Nuclear magnetic resonance metabolomic profiling of Day 3 and 5 embryo culture medium does not predict pregnancy outcome in good prognosis patients: a prospective cohort study on single transferred embryos

K. Kirkegaard1,4,*, A.S.P. Svane2, J.S. Nielsen2, J.J. Hindkjær1,3, N.C. Nielsen2, and H.J. Ingerslev1,3

1Centre for Preimplantation Genetic Diagnosis/The Fertility Clinic, Aarhus University Hospital, Brendstrupgaardsvej 100, Aarhus N 8200, Denmark 2Center for Insoluble Protein Structures, Interdisciplinary Nanoscience Center (iNANO) and Department of Chemistry, Aarhus University, Gustav Wieds Vej 14, Aarhus C 8000, Denmark 3Health, Aarhus University, Vennelyst Boulevard 9, Aarhus C DK-8000, Denmark 4Present address: Department of Clinical Biochemistry, Aarhus University Hospital, Aarhus, Denmark

*Correspondence address. E-mail: kirstine.kirkegaard@clin.au.dk

Submitted on October 16, 2013; resubmitted on August 15, 2014; accepted on August 26, 2014

STUDY QUESTION: Does the metabolomic profile, obtained with nuclear magnetic resonance (NMR), of spent culture media from human embryos correlate with reproductive potential in a cohort of good prognosis patients?

SUMMARY ANSWER: In a large cohort of single transferred blastocysts from a homogeneous group of good prognosis patients, we find a high degree of individual variation in the metabolome that, however, has no relation to pregnancy outcome.

WHAT IS KNOWN ALREADY: Differences among various specific metabolites have been linked to reproductive potential. Although results from retrospective near infrared (NIR) spectroscopy analyses of spent culture medias from transferred embryos were promising, randomized controlled trials were unable to demonstrate that NIR analysis improved pregnancy rates. Therefore, a more detailed investigation of the relation between embryo metabolism and reproductive potential is required. NMR is a powerful technique that provides detailed structural and dynamic information.

STUDY DESIGN, SIZE, DURATION: A prospective cohort study was conducted at the Fertility Clinic, Aarhus University Hospital between February 2011 and July 2012. Infertile patients aged < 38 years without endometriosis were offered participation and their embryos were included if greater than or equal to eight oocytes were retrieved. In total, 161 infertile patients were included in the cohort.

PARTICIPANTS/MATERIALS, SETTING, METHODS: Spent culture media was collected on Days 3 and 5 after oocyte retrieval from 148 single transferred embryos. NMR spectra were obtained from 12 μl of spent media. Data were quantitatively analysed using multivariate analysis with respect to pregnancy outcome, defined as a live fetus by ultrasound in gestational Week 8, along with patient and treatment related variables such as embryo score, age, BMI, fertilization method and cause of infertility.

MAIN RESULTS AND THE ROLE OF CHANCE: A total of 148 cycles were included in the analysis [embryo transfer cancelled (n = 12), no media collected (n = 1)]. Clinical pregnancy was confirmed in 47 patients (32%). We obtained high quality NMR spectra for 141 Day 3 and 137 Day 5 samples. Our spectra show a high degree of individual variation. Multivariate data analysis was performed on spectral data with several different pre-processing combinations, i.e. binning, alignment, normalization and scaling in the attempt to develop a valid prediction model. Different strategies of multivariate analysis showed, however, no correlation between the NMR profiles and pregnancy outcome, patient or treatment characteristics. No model could therefore be developed for prediction of pregnancy outcome. We conclude that within this group of good prognosis patients, large-scale metabolic variations between embryos detected with NMR have no apparent association with pregnancy outcome.
LIMITATIONS, REASONS FOR CAUTION: Although this study is the largest we know of using NMR to investigate metabolomic profiles of single-transferred embryos, there may be differences that would be detected with a larger study. When analysing such a small sample volume, even small variations in the amount of media and dilution may introduce a large uncertainty in the results.

WIDER IMPLICATIONS OF THE FINDINGS: Our study questions the usefulness of the entire metabolome for embryo selection, which should direct the search for viability markers in the culture media towards individual components.

STUDY FUNDING/COMPETING INTERESTS: Funding was provided by Aarhus University, the Lippet Foundation, the Toyota Foundation, the Aase og Einar Danielsen foundation. Research at the Fertility Clinic, Aarhus University Hospital is supported by an unrestricted grant from MSD and Ferring. The authors declare no competing interest.

TRIAL REGISTRATION NUMBER: NCT01139268.

Key words: assisted reproduction / human / nuclear magnetic resonance / embryo selection / pregnancy

Introduction

Elective transfer of a single embryo (SET) is increasingly recommended as best practice in IVF treatment and is in several countries mandatory when treating certain patient groups (Maheshwari et al., 2011; Gianaroli et al., 2012). The unequivocal aim is to avoid multiple pregnancies while maintaining high pregnancy rates, which requires reliable selection of competent embryos. Embryo selection is currently based on morphology (ESHRE/ALPHA, 2011). Although assessment of morphological parameters at given, discrete inspection points is correlated fairly closely with the reproductive potential (ESHRE/ALPHA, 2011), the inability to reliably predict embryo viability is often considered a main reason for the relatively low pregnancy rate following assisted reproduction treatment. As a result, new ways to identify the most viable embryo has become a subject of immense interest. Among several promising strategies, the search for specific informative metabolites has been evaluated as a non-invasive approach that not only potentially could correlate with viability, but also contribute information on embryo physiology (Gardner and Wale, 2013). Another approach has been to analyse metabolomic spectra with near infrared (NIR) spectroscopy. Results from several retrospective trials assessing embryo viability by NIR were encouraging, which led to the development of a prototype NIR spectrometer intended for clinical IVF (Molecular Biometrics, Inc., Norwood, MA, USA). However, two randomized controlled trials (RCTs) conducted with the prototype were unable to demonstrate that NIR analysis improved pregnancy rates (Hardarson et al., 2012; Vergouw et al., 2012). These studies raised the question of whether the absence of improved clinical outcome using the NIR technology reflected the use of an inappropriate method or rather was the result of inherent limitations in identifying viable embryos based on their metabolism (Hardarson et al., 2012; Vergouw et al., 2012). To pursue the metabolomic approach, a more detailed analysis of the embryo-induced changes in media components would therefore be required. Nuclear magnetic resonance (NMR) spectroscopy is intrinsically rich in information and is therefore used to obtain metabolic profiles with structural and quantitative information enabling metabolite identification. The aim of the present study was therefore to identify metabolites in the culture media potentially related to viability by performing NMR analysis on spent culture media from single-transferred blastocysts in a cohort of prospectively recruited good prognosis patients.

Materials and Methods

Study design and participants

Between February 2011 and July 2012, infertile patients were recruited prospectively at the Fertility Clinic, Aarhus University Hospital, Denmark. The cohort was established for evaluation of three approaches to improve selection of the embryo with the greatest implantation potential, i.e. time-lapse analysis (Kirkegaard et al., 2013), NMR and gene expression analysis of trophectoderm (TE) biopsies (O-215). Patients were offered participation if the woman was aged <38 years and had no endometriosis. Embryos were included if greater than equal to eight oocytes were retrieved. Eligible patients could contribute to the study with one treatment cycle only. Data related to patient characteristics were obtained for the current treatment cycle. In total, 161 patients were included in the cohort.

Ethical approval

Written informed consent was obtained from all participants before inclusion. Patients consented to blastocyst culture (Day 6), SET, time-lapse imaging, analysis of the spent culture media and blastocyst biopsy. The Central Denmark Region Committees on Biomedical Research Ethics and the Danish Data Protection Agency approved the study. The study was registered at ClinicalTrial.gov with accession number NCT01139268.

IVF, embryo culture and collection of media samples

Ovarian stimulation and oocyte retrieval were performed according to standard procedures as previously described (Kirkegaard et al., 2012). Following retrieval, oocytes were placed in Cook fertilization medium (COOK®, Brisbane, Australia) and fertilized with conventional ICSI or IVF procedures. ICSI fertilized embryos were placed in individual wells (EmbryoSlide, Unisense Fertitech, Aarhus, Denmark) immediately after injection and cultured in a tri-gas time-lapse incubator (EmbryoScope, Unisense Fertitech, Aarhus, Denmark). To secure optimal conditions for confirmation of fertilization before denudation, IVF embryos were cultured ~18 h in a conventional incubator (Galaxy R, RS Biotech, CM Scientific, West Lothian, UK) under oil at 37°C, 20% O2 and 6% CO2 before transfer to the EmbryoScope. In the EmbryoScope, embryos were cultured at 37°C, 5% O2 and 6% CO2 in 25 µl individual wells containing droplets of Sydney IVF Cleavage Medium (COOK®) under oil with change to Sydney IVF Blastocyst Medium (COOK®) 68 h (Day 3) after fertilization. On Day 5, a TE biopsy was obtained from the embryo intended for transfer, if the procedure was logistically feasible. The biopsy was obtained for research purposes only. In total,
Nuclear magnetic resonance of culture media and pregnancy

23 transferred embryos were biopsied. Blastocysts were graded according to the Gardner criteria; in brief, based on expansion of the blastocoeel cavity (1–6), number and cohesiveness of the inner cell mass (ICM) and TE (A–C). The single embryo with the highest morphological grade was selected for transfer. The obtained time-lapse data were not used for selection and the results have been presented in a separate publication (Kirkegaard et al., 2013). All transfers were performed on Day 6, motivated by the requirement for regeneration in case of a biopsy (Kokkali et al., 2005). The spent culture medium was collected individually on Days 3 and 5 from all embryos that displayed normal fertilization (two pronuclei) and completed the first division, and stored in individually labelled vials at −80°C.

Biochemical pregnancy was confirmed by serum β-hCG measurement 16 days after oocyte retrieval. Clinical pregnancy rate was registered as number of ongoing pregnancies per embryo transfer, based on presence of a fetal heart beat (FHB) visualized by ultrasound 8 weeks after embryo transfer.

NMR analysis

The collected media from the transferred embryos were thawed, vortexed and centrifuged. Growth medium (12 µl) was subsequently mixed with 488 µl of a solution of 30 µM DSS (4,4-dimethyl-4-silapentane-1-sulphonic acid) in 10/90% D2O/MiliQ water. DSS was used as a standard reference compound. NMR spectra were collected on a Bruker Avance III spectrometer (500.13 MHz) equipped with a temperature-regulated Bruker SampleJet and using a standard inverse triple-resonance TXI 5-mm probe. 1D spectra at 278 K were collected with 768 scans using the WS pulse sequence (Liu et al., 1998) to suppress water.

Pilot studies on comparable samples were conducted to establish optimal parameters and ensure the stability of the measurements. The spectra proved highly reproducible and it was possible to easily distinguish cleavage media samples from blastocyst media samples (Supplementary data, Fig. S1), as well as spent media from unspent samples (Supplementary data, Fig. S2). Three large signals between 2.9 and 3.6 ppm showed substantial variation in chemical shift and line broadening, most likely caused by small variations in pH between samples. Several compensatory strategies were applied, including binning, alignment and removal of the spectral region containing the shifting signals. Water and DSS signals were removed before data analysis.

Statistical analysis

Processing, referencing and phasing of the 1D spectra were performed in order to minimize any differences in the spectra by comparing spectral overlays using Topspin 2.1 (Bruker Biospin). The shifting signals between 2.9 and 3.6 ppm were aligned in MATLAB 7.13 (MathWorks, Inc., Natick, MA, USA) using the icoshift tool (Savorani et al., 2010). Both regular and dynamic bins were applied. Dynamic bins were determined using the Dynamic Adaptive Binning (DAB) Algorithm (Anderson et al., 2011) in MATLAB 7.13 (MathWorks, Inc.). The intensities were normalized to the overall intensity of each spectrum in order to adjust for small variations in sample volume and to DSS (Craig et al., 2006). Outliers were identified and removed using principal component analysis (PCA) score-plots, residuals and Hotellings T2. The number of excluded outliers was variable and depended on which type of model was attempted. Initially a conservative approach to outlier exclusion was preferred, where we removed only samples where the cause of the outlier behaviour was experimental, e.g. contamination of the growth media by the oil used during incubation. We typically excluded between 10 and 30 samples from the models. The binned and normalized data were quantitatively analysed using several multivariate data analysis methods, such as PCA, partial least squares (PLS), PLS discriminant analysis and orthogonal projections to latent structures (OPLS), to determine whether variations in the spectra were significantly related to pregnancy outcome. Patient- and treatment-related variables, such as embryo score, age, BMI, fertilization method and cause of infertility, were also included in the analysis as possibly predictable factors. Both pareto and unit variance scaling was used in different models in the multivariate data analysis. Development of prediction models was attempted for prediction of pregnancy outcome or patient-/treatment-related variables. The statistical analysis was conducted using the Simca-P+ software (Umetrics, Umeå, Sweden).

Differences in baseline and demographic data between the pregnant and non-pregnant group were tested with Student’s t-test (continuous, normal distributed data), Wilcoxon rank-sum test (continuous, non-normal distributed data) or Fisher’s exact test (categorical data) in STATA, version 11.0 (StataCorp, College Station, TX, USA). Two-tailed P-values <0.05 were considered significant.

Results

In total, 161 infertile patients were included in the cohort. Embryo transfer was cancelled for 12 patients that were consequently excluded from the analysis. In 2 of the cancelled cycles, good quality blastocysts were frozen for later transfer, whereas the remaining 10 cycles were cancelled due to no blastocyst formation. In 1 cycle, no media was collected, which left 148 SET cycles for the analysis. Clinical pregnancy was confirmed for 47 patients (32%), with no pregnancy for 101 transfers. Female age was lower in the pregnant group, where the embryos were more expanded and had a higher TE and ICM score on the day of transfer compared with embryos from the non-pregnant group (Table 1). A stratified analysis showed no effect modification of the biopsy procedure (P = 0.23; RR biopsy/no biopsy 1.5 (0.86; 2.5)) (Supplementary data, Table S1). We obtained high quality spectra for 141 Day 3 and 137 Day 5 samples (Fig. 1). Multivariate data analysis was performed on spectral data with several different pre-processing combinations, i.e. binning, alignment, normalization and scaling as described in Materials and Methods, in the attempt to develop a valid prediction model. Moreover, development of prediction models for pregnancy outcome was attempted on both full-spectral-range1H NMR spectra and after narrowing spectral regions down to regions of interest or with high variation. Regions of interest were regions containing the signals from some of the particular metabolites that have been correlated previously with pregnancy outcome, such as carbohydrates, pyruvate and different amino acids, or regions that in our spectra displayed a high degree of variation. The multivariate analyses showed no correlation between the spectral variance of the datasets and pregnancy outcome for any of the described strategies. For none of the attempted models did Q2 values exceed the recommended value of 0.5. Even for the models with the highest Q2 values (0.3 at maximum), we found no clear clustering of data that could separate pregnant from non-pregnant samples. Therefore, no model could be developed for prediction of pregnancy outcome.

Figure 2 represents an OPLS plot, which is included to illustrate the lack of correlation with pregnancy outcome. These figures represent examples of attempted models, but similar distributions were seen when plotting other principle component combinations or when colouring according to other patient data such as BMI. Several signals stand out in particular, being present in only a few samples: one metabolite with peaks at 3.74 and 1.46 ppm, and another with peaks at 3.72, 2.80 and 2.60 ppm: plausible metabolites with these peaks...
could be alanine and aspartate, respectively. We ruled out errors in sample handling as the source of these metabolites, i.e. the amounts were on the same scale as the other metabolites, and the signals only appeared in very few samples, which were not related to preparation time. The up-regulation of these metabolites was not observed consistently at either Day 3 or 5, or at both Days 3 and 5 for individual embryos. There was no correlation between patient and treatment related variables (age, BMI, cause of infertility, fertilization method, embryo score) or pregnancy outcome and the presence of up-regulation of either of these two metabolites.

Discussion

We successfully obtained highly detailed spectral information from small amounts of spent culture medium from single-transferred embryos in this large, prospectively recruited cohort study of good prognosis IVF patients. The spectral analysis was conducted on both Day 3 and 5 media from single-transferred blastocysts, which provides a direct correlation between pre- and post-compaction metabolism and viability. We found no relationship between spectral variations and pregnancy outcome.

In contrast to our findings, several studies have previously reported a correlation between viability and various aspects of embryo metabolism. Being related to the production of ATP, oxygen consumption has been used as an indicator of overall metabolic activity and oxygen consumption has been found to be higher among oocytes with successful fertilization (Scott et al., 2008) and implantation (Meseguer et al., 2011) than among unsuccessful oocytes. Several challenges in designing a probe that can accurately measure the oxygen consumption of individual embryos at reasonable costs have so far limited the widespread use of this technique.

Another strategy has been to investigate embryo metabolism by measuring the turnover of individual metabolites or the complete inventory of small-molecule metabolites, the metabolome, in the spent culture media. Correlations have been found between viability and the metabolism of media components such as carbohydrates, pyruvate and different amino acids (Hardy et al., 1989; Conaghan et al., 1993; Gardner et al., 2001, 2011; Houghton et al., 2002; Brison et al., 2004).

### Table I

<table>
<thead>
<tr>
<th></th>
<th>Pregnant (n = 47)</th>
<th>Non-pregnant (n = 101)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female age (years)</td>
<td>29.8 (3.3)</td>
<td>31.1 (3.4)</td>
<td>0.03</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>22.65 (21.2; 27.6)</td>
<td>22.55 (21.2; 26.0)</td>
<td>0.93</td>
</tr>
<tr>
<td>Number of previous IVF treatment cycles</td>
<td>1 (0; 1)</td>
<td>0 (0; 1)</td>
<td>0.71</td>
</tr>
<tr>
<td>Number of aspirated oocytes</td>
<td>12 (10; 15)</td>
<td>13 (10; 16)</td>
<td>0.14</td>
</tr>
<tr>
<td>Number of fertilized oocytes (two pronuclei)</td>
<td>6 (4; 9)</td>
<td>6 (4; 8)</td>
<td>0.92</td>
</tr>
<tr>
<td>Method of fertilization</td>
<td>ICSI 28</td>
<td>63</td>
<td>0.86</td>
</tr>
<tr>
<td></td>
<td>IVF 19</td>
<td>38</td>
<td></td>
</tr>
<tr>
<td>Total FSH dose</td>
<td>1500 (1200; 2000)</td>
<td>1720 (1300; 2280)</td>
<td>0.08</td>
</tr>
<tr>
<td>Cause of infertility</td>
<td>Male factor 30</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tubal factor 8</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Other, unexplained 9</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>Degree of expansion on transfer (Day 6)</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>26</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>21</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>22</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Morphological grade of the ICM</td>
<td>&lt;0.001</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>33</td>
<td>36</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>14</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>Morphological grade of TE</td>
<td>0.06</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>32</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>15</td>
<td>37</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Number of embryos cryopreserved</td>
<td>2 (1; 4)</td>
<td>2 (0; 3)</td>
<td>0.02</td>
</tr>
<tr>
<td>Media missing/excluded</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Day 3</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Day 5</td>
<td>2</td>
<td>4</td>
<td></td>
</tr>
</tbody>
</table>

Continuous data are expressed as mean and SD, or median and interquartile range if the assumption of normality was not fulfilled. Categorical data are presented as number of cases. For testing differences between the two groups, Fisher’s exact test was used for categorical data and Student’s t-test for continuous normal distributed data, and Wilcoxon rank-sum test for non-normal distributed data.
In particular, the metabolism of glucose has attracted attention as post-compaction glucose consumption has been reported to predict blastocyst development and, recently, live birth (Gardner et al., 2001, 2011), but the findings remains to be validated. Findings from studies on amino acid turnover are inconsistent with regard to which amino acids might predict viability. In the present study, different peaks that varied among individual embryos could be identified in the spectra, which allowed for an identification of specific metabolites. As none of these

---

**Figure 2** OPLS-discrimination analysis plot for Day 5 samples, as an example of a model for human pregnancy outcome. Dotted line is 2 SD, stapled line is 3 SD. In this example, only non-implanted embryos and embryos with FHB were included, i.e. missed abortions and biochemical pregnancies were excluded ($n = 18$), as it gave the best separation between the groups. After a PCA outliers ($n = 15$) were excluded, which left 109 samples for the model (one component model, pareto scaling, $R^2_X = 0.452$, $R^2_Y = 0.0947$, $Q^2 = 0.0522$), clearly supporting the lack of association.

**Figure 3** PCA score-plot of the first two principal components ($t_1$ and $t_2$) of data from human cleavage samples (Day 3), coloured according to FHB: positive = green triangle, negative = red circle. Spectra have been Fourier transformed using a 2 Hz exponential line-broadening and binned using DAB. Five component model $n = 129$, $R^2_X(cum) = 0.72$ and $Q^2(cum) = 0.60$. Clear outliers were removed before making the two-class model between media from embryos with negative transfer results and from embryos developing to display FHB. OPLS was used, but no orthogonal component was found to strengthen the model. Thus, the $X$-axis displays sample number.
individual metabolite concentrations correlated with pregnancy outcome, we did not pursue this any further.

The limitations related to the evaluation of specific metabolites have led others to investigate the metabolome, as it represents the end product of the cellular processes in the embryo, and thus provides a more general overview of the embryo metabolism. Most studies investigating the metabolome have used vibrational techniques, such as NIR and Raman spectroscopy, since they are relatively simple and inexpensive platforms and thus potentially applicable in a clinical setting. In a pioneer study conducted using both Raman and NIR spectroscopy, key spectral regions related to viability were identified and an algorithm for generating a viability score was developed (Seli et al., 2007). The positive relationship between a high viability score and implantation was later confirmed in several other studies using NIR spectroscopy (Vergouw et al., 2008, 2011; Seli et al., 2010) and appeared to be independent of culture medium and culture conditions (Ahstrom et al., 2011). Two RCTs conducted with a clinical bench-top prototype were however unable to demonstrate that NIR analysis improved pregnancy rates (Hardarson et al., 2012; Vergouw et al., 2012). The failure to reproduce the findings from previous retrospective studies was partly attributed to a limited sensitivity of the prototype. NIR analysis is in general used for overall spectral profile comparison, primarily aimed at distinguishing viable from non-viable embryos with regard to either developmental or pregnancy potential, while it is not typically used for target metabolite identification (Botros et al., 2008). In contrast, NMR spectroscopy provides comprehensive chemical information about the composition of unknown materials, even with a small sample volume, provided that a powerful magnetic field is applied. NMR is therefore considered a suitable choice for exploratory studies on metabolism. The number of NMR studies performed on human embryo samples is, however, limited and the findings are ambiguous. In two studies, a positive correlation was found between implantation and NMR profiling (Seli et al., 2008; Pudakalakatti et al., 2013), whereas another two studies found that NMR profiles were not predictive of implantation (Rinaudo et al., 2012) or developmental arrest (Nadal-Desbarats et al., 2013). The study by Seli et al. (2008) was limited by a small size as only 18 patients were included. The study by Pudakalakatti et al. (2013) included 48 patients, but only reported analysis on the specific metabolites lactate, pyruvate and alanine. A drawback was that the study included a mixture of Day 2 and 3 collected media and that the analysis included cycles where not all of the transferred embryos were implanted. Consequently, the particular embryo that resulted in successful implantation could not be identified. The positive correlation between implantation and NMR profiling was questioned by Rinaudo et al. (2012) in a study of 108 patients. The media were, however, divided into five different sets of smaller size that differed either with regard to acquisition of NMR data or culture media. In conclusion, none of the previously conducted NMR studies have evaluated media from single transferred embryos, and the study populations from the two studies reporting a positive correlation between NMR profiles and implantation were poorly characterized. In contrast, our study was performed on spent culture media from single-transferred blastocysts and the large study cohort consisted of patients positively selected by good prognosis factors, which overcomes the limitations of previous studies. Although several different strategies were attempted, we found no relationship between spectral variations and pregnancy outcome and thus conclude that NMR profiling of spent culture media is not sufficiently powerful to predict pregnancy outcome in this cohort. In our interpretation, the failure of the PLS/PCA prediction is that patient/sample variation not related to pregnancy potential dominates the data analysis. If the human variation is large, which our analysis suggests is the case—even in a group of good prognosis patients such as the present study—a very large sample size would be needed for the variation to become background only, and thus allow the possible correlation with pregnancy appear.

While the non-invasive analysis of spent culture medium is appealing, as it poses no risk to the embryo, an obvious disadvantage is that no amplification of the sample material can be performed in contrast to invasive RNA and DNA studies. Therefore, the small sample volume constitutes a challenge common to all studies of preimplantation embryo metabolism. In this study, detailed spectra with distinct variation were obtained, which confirms that a useful metabolic profile is achievable even with the limited media volume. There are, however, some limitations that require attention. With a very small sample volume, even small variations in the amount of media and dilution may introduce a large uncertainty in the results. This uncertainty can be minimized by normalizing the spectra to, for example, integrals of the entire spectrum, as was done in the present study. This strategy gave the best grouping of data, but has the weakness that the method assumes that each sample contains equal amounts of metabolites. The difference that could have been disguised with our choice of normalization would have been an overall fall/rise in all metabolites. Based on previous finding in the literature, we did not consider this likely or a problem for this particular study, since our aim was to investigate potential differences between individual spectral peaks rather than evaluate overall metabolism in terms of general up or down-regulation of all metabolites.

A second challenge invariably linked with the small sample volume and the lack of sample amplification options is that small variations may not be detected or that detection requires highly sensitive equipment and high resolution. A likely explanation for the failure of NIR analysis to improve clinical outcome in the conducted RCTs was reported to be a variation in the analysis performed on the particular instrument used in the trials that was larger than the relevant biological variation evaluated. The instrument, a bench-top NIR spectrometer, was later retracted from the market, yet the question that remains unanswered is whether the method or the instrument was responsible for the lack of positive results. The NIR spectrometer is a relatively inexpensive, rapid and simple platform that can easily be operated by users with minimal training, which makes the method an obvious candidate for clinical application (Botros et al., 2008). In exploratory research studies, however, NMR remains the method of choice, since it has powerful identification capabilities. In order to achieve sufficient resolution when the metabolite concentration is low, the analysis requires a large instrument with a powerful magnet, and is time-consuming and costly. The variations reported in the present study were obtained using such an equipment at a highly specialized centre, and the variations would most likely not have been detected with a smaller instrument. In other words, several of the challenges facing the bench-top NIR instrument would apply to a bench-top NMR instrument as well.

The embryos were cultivated in a time-lapse incubator, which allowed for a concomitant analysis of the correlation between time-lapse parameters and pregnancy (Kirkegaard et al., 2013). Interestingly, a logistic regression analysis of early time-lapse parameters chosen by indication suggested that while a subset of the evaluated time-lapse parameters
predicted blastocyst development, the parameters did not predict pregnancy, which might be explained by the homogeneity of the population. As neither NMR nor time-lapse analysis resulted in a predictive model, a meaningful comparison is not possible.

A limitation of our study that potentially could affect the external validity of the study is that blastocysts were transferred on Day 6, as opposed to Day 5, which is standard in most clinics. Non-elective Day 6 transfer has been associated with lower pregnancy rates compared with Day 5 transfers (Dessolle et al., 2011), yet another study found no difference in pregnancy outcome when Day 6 was elective (Elgindy and Elseedeek, 2012). In the present study, Day 6 transfer was motivated by the intention of TE biopsy and thus applied on all embryos in the study. The effect, if any, of delayed transfer, therefore would not cause any bias. Only a small subset of embryos was biopsied, which might potentially introduce an effect modification. A stratified analysis showed no difference in pregnancy outcome between biopsied and non-biopsied embryos, which allowed us to include the biopsied embryos. The study was conducted as a prospective cohort study, where patients were recruited and media collected before the clinical outcome was known. Patients who became pregnant were generally younger, and had an embryo of higher quality transferred than patients who did not achieve a pregnancy (Table I). Both age and embryo quality are known confounders for clinical outcome. Had we been able to develop a model, it could be questioned whether the model was related to clinical outcome or the confounders. As no model was developed, the observed differences in age and embryo quality between the pregnant and non-pregnant merely illustrates the known influence of age and embryo quality on clinical outcome.

In conclusion, our results indicate that NMR profiling of spent culture media is not sufficiently powerful to predict pregnancy outcome. As NMR is presently the choice of method for spectroscopic analysis, our study therefore questions the predictive value of using variations in the metabolome for embryo selection. Improvements in spectroscopic techniques might challenge this view.

Supplementary data

Supplementary data are available at http://humrep.oxfordjournals.org/.

Acknowledgements

The authors wish to thank the clinical, paramedical and laboratory team of the Fertility Clinic, Aarhus University Hospital.

Authors’ roles

K.K., N.C.N., J.H. and H.J.I. designed the study. K.K., H.J.I. and J.J.H. were responsible for media collection and for conducting the clinical trial. A.S.P.S. and J.S.N. performed the NMR analysis and related statistics. All authors interpreted the findings. K.K. and A.S.P.S. wrote the first draft. All authors critically reviewed and approved the final version of the manuscript.

Funding

Funding for the present study was provided by Aarhus University, the Lippert Foundation, the Toyota Foundation, the Aase og Einar Danielsen Foundation. Research at the Fertility Clinic, Aarhus University Hospital is supported by an unrestricted grant from MSD and Ferring.

Conflict of interest

The authors declare no conflict of interest.

References


Elgindy E, Elseedeek MS. Day 5 expanded blastocysts transferred on same day have comparable outcome to those left for more extended culture and transferred on day 6. J Assist Reprod Genet 2012; 29:1111–1115.


