Title

1H NMR based profiling of spent culture media cannot predict success of implantation for day 3 human embryos

Journal

Journal of Assisted Reproduction and Genetics, 29(12)

ISSN

1058-0468

Authors

Rinaudo, Paolo
Shen, Shehua
Hua, Jia
et al.

Publication Date

2012-12-01

DOI

10.1007/s10815-012-9877-9

Peer reviewed

J Assist Reprod Genet (2012) 29:1435–1442
DOI 10.1007/s10815-012-9877-9

TECHNOLOGICAL INNOVATIONS

^{1}H NMR based profiling of spent culture media cannot predict success of implantation for day 3 human embryos

Paolo Rinaudo · Shehua Shen · Jia Hua · Su Qian · Uday Prabhu · Erwin Garcia · Marcelle Cedars · Dinesh Sukumaran · Thomas Szyperski · Chris Andrews

Received: 3 September 2012 / Accepted: 9 October 2012 / Published online: 23 October 2012

Abstract
Background Identification of a non-invasive technique to assess embryo implantation potential in assisted reproduction would greatly increase success rates and lead more efficiently to single embryo transfer. Early studies suggested metabonomic analysis of spent culture media could improve embryo selection. The goal of this study is to assess if embryo implantation can be predicted based on proton nuclear magnetic resonance (^{1}H NMR) profiles of spent embryo culture media from patients undergoing transfer of multiple embryos on cleday 3.

Method We conducted a retrospective study in an academic assisted reproduction technology (ART) program and analyzed the data in a university research center. Two hundred twenty-eight spent culture media samples originating from 108 patients were individually analyzed. Specifically, five distinct sets (1 to 5) of different types of spent media samples (volume ~14 μL) from embryos that resulted in clinical pregnancy (positive heart rate at 6 weeks gestation) (n_1=29; n_2=19; n_3=9; n_4=12; n_5=33; n_{total}=102) and from embryos that did not implant (n_1=28; n_2=29; n_3=18; n_4=15; n_5=36; n_{total}=126) were collected on day 3 of embryo growth. The

Capsule ^{1}HNMR media profiling does not predict embryo implantation.

Support: NICHD-R21: HD054956-01A2 (to PR, SS, and TS)

Electronic supplementary material The online version of this article (doi:10.1007/s10815-012-9877-9) contains supplementary material, which is available to authorized users.

P. Rinaudo (✉) · S. Shen · M. Cedars
Department of OB GYN and Reproductive Sciences,
University of California,
San Francisco, CA, USA
e-mail: rinaudop@obgyn.ucsf.edu

J. Hua · U. Prabhu · E. Garcia · D. Sukumaran · T. Szyperski (✉)
Departments of Chemistry and Biostatistics,
State University of New York at Buffalo,
Buffalo, NY, USA
e-mail: szypersk@buffalo.edu

S. Qian · C. Andrews
Departments of Biostatistics,
State University of New York at Buffalo,
Buffalo, NY, USA

Present Address:
S. Shen
Department of Clinical Research and Development, Auxogyn. Inc.,
Menlo Park, CA, USA

Present Address:
J. Hua
Fresenius Kabi,
Grand Island, NY 14072, USA

Present Address:
U. Prabhu
Bruker India,
Mumbai 400093, India

Present Address:
E. Garcia
LipoScience Inc.,
2500 Sumner Blvd.,
Raleigh, NC 27616, USA

Present Address:
C. Andrews
Department of OVS, Room 8320, University of Michigan,
1000 Wall St,
Ann Arbor, MI 48105, USA

media samples were profiled using ^{1}H NMR spectroscopy, and the NMR profiles of sets 1 to 5 were subject to standard uni- and multi-variate data analyses in order to evaluate potential correlation of profiles with implantation success.
Results For set 1 of the media samples, a borderline class separation of NMR profiles was obtained by use of principal component analysis (PCA) and logistic regression. This tentative class separation could not be repeated and validated in any of the other media sets 2 to 5.
Conclusions Despite the rigorous technical approach, ^{1}H NMR based profiling of spent culture media cannot predict success of implantation for day 3 human embryos.

Keywords ART · ^{1}H NMR · Metabonomics · IVF · Implantation success

Introduction

Wide-spread use of in vitro fertilization (IVF) has significantly increased the number of multiple births [12]. In fact, in the United States more than 30 % of pregnancies resulting from assisted reproduction technology (ART) are twins or higher-order multiple gestations; additionally, 51 % of all ART neonates are the products of multiple gestations, having a significant impact on individual/public healthcare and costs [34]. Multiple pregnancies are principally caused by the transfer of more than one embryo in a single cycle, given that monozygotic twinning occurs only in 1–2 % of embryos transferred [14].

A potential solution to this problem is to develop improved methods for embryo selection to permit single embryo transfer. One such method is to culture embryos to the blastocyst stage. Indeed blastocyst transfer has been associated with improved pregnancy rates [4]. However, not all patients qualify for blastocyst transfer and therefore cleavage embryo transfer remains common, in particular in patients with less than optimal prognosis, and may be preferred in some situations to prolonged in vitro culture [13]. Since it is well known the routinely used morphological criteria are not sufficient to identify implantable embryos, techniques that improve embryo selection at the cleavage stage, when multiple embryos are often transferred, would be particularly useful [22]. As a result, elective single embryo transfer (eSET), would be more widely utilized and greatly reduce the risk of multiple pregnancies.

Recently, several studies focused on developing novel non-invasive methods for predicting implantability and thus improved embryo selection. The measurement of amino acid concentrations, soluble human leukocyte antigen G expression, and oxygen consumption in culture media have all been reported to be associated with improved embryo implantation [16]. More recently the use of metabonomics of spent culture media has been studied by several groups. In particular, profiling using Raman, near-infrared spectroscopy (NIR) [25,28,32] or proton nuclear magnetic resonance spectroscopy (^{1}H NMR) [18,26] have been proposed to predict implantability. However, none of these techniques has yet been validated for clinical application [17,21] and two randomized control trials using Near Infrared (NIR) technology have recently failed to find a beneficial effect [11,31,33].

The primary goal of the paper was to assess if ^{1}H NMR profiling of day 3 spent culture media would enable prediction of implantation success in patients undergoing multiple embryo transfer. NMR technology has distinct advantage in analyzing metabolic profiles in small volume of specimens with high precision. Unlike Near Infrared (NIR) or Raman technology, which distinguishes unique molecule groups, NMR is able to define specific metabolites. Although the technology is not currently for real-time diagnosis due to its expensive and large-sized equipment, low throughput, and special expertise required, it is a great exploratory research tool.

A secondary goal was to analyze if different components of culture media (addition of different protein sources or different composition of the media) would alter the prediction potential, since media and protein supplement varies in different laboratories and commercially available media frequently change product compositions. An effective diagnostic metabonomics technology should be able to be applied to different types of culture media that are routinely used in the embryology laboratory.

Materials and methods

Patient selection, treatment and media sample collection

Patients were treated at the University of California, San Francisco (UCSF) Center for Reproductive Health. Institutional Review Board approval was obtained before the initiation of the study. All patients undergoing IVF were considered for participation in the study. However, only patients who were planning for a day 3 embryo transfer were enrolled in the study, given that day 3 embryo transfer is more often associated with high order multiple pregnancies [6,20].

Controlled ovarian stimulation was performed using gonadotropins and different protocols as previously published [23]. Briefly, follicular growth was induced with gonadotropins and egg retrieval was performed when two or more follicles had reached 18 mm or greater in maximal diameter. Oocytes were collected by transvaginal ultrasound guided needle aspiration under deep conscious sedation.

Conventional insemination or intra-cytoplasmic sperm injection (ICSI) was performed according to clinical indications. Normally fertilized oocytes, defined by visible two

pronuclei found between 16 to 19 h post insemination, were placed into individual 25 μL media droplets covered with Ovoil (VitroLife, Englewood, CO) for culture. For inclusion in the present study, patients had to have (i) a normal uterus; and (ii) undergo a non-traumatic embryo transfer procedure on cycle day 3. Patients characteristics are summarized in Table 1. Importantly, embryos were cultured in individual drops, so embryo morphology and pregnancy outcome could be readily assessed. The study was carried out in two stages. Results from preliminary data (set 1) prompted the expanded investigation with a more sensitive NMR probe and a wider variety of IVF conditions (sets 2 through 5) that represent the changing conditions in an IVF clinical setting. The overarching assumption that lead to the current study design was that the spent media of viable embryos (defined as embryos able to implant) would have a very distinct metabolomic profile compared to the spent profile of embryos that do not implant independently of any individual characteristics of the patients (including age) or stimulation protocols.

Table 1 Summary of patients' clinical characteristics. n refers to the number of patients

	0 % Implanted (n=54)	100 % implanted (n=54)	P value[a]
Age (years)	37.7 (4.0)	36.5 (5.6)	0.22
Cause of infertility			0.11
Unexplained	20	22	
Endometriosis	1	4	
DOR	18	6	
Male factor	7	6	
Uterine	1	1	
Tubal factor	0	2	
Ovulation dysfunction	1	2	
Combined females	4	5	
Combination male female	2	6	
Protocols			0.038
Long luteal	34	47	
Microdose flare	1	0	
Antagonist	6	2	
Demi halt	3	2	
E2 priming	10	3	
Number of prior attempts	0.41 (0.74)	0.20 (0.53)	0.10
Stimulation days	10.3 (1.9)	11.3 (3.0)	0.06
Ampules used (IU)	3562 (1523)	2843 (1455)	0.016
Number of eggs retrieved	11.2 (7.4)	11.0 (9.7)	0.92
Estrogen levels (pg/ml)	2374 (1083)	2969 (1254)	0.011
Fertilization method			0.90
ICSI	34	33	
Split ICSI	3	2	
IVF	17	19	
Number of 2PN	7.8 (4.9)	10.5 (6.2)	0.014
Average cell number	7.8 (1.1)	8.1 (0.6)	0.15
Average Fragmentation	1.87 (0.67)	1.68 (0.54)	0.10
# of Embryos Transferred	2.59 (0.86)	1.96 (0.43)	<0.001
Endometrial thickness (mm)	8.9 (2.5)	9.9 (1.3)	0.014

[a] p values from Student's t-test (continuous variables) or Fisher's exact test (categorical variables)

On culture day 3, embryos were graded based on conventional morphological criteria [19]. The number of embryos to transfer in each patient was based on age, prior fertility history and ASRM guidelines [1]. Overall, patients had an average transfer of 2 embryos (2.5 in the non-implanted group and 1.9 in the implanted group) The embryos identified for transfer were removed from the droplets with minimal volume of media and placed into another petri dish to be loaded into transfer catheter. The spent media (14 microliters) was collected into labeled Eppendorf tubes and immediately stored in −80 °C freezer. At the same time, a control media droplet, which had been incubated in the same culture dish as the spent media samples, was collected. The NMR signal of the control droplet was subtracted from the NMR signal of the droplet containing an embryo to control for potential changes in media composition present between different lots of media.

Each sample was labeled by its embryo's implantation status, which was determined by presence of fetal heart beat at 6+ weeks on ultrasound examination. The media of transferred embryos from patients for whom not a single implantation was successful were labeled as "0 % implantation". The media of transferred embryos from patients for whom all transferred embryos implanted were labeled as "100 % implantation". Only media samples from these patients for whom *all* embryos transferred either did or did not implant were considered for the present study. Media samples were shipped on dry ice to the State University of New York, Buffalo. Acquisition of the NMR spectra was performed blindly, with the investigators not knowing the implantation status of individual media drops.

A total of 228 spent media samples from 108 patients were collected in five different sets (Table 2). Results of Set 1 experiment (n=29 implanted embryos and n=28 non-implanted embryos) provided motivation for expanding the investigation. The five different sets reflected changes in the culture media implemented by the manufacturer. **Set 1** (57 samples from 26 subjects) contained $G_{1.3}$ media (VitroLife, Englewood, CO) samples supplemented with 5 % human serum albumin (HSA; Sage In-Vitro Fertilization, Cooper Surgical Company, Trumbull, CT, USA). **Set 2** (48 samples from 25 subjects) contained the same media and supplement, but NMR data were acquired differently (see below). **Set 3** (27 samples from 12 subjects) contained $G_{1.5}$ medium (VitroLife, Englewood, CO) supplemented with 5 % human serum albumin (HSA; Sage In-Vitro Fertilization, Cooper Surgical Company, Trumbull, CT, USA). **Set 4** (27 samples

Table 2 Summary of five Sets of samples including media and supplement used during incubation and equipment used for acquisition of NMR spectra. HSA: human serum albumin; SSS: serum supplement substitute. n refers to the number of spent culture media examined from embryos that implanted or did not implant

SET	Group	Media	Supplement	Spectrometer / Probe	Acquisition Time (hr)
1	Non-implanted n=28 Implanted n=29	$G_{1.3}$	HSA	600 MHz Microflow	354
2	Non-Implanted n=29 Implanted n=19	$G_{1.3}$	HSA	600 MHz Cryogenic	768
3	Non-implanted n=18 Implanted n=9	$G_{1.5}$	HSA	600 MHz Cryogenic	768
4	Non-implanted n=15 Implanted n=12	$G_{1.3}$	HSA	600 MHz Cryogenic	198
5	Non-implanted n=36 Implanted n=33	$G_{1.5}$	SSS	600 MHz Cryogenic	198

from 14 subjects) contained the same media and supplement as Set 1 and 2. **Set 5** (69 samples from 31 subjects) contained $G_{1.5}$ medium (VitroLife, Englewood, CO) supplemented with synthetic serum substitute (SSS; Irvine Scientific, Irvine, CA).

NMR data acquisition and processing

For set 1 of media samples, NMR data were acquired (total measurement time 354 h; Table 2) at 25 °C using a Varian Inova 600 spectrometer equipped with a PROTASIS® micro-flow probe connected to an auto-sampler as described previously [29], processed following established standard operating procedures [29], and high spectral quality was validated as described elsewhere [7].

For Sets 2 to 5, NMR samples were prepared by combining 14 μL of media sample with 28 μL of a D_2O stock solution containing the internal standard formate (1.0 mM) and NaCl (0.9 %w/v). The NMR sample was transferred to a capillary tube (New Era Enterprises, Inc. NJ, USA) with inner diameter 1.2 mm, which was then inserted into a regular NMR tube with 5 mm inner diameter containing 500 μl of D_2O for stable 'locking' of the spectrometer. NMR data were acquired (total measurement time 966 hours; Table 2) at 25 °C using a Varian Inova 600 spectrometer equipped with cryogenic probe. Spectra were processed and high quality was validated as described for the spectra for Set 1. 1D ^{1}H Carr-Purcell-Meiboom-Gill (CPMG) spectra was acquired for all sample sets, as we have done before [7]. NOESY spectra [7] acquisition was limited to Set 1, 2 and 3 given that no additional information was gained by this analysis after preliminary review. To alleviate the impact of variations of the complex ^{1}H NMR profile of SSS (Set 5), spectra acquired for control samples were subtracted from the spectra acquired for spent media samples. Representative 1D ^{1}H CPMG spectra is shown in Figure S1.

Statistical analysis

Spectral data integrity was investigated graphically without knowledge of outcome variable (implantation status). Spectra were normalized to unit integral and subdivided in small regions called bins (with 0.005 ppm resolution) to reduce effects arising from variations of total signal and signal positions [15]. Univariate analysis of spectra included logistic regression analysis to calculate the likelihood of implantation for each bin, followed by Bonferroni correction [15]. Bins with False Discovery Rate (fdr) of less than 0.05 were noted as "promising candidates" for biomarkers differentiating implanted from non-implanted embryos [3]. Similar analyses controlling for clinical covariates (e.g., egg age) were completed. Difference spectra (sample–control) obtained for set 5 were analyzed in the same way.

The primary multivariate analysis was a principal component analysis (PCA) of the spectra to reduce the dimensionality of the data followed by logistic regression to predict implantation status. Additionally we pursued supervised methods including k-nearest neighbors, and partial least squares discriminant analysis with and without a genetic algorithm for variable selection [30]. Predictive models were evaluated by the area under the Receiver Operating Characteristic (ROC) curve and by cross-validated predictive accuracy. Data sets could not be combined as culture media and NMR conditions varied. Cross-validation sets were chosen at the patient level rather than the sample level to ensure independence of training and test sets [2]. Difference spectra obtained for set 5 (i.e., samples containing SSS) by subtracting the spectrum of the control sample were analyzed in the same way. Statistical analyses were completed in the R (http://www.R-project.org) and MATLAB® (The Mathworks, Natick MA USA) computing environments.

Results

Patient characteristics are described in Table 1. Overall, patients whose embryos implanted were more often stimulated with a downregulated Lupron protocol (47/54, 87 % vs. 34/54, 63 %), had a thicker endometrial lining (9.9 vs. 8.9 mm), more 2PN embryos (10.5 vs. 7.8) but fewer embryos transferred (1.9 vs. 2.5) compared to patients whose embryos did not implant. In addition there was a tendency for longer stimulation (11.3 vs. 10.3 days, p=0.06). However, the two groups of patients were similar in all the remaining characteristics, and importantly they had a similar age (36.5 vs. 37.7) and their embryos had similar cell number (8.1 vs. 7.8) and fragmentation score (1.68 vs. 1.87).

Univariate and multivariate data analyses indicate that 1D ^{1}H NMR-based profiling of metabolites in culture media do not identify implantable embryos

Five datasets were used to investigate the research hypothesis that metabolic profiles of spent media culture differentiate implanted and non-implanted embryos. Univariate analysis of set 1 samples found the mean signal of one region of the NMR spectrum (bin 5.65) to be different (p=0.009 *post* Bonferroni correction) between implanted and non-implanted embryos. Similarly, multivariate analysis of set 1 samples provided a predictive model with area under ROC curve (AUC) of 0.75 (p=0.04). The numerical value of this area would be close to 1 if the prediction was excellent and close to 0.5 if it was poor. The model differentiated between implanted and non-implanted embryos (Tables 3 and 4, Fig. 1) with 69 % sensitivity (correctly identifying an implanted embryo) and 40 % specificity (correctly identifying a non-implanted embryo). Hence, although the class means in Set 1 could be distinguished, individual predictions were poor. Thus additional studies were performed using a spectrometer with higher sensitivity (i) to reproduce and validate the class separation, which is a key step in studies with multiple predictors, and (ii) to improve the ability to make accurate individual predictions.

Univariate analyses in Sets 2 through 5 identified, respectively, 0, 1, 32, and 0 *different* bins as candidate

Table 3 Summary of univariate analyses of Sets 1–5. Significance test applied to each bin and those with False Discovery Rate (fdr)<0.05 are noted in table. Note that any statistically significant result in one Set was not reproduced or validated in any other Set

Set	location of bins with fdr <0.05[a] (n)
1	5.650 (n=1)
2	NA (n=0)
3	7.200 (n=1)
4	3.825 3.790 3.625 3.610 3.565 3.520 3.515 3.495 3.460 3.435 3.355 3.350 2.770 2.530 2.350 2.040 3.340 3.320 3.305 3.290 3.285 3.280 3.275 3.265 3.140 3.110 3.020 2.995 2.970 2.915 2.880 2.835 (n=32)
5	NA (n=0)

[a] Chemical shift (ppm) with fdr<0.05; bins are 0.005 ppm wide

Table 4 Summary of multivariate analyses of Sets 1–5. Significance test and AUC of ROC curve for logistic regression model with first three principal components reported. The numerical value of the AUC would be close to 1 if the prediction was excellent and close to 0.5 if it was poor. Predictive models display inconsistent ability to differentiate implanted embryos and non-implanted embryos

Set	AUC (95 % CI)[a]	p-value[b]
1	0.75 (0.62, 0.88)	0.04
2	0.45 (0.29, 0.62)	0.753
3	0.73 (0.54, 0.92)	0.469
4	0.87 (0.74, 1.00)	0.002
5	0.56 (0.42, 0.69)	0.992

[a] Area Under the ROC Curve

[b] From likelihood ratio test statistic of logistic regression of implantation status on first 3 principal components

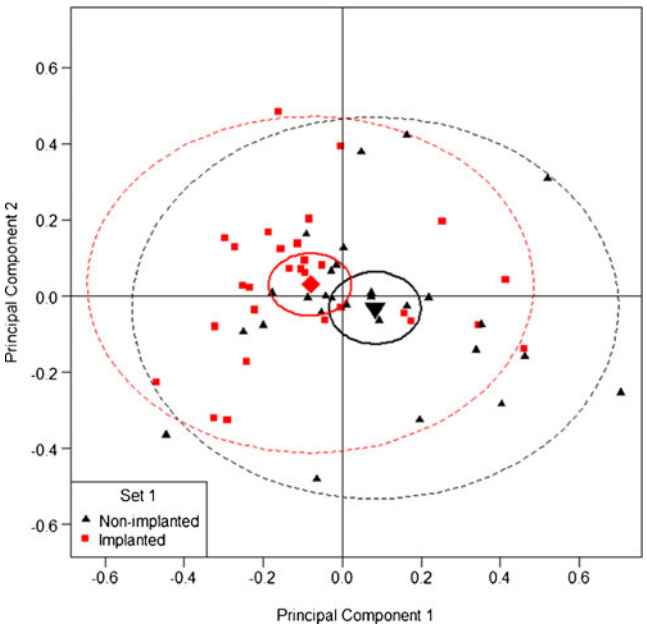

Fig. 1 Score plot of first two principal components derived from the 1D ^{1}H NMR spectra of spent culture media of Set 1 embryos. Principal components are linear combinations of variables that best summarize a large collection of variables in fewer dimensions. Non-implanted implantation embryos are shown as black triangles and cluster on the right (n=28); implanted embryos are shown as red squares and cluster on the left (n=29). The two group means are displayed by larger symbols. Solid ellipses represent 95 % confidence regions and are fairly distinct. The group means are statistically different in the direction of the first Principal Component (p=0.04). Dashed ellipses represent 95 % prediction regions and overlap substantially, indicating inadequate predictive ability

biomarkers to separate implanted versus non-implanted embryos. However, no bin was identified as promising (i.e., had fdr<0.05) in more than one Set. That is, Sets 2 through 5 did *not* validate Set 1 or each other (Table 3).

Multivariate analyses of Sets 2 through 5 did not identify good class separation with the exception of Set 4 (p=0.002, area under the curve=0.87, Table 4). However, Set 4 was the smallest of the five datasets (27 samples from 13 patients). In the remaining Sets (2, 3, and 5), implanted embryos did not have a different signal compared to non-implanted embryos (p>0.25, AUC near 0.5, Fig. 2). Additional statistical analysis using different methods of multivariate analyses (Partial least squares regression discriminant analysis with and without a genetic algorithm for variable selection; k-nearest neighbors; data not shown) confirmed these conclusions [24]. Analysis of difference spectra (i.e., subtracting the signal of the control sample) did not improve separation (not shown).

Discussion

In this work, we show that ^{1}H NMR spectra of spent culture media of day 3 embryos that implant (100 % implantation) are not consistently different from the spectra of embryos that fail to implant (0 % implantation). Specifically, the discrimination offered by NMR-based metabolomics could

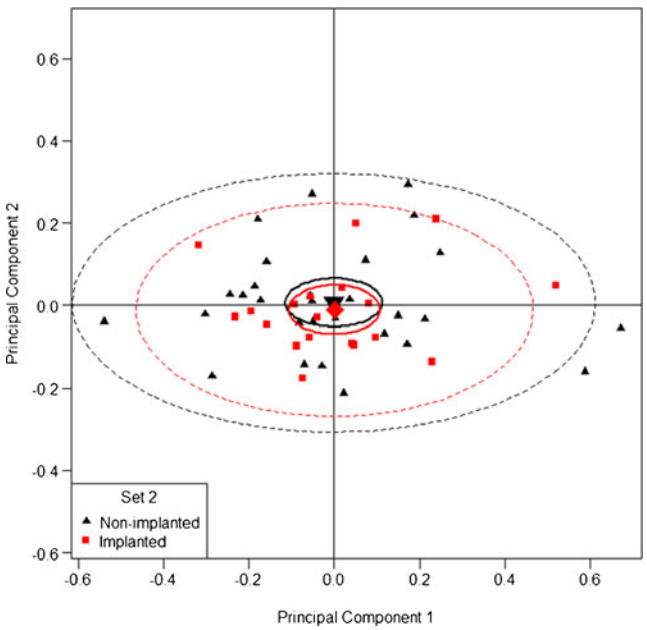

Fig. 2 Score plot of first two principal components derived from the 1D ^{1}H NMR spectra of Set 2 embryos. Non-implanted implantation embryos are shown as black triangles (n=29) and implanted embryos are shown as red squares (n=19). The two group means are displayed by larger symbols and are nearly equal. Solid ellipses represent 95 % confidence regions and overlap substantially (contrast to Fig. 1). Dashed ellipses represent 95 % prediction regions and similarly overlap

not increase the efficiency of embryo selection and thus IVF success rate.

Our study contrast several publications that showed a correlation between specific metabolic profiles of embryo spent culture media and successful implantation [18,25–28,31,33]; however, our findings are consistent with the results from two recent randomized controlled trials [11,31,33] showing no improvement in pregnancy outcome when media profiling was added to the conventional embryo selection protocol. In particular, our findings are contrary to the study of Seli et al, [26] and to the study of Pudakalakatti et al [18] where a positive correlation was found between implantation and NMR profiling. The reason for the difference is unclear. However, it is important to note that our study had a much larger sample size. In our study we included 108 patients: more specifically we included 106 embryos (from 54 patients) in the 100 % implantation and 140 embryos (from 54 patients) in the zero implantation group. The Seli study only included 18 patients: 17 embryos (from 10 patients) implanted and led to delivery (100 % sustained implantation), whereas 17 embryos (from 8 patients) did not implant (zero percent implantation). The Pudakalakatti study included 48 patients (and the profile of 43 implanted embryos and 57 not implanted embryos) and specifically analyzed only few metabolites: lactate, pyruvate and alanine. In addition, the authors used a "per patient" definition of successful implantation: because multiple embryos were transferred in the same patient, a group of embryo was defined as implanted if at least one embryo in its cohort was implanted successfully. The particular embryo that resulted in successful implantation could not be determined. However, for patients resulting in pregnancy (successful implantation), it was found that all embryos had a similar metabolic signatures in the NMR spectrum. This point is important and we believe it represent a critical limitation of the Pudakalakatti study, since our data analysis showed that even among sibling embryos that resulted in implantation the NMR signature can be very different.

Most importantly, none of the reported studies performed validation of their prediction model in an independent data set. Our study was specifically designed to validate results of the initial experiments in independent data sets. NMR data were acquired for five independent sets of media samples in order to make a most comprehensive attempt to validate the findings for the individual sets. While Sets 1, 2 and 4 (Table 2) were conducted on embryos cultured on the same type of media ($G_{1.3}$ and HSA), set 3 were performed on embryos cultured on $G_{1.5}$+HSA medium and set 5 on $G_{1.5}$+SSS medium. Overall, it was not possible to jointly subject NMR profiles from different sets to multivariate data analysis since the differences of the culture media composition and their supplements clearly dominate class separation. The inability to compare metabolomic profiles between media containing different components is

a key finding, since culture media manufacturers routinely change and improve culture media. Furthermore, different IVF laboratories might use different culture media; here we show that changes of media composition profoundly affect the metabolomic profile, even when a very rigid standard operating procedure in a single laboratory was adopted. A robust embryo predictor technology for routine use in a clinical setting should, however, not be influenced by these differences.

In addition, even when limiting analysis to embryos cultured in same media but in different sets (Sets 1, 2, 4), prediction model could not be confirmed in independent data sets (Tables 3 and 4).

Several studies have analyzed the differences between metabolites present in media from implantable versus non-implantable embryos [5,8,9]. For example, Conaghan showed that pyruvate uptake was lower, on average, in embryos that implant compared to embryos that did not implant [5]. However, pyruvate uptake in individual embryos differed widely, indicating that pyruvate uptake is a rather poor marker for embryo selection [5]. Pudakalakatti et al also analyzed the concentration of pyruvate in spent culture media; in addition, lactate and alanine were assessed in 57 embryos that implanted and 43 embryos that did not implant using NMR [18]. They found that embryos transferred on day 3 (after 72 h in vitro culture) with successful implantation exhibited significantly higher pyruvate uptake (not lower, as in the Conaghan study) and lower pyruvate/alanine ratios compared to those that failed to implant. Lactate levels in media were similar for all embryos. On the contrary, Seli et al, also using NMR, did not find any difference in pyruvate or pyruvate/alanine ratio but found that implanted embryos had higher glutamate and lower alanine levels, and a decrease in the alanine to lactate ratio [26]. In a more recent study Gardner et al analyzed the daily glucose consumption of post compaction embryo in spent culture media using microfluorimetry [9]. The author found that glucose consumption in embryo that implanted was significantly higher on both Day 4 and Day 5 than that by embryos which failed to develop post-transfer ($p<0.01$) and that glucose uptake was independent of embryo grade. In summary the above mentioned studies indicate that the mean (average) concentrations of several metabolites can be different in the two groups; however each study showed a large intra-group variability that does not support robust individual predictions. Importantly, none of the research group attempted to validate their findings in independent data sets. More concerning, the metabolites found to be altered in spent culture media of implanted embryo in different NMR studies are not consistent, indicating lack of reproducibility between different studies.

A potential limitation of our study is inclusion of single as well as multiple embryo transfer cases. Our experimental design was however modeled to reflect the more common practice of multiple embryo transfer prevalent in US [10]; moreover the studies published using NMR technology used a similar multiple embryo transfer design [18,26].

One additional consideration to ponder while analyzing our study and embryo selection and implantation studies in general, is the fact that the endometrium could play a role in allowing an embryo to implant. However, it is important to note that all patients in our study had evidence of normal endometrium defined as 1) lining of more set the than 7 mm and 2) had to have a normal saline sonogram. A potential confounding factor is that patients whose embryos implanted showed some evidence of better clinical prognosis (had higher estrogen levels on day of hCG, a thicker endometrial lining, more 2PN embryos and had less number of embryos transferred). However, other important clinical markers (age, fragmentation score or number of embryonic cells) were not different among the groups.

In conclusion, we have found no significant difference between ^{1}H NMR profiles of spent culture media of implanted versus non-implanted embryos while following a rigorous protocols and using one of the most sensitive technologies available today to study metabolite profiling. Although it might be possible that future advancements of NMR hardware may improve class separation and thus predictive accuracy of NMR-based profiling, our findings suggest that variations of embryo-dependent metabolite concentrations in media are too small to be of predictive value for embryo selection.

References

1. 2009 Guidelines on number of embryos transferred. Fertil Steril. 92 1518–9.
2. Adler W, Brenning A, Patapov S, Schmid M, Lausen B. Ensemble classification of paired data. Comput Stat Data Anal. 2011;55:1933–41.
3. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodological). 1995;57:289–300.
4. Blake DA, Farquhar CM, Johnson N, Proctor M. Cleavage stage versus blastocyst stage embryo transfer in assisted conception. Cochrane Database Syst Rev. 2007;CD002118.
5. Conaghan J, Hardy K, Handyside AH, Winston RM, Leese HJ. Selection criteria for human embryo transfer: a comparison of pyruvate uptake and morphology. J Assist Reprod Genet. 1993;10:21–30.
6. Frattarelli JL, Leondires MP, McKeeby JL, Miller BT, Segars JH. Blastocyst transfer decreases multiple pregnancy rates in in vitro fertilization cycles: a randomized controlled trial. Fertil Steril. 2003;79:228–30.
7. Garcia E, Andrews C, Hua J, Kim HL, Sukumaran DK, Szyperski T, Odunsi K. Diagnosis of early stage ovarian cancer by 1H NMR metabonomics of serum explored by use of a microflow NMR probe. J Proteome Res. 2011;10:1765–71.

8. Gardner DK, Lane M, Stevens J, Schoolcraft WB. Noninvasive assessment of human embryo nutrient consumption as a measure of developmental potential. Fertil Steril. 2001;76:1175–80.
9. Gardner DK, Wale PL, Collins R, Lane M. Glucose consumption of single post-compaction human embryos is predictive of embryo sex and live birth outcome. Hum Reprod. 2011;26:1981–6.
10. Grady R, Alavi N, Vale R, Khandwala M, McDonald SD. Elective single embryo transfer and perinatal outcomes: a systematic review and meta-analysis. Fertil Steril. 2012;97:324–31.
11. Hardarson T, Ahlstrom A, Rogberg L, Botros L, Hillensjo T, Westlander G, Sakkas D, Wikland M. Non-invasive metabolomic profiling of Day 2 and 5 embryo culture medium: a prospective randomized trial. Hum Reprod. 2012;27:89–96.
12. Janvier A, Spelke B, Barrington KJ. The epidemic of multiple gestations and neonatal intensive care unit use: the cost of irresponsibility. J Pediatr. 2011;159:409–13.
13. Kallen B, Finnstrom O, Lindam A, Nilsson E, Nygren KG, Olausson PO. Blastocyst versus cleavage stage transfer in in vitro fertilization: differences in neonatal outcome? Fertil Steril. 2010;94:1680–3.
14. Kawachiya S, Bodri D, Shimada N, Kato K, Takehara Y, Kato O. Blastocyst culture is associated with an elevated incidence of monozygotic twinning after single embryo transfer. Fertil Steril. 2011;95:2140–2.
15. Lindon JC, Nicholson JK, Holmes E. The Handbook of metabonomics and metabolomics. Amsterdam: Elsevier; 2007.
16. Mastenbroek S, van der Veen F, Aflatoonian A, Shapiro B, Bossuyt P, Repping S. Embryo selection in IVF. Hum Reprod. 2011;26:964–6.
17. Nel-Themaat L, Nagy ZP. A review of the promises and pitfalls of oocyte and embryo metabolomics. Placenta. 2011;32 Suppl 3: S257–63.
18. Pudakalakatti SM, Uppangala S, D'Souza F, Kalthur G, Kumar P, Adiga SK, Atreya HS. NMR studies of preimplantation embryo metabolism in human assisted reproductive techniques: a new biomarker for assessment of embryo implantation potential. NMR Biomed. 2012
19. Racowsky C, Vernon M, Mayer J, Ball GD, Behr B, Pomeroy KO, Wininger D, Gibbons W, Conaghan J, Stern JE. Standardization of grading embryo morphology. Fertil Steril. 2010;94:1152–3.
20. Reh A, Fino E, Krey L, Berkeley A, Noyes N, Grifo J. Optimizing embryo selection with day 5 transfer. Fertil Steril. 2010;93:609–15.
21. Revelli A, Delle Piane L, Casano S, Molinari E, Massobrio M, Rinaudo P. Follicular fluid content and oocyte quality: from single biochemical markers to metabolomics. Reprod Biol Endocrinol. 2009;7:40.
22. Rijnders PM, Jansen CA. The predictive value of day 3 embryo morphology regarding blastocyst formation, pregnancy and implantation rate after day 5 transfer following in-vitro fertilization or intracytoplasmic sperm injection. Hum Reprod. 1998;13:2869–73.
23. Rosen MP, Zamah AM, Shen S, Dobson AT, McCulloch CE, Rinaudo PF, Lamb JD, Cedars MI. The effect of follicular fluid hormones on oocyte recovery after ovarian stimulation: FSH level predicts oocyte recovery. Reprod Biol Endocrinol. 2009;7:35.
24. Schervish M. A review of multivariate analysis. Stat Sci. 1987;2:396–413.
25. Scott R, Seli E, Miller K, Sakkas D, Scott K, Burns DH. Noninvasive metabolomic profiling of human embryo culture media using Raman spectroscopy predicts embryonic reproductive potential: a prospective blinded pilot study. Fertil Steril. 2008;90:77–83.
26. Seli E, Botros L, Sakkas D, Burns DH. Noninvasive metabolomic profiling of embryo culture media using proton nuclear magnetic resonance correlates with reproductive potential of embryos in women undergoing in vitro fertilization. Fertil Steril. 2008; 90:2183–9.
27. Seli E, Bruce C, Botros L, Henson M, Roos P, Judge K, Hardarson T, Ahlstrom A, Harrison P, Henman M, Go K, Acevedo N, Siques J, Tucker M, Sakkas D. Receiver operating characteristic (ROC) analysis of day 5 morphology grading and metabolomic Viability Score on predicting implantation outcome. J Assist Reprod Genet. 2011;28:137–44.
28. Seli E, Sakkas D, Scott R, Kwok SC, Rosendahl SM, Burns DH. Noninvasive metabolomic profiling of embryo culture media using Raman and near-infrared spectroscopy correlates with reproductive potential of embryos in women undergoing in vitro fertilization. Fertil Steril. 2007;88:1350–7.
29. Sukumaran DK, Garcia E, Hua J, Tabaczynski W, Odunsi K, Andrews C, Szyperski T. Standard operating procedure for metabonomics studies of blood serum and plasma samples using a 1H-NMR micro-flow probe. Magn Reson Chem. 2009;47 Suppl 1:S81–5.
30. Varmuza K, Filzmoser P. Introduction to multivariate statistical analysis in chemometrics. Boca Raton: CRC Press; 2009.
31. Vergouw CG, Botros LL, Judge K, Henson M, Roos P, Hanna Kostelijk E, Schats R, Twisk JW, Hompes PG, Sakkas D, Lambalk CB. Non-invasive viability assessment of day-4 frozen-thawed human embryos using near infrared spectroscopy. Reprod Biomed Online. 2011;23:769–76.
32. Vergouw CG, Botros LL, Roos P, Lens JW, Schats R, Hompes PG, Burns DH, Lambalk CB. Metabolomic profiling by near-infrared spectroscopy as a tool to assess embryo viability: a novel, non-invasive method for embryo selection. Hum Reprod. 2008;23:1499–504.
33. Vergouw CG, Kieslinger DC, Kostelijk EH, Hompes PG, Schats R & Lambalk CB. Metabolomic profiling of culture media by near infrared spectroscopy as an adjunct to morphology for selection of a single day 3 embryo to transfer in Ivf: a double blind randomised trial. In *American Society of Reproductive Medicine O:7*, 2011.
34. Wright VC, Chang J, Jeng G, Macaluso M. Assisted reproductive technology surveillance–United States, 2005. MMWR Surveill Summ. 2008;57:1–23.