viability index was calculated and assigned to each of the blinded samples. These results were provided to the second clinical center, and the blind was broken to allow evaluation of the results.

**Raman Biospectroscopy**

Before analysis, samples were thawed at room temperature (25°C) for 30 minutes. Raman analysis was conducted using a HeNe based Raman system (DeltaNu Advantage 100, Laramie, WY), with a reflective spectrograph with a CCD detector. Raman-compatible 1-mm diameter quartz sample cells were filled with 15 μL of sample media for the spectral measurements. Spectra were recorded from 50 to 3,450 cm⁻¹ at room temperature (25°C ± 1°C). Integration time was 5 minutes per sample. Both an empty sample cell and a water spectrum were repeated with every four samples to monitor dark signal and laser drift in the system.

**Spectral Analysis**

Spectral analysis was begun by subtracting dark signal background from every spectrum. The means of the spectra from all samples were determined and subtracted from all of the samples obtained at the respective site. These mean centered Raman spectra were then used for all subsequent calculations.

Quantification of properties from each Raman spectra involved conversion of the data at each wavelength into a wavelet domain that facilitates the application of the genetic algorithms used for modeling. To obtain data in the wavelet domain, a “son” Haar wavelet transformation was employed. Details of the method are described in detail elsewhere (14, 16). Briefly, a custom genetic algorithm program was written in MatLab, (The Math Works, Inc., Natick, MA), and a specific mathematical model using the calculation of viability derived in a prior study involving open evaluation of samples with known clinical outcomes (14). The algorithm was applied to the subset of coefficients obtained from the son Haar Transform and related to the embryos implantation rate. Transformed data throughout the spectra were evaluated and a unique viability index was obtained for each sample.

**Statistical Analysis**

The viability indices of those samples that contained embryos that had implanted were compared with those that did not using Student's t-test. Similar comparisons were also made of the results obtained from samples collected on day 3 versus those collected on day 5. Possibly more important in this setting is determining if a threshold value might be used to predict major differences in implantation potential. To this end, receiver operator characteristic (ROC) curves were created and used to identify the threshold value with overall best mix of sensitivity and specificity. This threshold value was then applied to the data set to calculate positive and negative predictive values, and overall accuracy. Statistical analyses were done using Analyse-it statistical software (Analyse-it Software LTD, Leeds, UK). Alpha error of <0.05 was considered significant for all comparisons.

**RESULTS**

Eleven consenting patients had 100% implantation rates during the study interval. Eleven matched controls were subsequently selected based on age and basal FSH levels. Subsequently three patients, all from the matched control group, were noted to have abnormal endometrial development. One patient had a peak endometrial thickness of 5 mm, and two patients failed to develop a trilaminar endometrium. All were excluded before analysis of the biospectroscopy results.

A total of 44 blinded samples obtained from 19 patients were analyzed using Raman biospectroscopy. Thirty five samples from 14 patients (74%) had been collected on day 3, whereas nine samples from five patients were obtained on day 5 (26%). Raman spectra were obtained on each sample. A large variation was observed in signals obtained from different samples consistent with variations expected from embryo metabolism. The mean centered spectra are presented (Fig. 1). The established mathematical model (14) was then applied and each sample assigned a viability index. The samples and spectra were completely blinded during this process.

**Day 3 versus Day 5 Samples**

The results were unblinded and in the initial analysis, the day 3 samples were compared with the day 5 samples. The viability indices on day 3 (0.71 ± 0.07) were significantly higher than those on day 5 (0.56 ± 0.07; P<.0001) (Fig. 2A). This likely reflects the known significant differences in embryonic metabolism at these stages of development. The data were separated relative to the day of transfer for all further analyses.

**Analysis of Day 3 Samples**

Evaluation of the day 3 samples demonstrated significantly higher viability indexes from the spent media of embryos with proven reproductive potential (0.875 ± 0.12) compared with those that failed to implant (0.56 ± 0.09) (P<.03). Two of the data points, one from each outcome group, were significant outliers (>3 SDs from their respective means) and may represent possible oil contamination of the test cell. These data points were excluded and the data were compared again with similar findings. Specifically, the viability indexes of those samples associated with implantation (N = 16) were significantly higher (0.94 ± 0.1) than those that were not (N = 17) (0.51 ± 0.07) (P<.002) (Fig. 2B).

A ROC curve was done to help select a threshold that might be used to then calculate sensitivity and specificity, predictive values, and overall accuracy. The ROC curve indicates that a viability index of 0.76 provides the greatest
discrimination (Fig. 3A). Using a threshold value of 0.76 to distinguish embryonic reproductive potential (higher values associated with implantation and lower samples with a failure to implant), the sensitivity for and specificity for predicting a failure to implant was 82.4% (14 of 17) and 69% (11 of 16), respectively. Overall accuracy (i.e., correctly predicting implantation) was 75.8% (25 of 33).

Analysis of Day 5 Samples

The samples obtained on day 5 at the time of blastocyst transfer were similarly evaluated. The samples associated with implantation appeared to have higher viability indexes ($-0.81 \pm -0.08$ vs. $-0.21 \pm -0.14; P<0.009$). The distributions of results are presented in Figure 2C. An ROC curve was completed that identified the best threshold value at $-0.43$ (Fig. 3B). Using that threshold, the sensitivity and specificity were both 100%. The overall accuracy was also 100% (8 of 8).

Overall Results

Evaluation of the data set as a whole confirms the strong association with the metabolomic profile of the spent culture media and clinical outcomes. The day 3 specimens were evaluated relative to their threshold and the day 5 specimens relative to their threshold. The overall sensitivity and specificity for predicting a negative result were 85.7% (18 of 21) and 80% (15 of 20), respectively. Overall accuracy for biospectroscopic metabolomic profiling was 80% (33 of 41).

DISCUSSION

In this study, we found that rapid, noninvasive analysis of embryo culture media using Raman biospectroscopy and bioinformatics is highly associated with the reproductive potential of in vitro cultured human embryos. A previously developed model (14) based on Raman biospectroscopic evaluation of samples with known outcomes was applied in this prospective blinded study. The overall accuracy of the technique for predicting outcome was 82% in a cross-sectional population. These findings indicate that embryos with greater reproductive potential impact their environment differently from those with lesser potential, and that the difference is detectable using Raman-based biospectroscopy.

Raman is a vibrational spectroscopy commonly used for assessing molecular motion and fingerprinting species. Raman is based on inelastic backscattering of a monochromatic excitation source within a routine energy range of 200–4000 cm$^{-1}$. Advantages of Raman spectroscopy include its compatibility with wet samples and normal ambient temperature. Vibrational spectroscopy has long been used for constituent analysis in the chemical industry. For example, agricultural products or tissue are commonly analyzed based on their spectra (16–20). Recently, with the availability of low-cost laser systems and high-sensitivity array detectors, there has been an increased interest in use of Raman spectroscopic signatures to allow qualitative as well as quantitative assessment of diverse samples ranging from food stuffs and pharmaceuticals to explosives. The high specificity of the vibrational modes of molecules accessible by Raman provides a convenient, rapid, and cost-effective source for constituent identification from the background. Raman spectroscopy has also been shown capable of detecting a variety of metabolites such as lactate, glucose, and fatty acids in biofluids (21). In addition, vibrational modes of readily modified constituents such as sulfydrrls are visible in the Raman spectra.

One of the major advantages of a spectroscopic approach to assessment of spent culture media is that whole sample matrices can simultaneously be monitored noninvasively without the need to evaluate any single or small group of
analytes. A spectral signature provides detailed multicomponent integrated information about a given sample in one simple measurement.

The technique has potentially powerful advantages relative to techniques that provide direct quantitative or qualitative analysis of any single or small group of analytes. It is capable of detecting many of the same quantitative or qualitative changes, but will allow detection of subtle metabolic changes in the sample. Alterations such as hydroxylation should be detectable without the necessity of detecting changes in the actual contents of the media. This might be very important, because oxidative stress at the cellular level can produce extensive modification of the molecules in the immediate milieu without changing their production.

One concern regarding the application of this technology to different IVF laboratories is the substantial differences in the content of commercial culture media that they might use. Even batch-to-batch variability within one formulation might impact such a sensitive technique. A valid and widely applicable methodology for the evaluation of spent culture media should be able to detect the changes associated with embryonic reproductive potential and not be impacted by these factors. In this study, two different media with two different protein supplements were used. Equivalent predictive values were attained, indicating that at least the differences in the composition of those two media formulations do not alter or mask the biospectroscopic results. This suggests, but does not yet demonstrate, that the difference being detected through analysis of the Raman spectra are the impact of oxidative stress or other similar processes on the media itself and not necessarily huge shifts in the production or clearance of any single analyte in the system.

The advantages of having a rapid and reliable adjunct to the assessment of embryo reproductive potential are numerous. A better understanding of the implantation potential of a given group of embryos would allow embryologists, clinicians, and patients to make more precise decisions about the number of embryos to transfer. This, in turn, should reduce

![Figure 2](image-url)

**FIGURE 2** Viability indexes calculated from Raman biospectroscopic analysis of blinded spent media samples that had contained embryos with known implantation potential (A). Viability indexes are significantly different from media collected on cycle day 3 compared with media collected on cycle day 5 ($P < .0001$) (B). Distribution of viability indexes on spent media collected on cycle day 3 demonstrating a significant difference in those embryos with 0 and 100% implantation rates ($P < .03$). Each point represents an individual sample, whereas different colors represent different patients (C). Distribution of viability indexes on spent media collected on cycle day 5 (blastocyst) demonstrating significant differences between those embryos with 0 and 100% implantation rates ($P < .009$). Each point represents an individual sample, whereas different colors represent different patients.
the risk of multiple gestation while maintaining or even increasing pregnancy rates.

Enhanced understanding of the reproductive potential will also be valuable when counseling patients about the cause of a failed treatment cycle. Many times patients have morphologically normal embryos transferred without attaining a pregnancy. Understanding the etiology of that failure is paramount to making sound treatment decisions about future treatment options. Although some patients will want to cycle again even with the knowledge that they have significant embryo quality issues, at least they might truly make an informed decision as possible. The data obtained through biospectroscopic assessment of embryonic reproductive potential might also be useful in patients with multiple unexplained failed cycles who are struggling with decisions regarding empiric gamete donation or the use of a gestational carrier.

Biospectroscopy-based metabolomic profiling might also be valuable in tracking equipment or technique drift in a given laboratory. Consistent changes in the reproductive potential of embryos from one incubator or following the procedures of a single technician might allow early and precise identification of the potential problem. This would allow corrective action to be taken at the earliest possible time and provide great protection for the patients undergoing treatment.

Further study will certainly be required to fully characterize the value and the limits of this technology in IVF. At this time, this technology offers great potential for development as a tool to allow rapid noninvasive assessment of embryonic reproductive potential before transfer.

REFERENCES
