Facultad de Medicina
y Odontología

Departamento de Pediatría, Obstetricia y Ginecología
Programa de Doctorado: 3139 Medicina

Nuevas estrategias para estimar la calidad embrionaria y el éxito de las embrio-transferencias mediante la evaluación no invasiva y selección automática en sistemas de time-lapse.

Novel strategies to estimate embryo quality and embryo-transfer success through non-invasive evaluation and automatic selection in time-lapse systems.

Autora:
Raquel Del Gallego Bonilla

Directores:
Marcos Meseguer Escrivá
José Alejandro Remohí Giménez
Antonio Pellicer Martínez

2022, Valencia (España)

Departamento de Pediatría, Obstetricia y Ginecología

PROGRAMA DE DOCTORADO: 3139 MEDICINA

Nuevas estrategias para estimar la calidad embrionaria y el éxito de las embrio-transferencias mediante la evaluación no invasiva y selección automática en sistemas de time-lapse.

Trabajo de Tesis Doctoral

Realizado por Dña. Raquel Del Gallego Bonilla

Graduada en Biología, para optar al grado de doctora por la Universidad de Valencia

Directores:

Dr. D. Marcos Meseguer Escrivá

Dr. D. José Alejandro Remohí Giménez

Dr. D. Antonio Pellicer Martínez

Valencia, Noviembre 2022

D/Dña. **Marcos Meseguer Escrivá**, Doctor en Biología por la Universidad de Valencia y Supervisor Científico y Embriólogo Senior en el Laboratorio de Fecundación *in vitro* del Instituto Valenciano de Infertilidad (IVIRMA Global, Valencia, España)

CERTIFICA:

Que el trabajo titulado:

Nuevas estrategias para estimar la calidad embrionaria y el éxito de las embrio-transferencias mediante la evaluación no invasiva y selección automática en sistemas de time-lapse.

Ha sido realizado íntegramente por Dña. **Raquel Del Gallego Bonilla**, bajo mi dirección y supervisión. Dicho trabajo está concluido y reúne todos los requisitos para su presentación y defensa como Tesis Doctoral ante un tribunal.

Y para que así conste a los efectos oportunos, se expide la presente certificación, en Valencia (España) a 14 de Noviembre de 2022.

Fdo.

Marcos Meseguer Escrivá

NIF. 73776663-T

D/Dña. **José Alejandro Remohí Giménez**, Catedrático de pediatría, obstetricia y ginecología de la Universidad de Valencia y presidente de IVIRMA Global.

CERTIFICA:

Que el trabajo titulado:

Nuevas estrategias para estimar la calidad embrionaria y el éxito de las embrio-transferencias mediante la evaluación no invasiva y selección automática en sistemas de time-lapse.

Ha sido realizado íntegramente por Dña. **Raquel Del Gallego Bonilla**, bajo mi dirección y supervisión. Dicho trabajo está concluido y reúne todos los requisitos para su presentación y defensa como Tesis Doctoral ante un tribunal.

Y para que así conste a los efectos oportunos, se expide la presente certificación, en Valencia (España) a 14 de Noviembre de 2022.

Fdo.

José Alejandro Remohí Giménez

NIF. 24307995-P

D/Dña. **Antonio Pellicer Martínez**, Catedrático de pediatría, obstetricia y ginecología de la Universidad de Valencia y presidente de IVIRMA Global.

CERTIFICA:

Que el trabajo titulado:

Nuevas estrategias para estimar la calidad embrionaria y el éxito de las embrio-transferencias mediante la evaluación no invasiva y selección automática en sistemas de time-lapse.

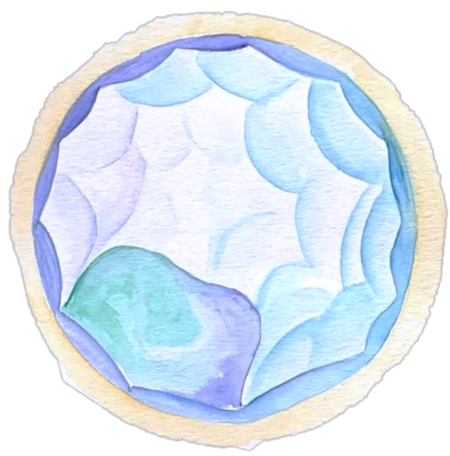
Ha sido realizado íntegramente por Dña. **Raquel Del Gallego Bonilla**, bajo mi dirección y supervisión. Dicho trabajo está concluido y reúne todos los requisitos para su presentación y defensa como Tesis Doctoral ante un tribunal.

Y para que así conste a los efectos oportunos, se expide la presente certificación, en Valencia (España) a 14 de Noviembre de 2022.

Fdo.

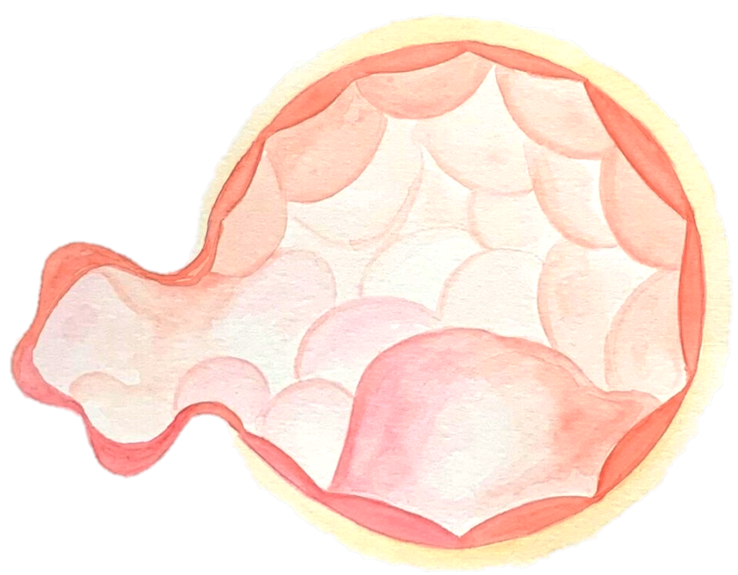
Antonio Pellicer Martínez

NIF. 19979620-A



"This extraordinary discrepancy between time on the clock
and time in the mind is less well known than it should be
and deserves fuller investigation."

Virginia Woolf



AGRADECIMIENTOS

ACKNOWLEDGEMENTS

AGRADECIMIENTOS *(ACKNOWLEDGEMENTS)*

Me gustaría empezar agradeciendo esta tesis a mis directores. Marcos, gracias por darme la oportunidad de comenzar esta aventura científica contigo. Me has abierto un mundo, literalmente, de posibilidades para poder dedicarme a lo que más me gusta, la investigación. Gracias a ti estoy donde estoy y esta tesis existe. A los doctores Pellicer y Remohí, los padres de la reproducción asistida. Sois toda una inspiración para cualquiera que se dedique a este campo. Es un honor teneros de directores.

Ahora, a los que quiero, mis agradecimientos serán cortos porque soy demasiado emotiva para ello.

Esta tesis es para ti, Abu. Tu ansia de conocimiento siempre inspiró mi curiosidad. Me he acordado mucho de ti leyendo hasta altas horas de la noche. Jamás podré explicar con palabras cuán agradecida estoy por todo lo que nos enseñaste, tu valor de la familia y tu sentido del humor.

A Madrina, tu alegría y cariño es contagioso. Gracias por ser tan valiente de escapar y trabajar tan duro en tu felicidad y tu familia, siempre con una sonrisa en la cara.

A mi Yaya, gracias por tu amor puro. Nos adoras y nos permites demasiado. Olvidas todos nuestros defectos y mantienes alta nuestra confianza. Gracias por animarnos siempre.

Esta tesis también es para ti, Mamá. Por todo lo que eres. Por crear un hogar basado en el amor. Por crear un mundo paralelo de diversión y comprensión al que siempre podemos acudir. Por apoyarnos incondicionalmente y enseñarnos todos nuestros valores. La vida son personas y tú eres mi estrella.

Para Papá. Gracias por darme la fortaleza para continuar esta tesis. Nunca has dejado de confiar en mí y mi capacidad para conseguir lo que me proponga. Gracias por siempre inspirarme a trabajar. No sería nadie sin tus consejos.

Para Adrián. Como siempre dice mamá, eres lo mejor de nuestra casa. Traes la cordura y estabilidad a nuestras locuras. Gracias por ser la voz de mi conciencia, mi sombra y mi mejor amigo. Y por supuesto, a mis otras hermanas, Daisy y Sera (las tenía que poner aquí).

عمرو، حياتي، بحبك. رسالة الدكتورا كانت مستحيلة من غيرك. آمنت فيي، حتّا لما أنا كنت رح أستسلم. أعّدتني المليون مرّة حتى أخلص الرسالة. شكراً انك أعطيتني درس الجهد و الصبر كل يوم. أنا كتير ممنونة انك لاقيتنى.

الى عائلتي الجديدة.
ميساء، قوّتكِ أن ترحلي لبلد تاني، تتعلّمي لغة جديدة و ترخصي حياتك لعناية الناس ملهمة كمحترفة و كامرأة، شكراً. و شكراً انك زرعتي في عمرو أكثر قلب حنون و طيب في العالم.
شبلي، شكراً انك دائماً بتنبع الفرح و الضحكة في العائلة. حس الفكاهة فيك حسّستني اني في بيتي و بين أهلي من أوّل يوم.
كريم، شكراً انك دائماً بتسمعلنا كلنا مع نكتة جاهزة. انت المستشار الرئيسي للعائلة، و انا بتشرّف و بحاجة اني أكون جزء منها.
تامر، جهدك و مثابرتك ملهمة. انت رح اتحقق كل أهدافك. شكراً انك شريكي في الجريمة بالمزح على عمرو.

Al resto de mi familia. Roberto, Pilar, Klara, Ariel, Zoel, Nacho, Laure+1, Alma, Alex, Sofi, Nacho, Leyre, Rafa, Marisol, Luis, Luisanti, Marta, Valeria, Cristina, Juanki, Lucía, Adrián, Conchi, Santi, Jimena y Rodrigo. Gracias por ser la mejor familia del mundo, os quiero muchísimo.

A Marta, mi mejor compañera en esta tesis. Este camino ha sido largo, pero tú has sido lo mejor de él. Me has enseñado que la amistad no entiende de kilómetros, ni 7,000 KM han podido con nosotras. Tu apoyo, chistes, memes y risas lo han sido todo. Gracias por estar siempre ahí.

A Miguel. Me has enseñado el significado del amor incondicional. Se que siempre voy a poder contar contigo, para animarnos el día, para aconsejarnos o para centrarnos y motivarnos a ser las personas más productivas del mundo. Estoy aquí para lo que necesites. Gracias por todo.

A mi familia de Dubái. Diego, Adrián, Aria, Fakhrou, Valeria+2, Thibaut, Raquel, Tomeu y Claudia. Gracias por siempre hacerme sentir en casa.

A todos los que se han cruzado en mi camino en estos años de tesis, gracias por vuestro apoyo:

A mis compañeros becarios precarios, ¡podéis con todo!

A mis amigas de Godella, gracias por siempre echarnos unas risas, aunque estemos lejos.

A mis amigos del cole, gracias porque no importa cuánto tiempo pase sin vernos, parece que no paso nada.

To my York friends, thank you for keeping our strong friendship from 4 different countries.

A mis amigas del pueblo, gracias por hacerme siempre un hueco en vuestras vidas.

A mis amigos del Lindy, gracias por alegrarme la vida.

Y, por último, gracias a aquellos que me inspiraron en mis primeros pasos científicos: María José Lorente, Pedro Luis Tineo, Miguel Ángel Silvestre y mi gran amiga, Sara Sadeghi.

Gracias a todos.

CERTIFICADO DEL CEIC IVI VALENCIA SOBRE EL ESTUDIO RETROSPECTIVO "VALIDACIÓN RETROSPECTIVA DEL HARDWARE GERI Y SU SOFTWARE DANA DE ANÁLISIS AUTOMATIZADO Y SELECCIÓN EMBRIONARIA. TECNOLOGÍA TIME-LAPSE." 1611-VLC-079-MM

Miguel Moreno Albiñana, Secretario del Comité Ético de Investigación Clínica de IVI Valencia,

CERTIFICA

I.- Que este Comité ha evaluado la propuesta del Promotor del Estudio Retrospectivo denominado:

- Título: "*Validación retrospectiva del Hardware GERI y su software DANA de análisis automatizado y selección embrionaria. Tecnología time-lapse*".
- Código del Protocolo del Promotor / Promotor: 1611-VLC-079-MM / IVI Valencia
- Investigador Principal: Dr. Marcos Meseguer Escrivá
- Fecha del Protocolo: enero 2017

II.- Que tomando en consideración:

- La pertinencia del Estudio, idoneidad de investigadores y colaboradores, idoneidad de instalaciones, teniendo en cuenta el conocimiento disponible, así como los requisitos de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal, de la Ley 14/2007, de 3 de julio, de Investigación Biomédica y la Ley 14/2006, de 26 de mayo, sobre Técnicas de Reproducción Asistida.

- Los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio, ausencia de riesgos y molestias para los sujetos, así como los beneficios esperados.

En su virtud, este Comité emite **DICTAMEN FAVORABLE** para la realización de dicho Estudio.

Lo que firmo en Valencia, a 6 de enero de 2017

FDO.: MIGUEL MORENO ALBIÑANA
SECRETARIO CEIC IVI VALENCIA

CERTIFICADO DEL CEIC IVI VALENCIA SOBRE LA ENMIENDA RELEVANTE 1 AL PROYECTO DE INVESTIGACIÓN 1308-C-127-FD "SECRETOMA DEL EMBRIÓN HUMANO COMO MARCADOR DE VIABILIDAD EMBRIONARIA E IMPLANTACIÓN."

D. Miguel Moreno Albiñana, Secretario del Comité Ético de Investigación Clínica de IVI Valencia,

CERTIFICA

I.- Que este Comité ha evaluado la propuesta del Promotor del Estudio denominado:

- Título: *"Secretoma del embrión humano como marcador de viabilidad embrionaria e implantación."*
- Código del Protocolo del Promotor / Promotor: 1308-C-127-FD / Fundación IVI – IVI Valencia
- Investigador Principal: Dr. Francisco Domínguez Hernández
- Versión / Fecha del Protocolo: Versión 2 / 7 de septiembre de 2016.
- Versión / Hoja de Información al Sujeto Consentimiento Informado: Versión 1 / 26 de septiembre de 2013.

II.- Que tomando en consideración:

La pertinencia del Estudio, idoneidad de investigadores y colaboradores, idoneidad de instalaciones, teniendo en cuenta el conocimiento disponible, así como los requisitos de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal, de la Ley 14/2007, de 3 de julio, de Investigación Biomédica, de la Ley 14/2006, de 26 de mayo, sobre Técnicas de Reproducción Asistida, el Real Decreto-ley 9/2014, de 4 de julio, por el que se establecen las normas de calidad y seguridad para la donación, la obtención, la evaluación, el procesamiento, la preservación, el almacenamiento y la distribución de células y tejidos humanos y se aprueban las normas de coordinación y funcionamiento para su uso en humanos; y las normas que las desarrollan.

Los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio, justificación de los riesgos y molestias previsibles para el sujeto, así como los beneficios esperados.

En su virtud, este Comité emite un **DICTAMEN FAVORABLE** a la citada Enmienda.

Lo que firmo en Valencia a de 26 de septiembre de 2016

FDO.: MIGUEL MORENO ALBIÑANA
SECRETARIO CEIC IVI VALENCIA

INDEX

INDEX

ABBREVIATIONS....... I

ABSTRACT (English)....... VII

RESUMEN (Castellano)....... XIII

INTRODUCTION....... 1

1. INFERTILITY....... 3

2. TIME-LAPSE SYSTEMS....... 9

2.1 Culture Conditions & Safety....... 9

2.2 Limitations of Conventional Embryo Scoring....... 11

2.3 Qualitative Parameters....... 13

3. EMBRYO SELECTION ALGORITHMS....... 15

3.1 Predictive Algorithms....... 15

3.1.1 Blastocyst Formation....... 15

3.1.2 Implantation Potential....... 16

3.1.3 Aneuploidy Risk....... 17

3.2 Embryo Selection Algorithms Concerns....... 19

3.2.1 Morphokinetic Parameters Variation....... 19

3.2.2 Universal Algorithm....... 21

4. AUTOMATION....... 25

5. ALTERNATIVE NON-INVASIVE TECHNIQUES....... 29

5.1 Introduction of Metabolomics....... 29

JUSTIFICATION....... 33

OBJECTIVES....... 39

MATERIAL & METHODS....... 45

1. IN VITRO FERTILIZATION STANDARD PROCEDURES....... 47

1.1 Controlled Ovarian Stimulation (COS) 47

1.2 Oocyte Pick-up & Fertilization Method....... 48

1.3 Time-lapse System Embryo Culture....... 48

1.4 Embryo Evaluation (Manual Morphokinetic Annotations) 49

1.5 Embryo Transfer & Pregnancy Status....... 49

2. AUTOMATED EMBRYO ASSESSMENT SOFTWARE.. 51
2.1 Software Development.. 51
2.2 Functionality... 52
3. OXIDATIVE STATUS OF THE SPENT CULTURE MEDIA... 57
STUDY PHASES... 63
SPECIFIC OBJECTIVE I. Characterization of the detection sensitivity & accuracy when performing the annotation of the human preimplantation embryo developmental events with an automated software.. 65
1.1 Purpose.. 65
1.2 Analysis.. 67
1.3 Results.. 69
1.4 Discussion... 77
SPECIFIC OBJECTIVE II. Clinical result prediction comparison when applying the manual and automated morphokinetic annotations of Known Implantation Data embryos........... 81
2.1 Purpose.. 81
2.2 Analysis.. 83
2.3 Results.. 85
2.4 Discussion... 93
SPECIFIC OBJECTIVE III. Non-invasive oxidative status analysis of the spent embryo culture medium in combination with Time-Lapse morphokinetics.. 99
3.1 Purpose.. 99
3.2 Analysis.. 101
3.2.1 Design.. 101
3.2.2 Oxidative Status Analysis... 101
3.2.3 Algorithm Development.. 101
3.2.4 Statistical Analysis.. 102
3.3 Results.. 105
3.3.1 Characteristics of Included Embryos.. 105
3.3.2 Oxidative Status Analysis... 106
3.3.3 Combined Embryo Selection Algorithm 106

3.4 Discussion…………………………………………………………………………………………………… 111
CONCLUSION……………………………………………………………………………………………………… 119
1. ACHIVEMENT OF OBJECTIVES & CLINICAL RELEVANCE……………………………………………… 121
2. FUTURE PERSPECTIVES…………………………………………………………………………………… 123
3. FINAL CONCLUSIONS (English)…………………………………………………………………………… 127
4. CONCLUSIONES FINALES (Castellano)…………………………………………………………………… 129
SUPPLEMENTAL MATERIAL…………………………………………………………………………………… 133
1. SUPPLEMENTAL FIGURES…………………………………………………………………………………… 135
2. SUPPLEMENTAL TABLES……………………………………………………………………………………… 139
3. SCIENTIFIC CONTRIBUTIONS………………………………………………………………………………… 141
3.1 Scientific Publications………………………………………………………………………………… 141
3.2 National or International Conferences……………………………………………………………… 142
3.3 Prices, Mentions and Distinctions…………………………………………………………………… 149
REFERENCES……………………………………………………………………………………………………… 153

FIGURES & TABLES INDEX

FIGURES INDEX

Introduction

Figure 1.
National Summary Report performed by the Society for Assisted Reproductive Technology (SART) describing the number of ART services, in the form of treatment cycles, deliveries and babies, in the US from 1985 up to 2019 (image from SART)………………………………………………… 5

Figure 2.
The Strengths of Computer Vision and Human Vision in Image Analysis (image from Danuser, 2011)…… 26

Material & Methods

Figure 3.
Description of the initial 3-phase validation of the software Dana…………………………………… 52

Figure 4.
(A) Pre-defined time ranges for each developmental event determining its inclusion in the annotations performed by Geri Assess 2.0. **(B)** Geri Assess 2.0 tab showing the automated annotations, including the out-of-range, in red with the warning sign (image from Geri Assess 2.0 Tech Note, 2019)…………………………………………………………………………………………………… 55

Figure 5.
(A) TCL device. **(B)** Sample preparation block and Analysis block. **(C)** Photochemical principles of thermochemiluminescence. **(D)** Example of a TCL kinetic curve. Parameters measured were amplitudes at 55 seconds, 155 seconds, and 255 seconds (**H1**, **H2** and **H3**, respectively)……. 59

Study Phases

Figure 6.
Patient Review page in Geri incubator. **(A)** Geri Assess 1.3 tab, to manually annotate events & observations. **(B)** Geri Assess 2.0, automatically annotates events & observations…………. 68

Figure 7.
Missing Data Analysis described as non-annotated events (NA). **(A)** Missing Data by stage and method represents the amount of non-annotated events per stage. **(B)** Missing data by Link represents the amount of non-annotated events per embryo analyzed……………………………… 70

Figure 8.
(A) Data distribution of the annotations performed for time to 6 cells (t6) by the IVIRMA embryologists' team and the Geri Assess 2.0 before and after applying the filtering tool for

the out-of-range annotations. **(B)** Superposed data distribution for both annotation techniques before filtering. **(C)** Superposed data distribution for Geri Assess 2.0 annotation before and after filtering.. 73

Figure 9.
Standard deviation calculated between the manual and automated annotations at the different embryo developmental stages... 75

Figure 10.
Summary of Specific Objective II.. 81

Figure 11.
Hierarchical classification of embryos based on: (i) Morphological screening; (ii) the new morphological criteria; (iii) timing of cell division to three cells (t3); (iv) duration of second cell cycle (cc2), i.e., the time from division to a two blastomere until division to a three blastomere embryo; (v) timing of cell division to five cells (t5). The classification generates 10 categories of embryos with increasing expected implantation potential (right to left) and almost equal number of embryos in each (image from Basile et al., 2015)... 85

Figure 12.
Embryo grading classification after using the morphokinetic manual annotations performed by the embryologist team and the automated ones performed by the software.................... 86

Figure 13.
Classification of embryos according to Rienzi et al. (2019) publication, based on the trophectoderm (TE) quality at time of blastulation (tB) and the time to morula (tM) (image from *Rienzi* et al.*, 2019)*.. 88

Figure 14.
Blastocyst classification using Rienzi et al.'s (2019) tM cut-off value with both, the embryologist's manual, and the software's automated annotations....................................... 89

Figure 15.
Blastocyst clinical rate prediction using Rienzi et al.'s (2019) tM cut-off value with both, the embryologist's manual, and the software's automated annotations....................................... 90

Figure 16.
Compaction in the human embryo. After several cell divisions (a–e), the blastomeres became flattened (f), and the intercellular boundaries became obscured (g–i), until they finally unified in one cluster (j,k). These morphological changes are called "compaction", and blastulation occurs only after complete compaction of the embryo (l) (image from Iwata et al., 2014).... 96

Figure 17.
(A) Thermochemiluminescence (TCL) parameters (H1sm, H2sm, H3sm, AVE-Hsm, Ratiosm) in relation to embryo outcome. **(B)** TCL parameters (H1sm, H2sm, H3sm, AVE-Hsm, Ratiosm) in relation to implantation.. 107

Figure 18.
Hierarchical classification of embryos based on morphokinetic screening parameters (tEB and t8–t5) (Motato et al., 2016) and thermochemiluminescence (TCL) parameter H2sm. The classification generates seven categories of embryos with differential implantation rates.. 108

Conclusions

Figure 19.
Factors affecting *in vitro* embryo development... 125

SUPPLEMENTAL FIGURES INDEX

Material & Methods

Supplemental Figure 1.
ASEBIR criteria (2015) for morphological assessment of blastocysts on day 5, between 112 and 120 hours………………………………………………………………………………… 135

Supplementary Figure 2.
Geri Assess 2.0 Timeline Bar with annotated developmental events and observations per embryo. For fragmentation observations, as it can occur numerous times or during a continuous period of time, a blue bar is placed on the timeline to reflect the event's presence………………………………………………………………………………… 136

Study Phases

Supplemental Figure 3.
Data distribution of the annotations performed, for all embryo developmental events analyzed, by the IVIRMA embryologist team and the Geri Assess 2.0 before and after applying the filtering tool for the out-of-range annotations……………………………………… 137

Supplemental Figure 4.
Standard deviation between the manual and automated annotations with the filtered and unfiltered dataset………………………………………………………………………………… 138

TABLES INDEX

Introduction

Table 1.
Morphokinetic Parameters.. 12

Table 2.
Published Embryo Selection Algorithms (ESAs).. 14

Table 3.
Inter-observer and intra-observer agreement for assessed parameters (image from Sundvall et al., 2013).. 20

Material & Methods

Table 4.
Summary of the key developmental events and observations annotated by Geri Assess 2.0 (Observations marked with *) *(Geri Assess 2.0 Tech Note, 2019)*... 53

Study Phases

Table 5.
Detection Rate Analysis of the morphokinetic developmental events by the Geri Assess 2.0 vs. a busy embryologist team in their routine clinical practice.. 71

Table 6.
Summary of filtered out-of-range annotations by embryo developmental event................... 72

Table 7.
Accuracy of the manual and automated annotations at the different developmental events, with the unfiltered and filtered dataset... 74

Table 8.
Embryo Selection Algorithms (ESAs) with application at the blastocyst stage......................... 87

Table 9.
Patient demographic characteristics... 105

Table 10.
Implantation success rates using our combined TCL and morphokinetic algorithm compared with implantation rates using the ASEBIR morphologic classification.................................... 109

SUPPLEMENTAL TABLES INDEX

Material & Methods

Supplemental Table 1.
ASEBIR criteria (2015) for morphological assessment of embryos on days 2-4.................... 139

Supplemental Table 2.
Time parameters assessed during TLS *(Motato* et al.*, 2016)*.. 140

ABBREVIATIONS

ABBREVIATIONS

aCGH	array-Comparative Genome Hybridization
AI	Artificial Intelligence
ART	Assisted Reproductive Technology
ASEBIR	Asociación para el Estudio de la Biología de la Reproducción
ASRM	American Society of Reproductive Medicine
AUC	Area Under the Curve
AVE	Average Media of points H1, H2 and H3
β-hCG	β - human Chorionic Gonadotropin
BE	Blastocyst Expansion
BEi	Initiating Blastocyst Expansion
BH	Hatching Blastocyst
BHi	Initiating Blastocyst Hatching
cc1	Formula = t2-t0; duration of the first round of cleavage; equal to t2
cc2	Formula = t3-t2; duration of the second round of cleavage
cc3	Formula = t5-t3; duration of the third round of cleavage
cc4	Formula = t9-t5; duration of the fourth round of cleavage
CCM	Continuous Culture Media
CI	Confidence Interval
CNN	Convolutional Neural Network
COS	Controlled Ovarian Stimulation
CPR	Clinical Pregnancy Rate
cps	Photons emitted per second
DIS	Discarded Embryos
dn	Dark noise
EES	Electronically Excited Species
EEVA	Early Embryo Viability Assessment
EIM	European IVF-monitoring Consortium
ESA	Embryo Selection Algorithm
ESHRE	European Society of Human Reproduction and Embryology
GnRH	Gonadotropin-Releasing Hormone
H1	TCL photon reading measurement after the first 55 seconds
H2	TCL photon reading measurement after the first 155 seconds
H3	TCL photon reading measurement after the first 255 seconds
ICC	Intra-Class Correlation Coefficient
ICM	Inner Cell Mass
ICSI	Intracytoplasmic Sperm Injection
IVI	Instituto Valenciano de Infertilidad
IVF	*In Vitro* Fertilization
IR	Implantation Rate
KID	Known Implantation Data
LBR	Live Birth Rate
NA	Non-Annotated Events
niPGT	Non-invasive Preimplantation Genetic Testing

NIRS	Near-Infrared Spectroscopic
OPR	Ongoing Pregnancy Rate
OPU	Ovarian Pick-Up
OR	Odds Ratio
P1	Duration of the first cytokinesis
P2	Formula = t3-t2; duration of the second round of cleavage
P3	Formula = t4-t3; duration of the second synchronization parameter
PGT-A	Preimplantation Genetic Testing for Aneuploidy
PGT-M	Preimplantation Genetic Testing for Monogenic (single gene disorders)
PGT-SR	Preimplantation Genetic Testing for chromosome structural rearrangements
PN	Pronuclei
RCT	Randomized Controlled Trial
ROC	Receiver Operating Characteristic Curve
ROS	Reactive Oxygen Species
s1	Formula = t2-t2; duration of the first synchronization parameter; equal to 0
s2	Formula = t4-t3; duration of the second synchronization parameter
s3	Formula = t8-t5; duration of the third synchronization parameter
SART	Society for Assisted Reproductive Technology
SET	Single Embryo Transfer
sm	Smoothing algorithm applied for photomultiplier fluctuations in TCL
SPSS	Statistical Package for the Social Sciences
T	Transferred embryos
t0	Time of insemination
TAC	Total Antioxidant Capacity
tB	Time to Blastocyst formation
TCL	Thermochemiluminescence
TE	Trophectoderm
tEB	Time to Expanded Blastocyst
tHB	Time to hatched Blastocyst
tHBi	Time to start of Hatching
tM	Time to the Morula Stage
tn	Time to n cell division
tPB2	Time to appearance of the second polar body
tPNa	Time to pronuclei appearance
tPNf	Time to pronuclei fading
tSB	Time to Start of Blastulation
TLS	Time-Lapse System
UAD	Units of Average Distance
US	United States
VIT	Vitrified Embryos
WHO	World Health Organization

ABSTRACT

// ABSTRACT

The introduction of time-lapse imaging to clinical in vitro fertilization practice enabled the undisturbed monitoring of embryos throughout the entire culture period. Initially, the main objective was to achieve a better embryo development. However, this technology also provided an insight into the novel concept of morphokinetics, parameters regarding embryo cell dynamics. The vast amount of data obtained defined the optimal ranges in the cell-cycle lengths at different stages of embryo development, which added valuable information to embryo assessment prior to transfer. Kinetic markers became part of embryo evaluation strategies with the potential to increase the chances of clinical success. However, the annotation of these parameters still depend on the subjectivity and experience of the professionals performing the annotations. The use of deep learning algorithms to analyze developmental events automatically is a step towards implementation of Artificial Intelligence into embryo assessment, which is becoming a significant trend in the future.

The present thesis major research target is the optimization of embryo selection though non-invasive protocols, aiming for the standardization of single embryo transfer as gold standard. Special focus being given to the validation of an Artificial Intelligence developed tool for the automatization of the morphokinetic parameter annotation and assessment of embryo metabolomics through the analysis of oxidative stress in the spent culture media. For this, three main specific objectives were stablished:

Specific objective I. Characterization of the detection sensitivity & accuracy when performing the annotation of the human preimplantation embryo developmental events with an automated software.

Specific objective II. Clinical result prediction comparison when applying the manual and automated morphokinetic annotations of known implantation data embryos.

Specific objective III. Non-invasive oxidative status analysis of the spent embryo culture medium in combination with Time-Lapse morphokinetics.

Thanks to our specific objective attainment, we can conclude the embryo assessment software can be an automated and objective alternative of annotating the morphokinetic parameters. Automated annotations would ease the embryologists' workload, especially during early events, which can be annotated with high accuracy. On the other hand, the more variable later events showed more clinical relevance in outcome prediction when using published embryo selection algorithms. Non-annotated events can still be annotated manually, increasing the accuracy of the data.

The thermochemiluminescence (TCL) assay was successfully validated as a non-invasive tool to perform the analysis of the oxidative stress of the spent culture media of the embryo. The combination of the TCL oxidative parameters with the time-lapse morphokinetic criteria presented a greater discriminatory power than morphological assessment in the identification of high-quality embryos, providing the foundation for a more objective and non-invasive embryo selection method to reduce multiple embryo transfers.

RESUMEN

RESUMEN

La introducción de la tecnología de lapso de tiempo en la práctica clínica de fecundación *in vitro* permitió la supervisión ininterrumpida de los embriones durante todo el período de cultivo. Inicialmente, el objetivo principal era lograr un mejor desarrollo embrionario. Sin embargo, esta tecnología también introdujo el novedoso concepto de la morfocinética, parámetros relacionados con la dinámica de las células embrionarias. La gran cantidad de datos obtenidos permitió definir los rangos óptimos de la duración del ciclo celular en diferentes etapas del desarrollo del embrión, lo cual añadió información complementaria para asistir en la evaluación del embrión antes de la transferencia. Los marcadores cinéticos se convirtieron en parte de la estrategia de evaluación embrionaria con un gran potencial en aumentar las probabilidades de éxito clínico. Sin embargo, la anotación de estos parámetros aún depende de la subjetividad y experiencia de los profesionales que realizan las anotaciones. El uso de algoritmos de aprendizaje profundo para analizar eventos de desarrollo automáticamente es un paso hacia la implementación de Inteligencia Artificial en la evaluación de embriones, que se está convirtiendo en una tendencia importante en el futuro.

El objetivo principal de la presente tesis es la optimización de la selección de embriones a través de protocolos no invasivos, con el objetivo de estandarizar la transferencia de un solo embrión en la práctica clínica. Se presta especial atención a la validación de una herramienta desarrollada mediante técnicas de Inteligencia Artificial para la automatización de la anotación de parámetros morfocinéticos y la evaluación de la metabolómica embrionaria mediante el análisis del estrés oxidativo en los medios de cultivo. Para ello, se establecieron tres objetivos específicos principales:

Objetivo específico I. Caracterización de la sensibilidad y precisión de detección de los eventos de desarrollo del embrión humano preimplantacional mediante un software automatizado.

Objetivo específico II. Comparación de la predicción de los resultados clínicos al aplicar las anotaciones morfocinéticas manuales y automatizadas de embriones con implantación conocida.

Objetivo específico III. Análisis no invasivo del estado oxidativo del medio de cultivo de los embriones en combinación con la morfocinética de sistemas de lapso de tiempo.

Gracias a la consecución de objetivos específicos, podemos concluir que el software de evaluación de embriones puede ser una alternativa automatizada y objetiva de anotar los parámetros morfocinéticos. Las anotaciones automatizadas aliviarían la carga de trabajo de los embriólogos, especialmente durante los eventos tempranos, los cuales se anotaron con gran precisión. Por otro lado, los eventos tardíos, pese a ser más variables, mostraron una mayor relevancia clínica en la predicción de resultados cuando se utilizaron algoritmos de selección publicados. Los eventos no anotados siempre pueden añadirse manualmente, lo cual aumenta la precisión de los datos.

El ensayo de termoquimioluminiscencia (TCL) se validó con éxito como una herramienta no invasiva para realizar el análisis del estrés oxidativo de los medios de cultivo del embrión. La combinación de los parámetros oxidativos de TCL con los criterios morfocinéticos de lapso de tiempo presentaron un mayor poder discriminatorio que la evaluación morfológica en la identificación de embriones de alta calidad, sentando las bases para un método de selección de embriones más objetivo y no invasivo para reducir las transferencias múltiples de embriones.

1

INTRODUCTION

INTRODUCTION

1. INFERTILITY

Human Infertility is defined by the American Society for Reproductive Medicine (ASRM) as:

> *"…the result of a disease (an interruption, cessation, or disorder of body functions, systems, or organs) of the male or female reproductive tract which prevents the conception of a child or the ability to carry a pregnancy to delivery. The duration of unprotected intercourse with failure to conceive should be about 12 months before an infertility evaluation is undertaken, unless medical history, age, or physical findings dictate earlier evaluation and treatment."*

The World Health Organization (WHO) estimation suggests that between 48 million couples and 186 million individuals live with infertility globally (Boivin *et al.*, 2007; Mascarenhas *et al.*, 2012; World Health Organization, 2018). This disorder can be considered as primary, in individuals or couples that have never achieved a pregnancy, and secondary, the incapacity to have a new pregnancy after a previous liveborn. A different number of factors, alone or combined, can constitute the etiology for infertility, which can be summarized as male, female, mixed or unexplained.

The most common causes associated to the male reproductive system are dysfunctionalities in the ejection of semen, hormonal disorders, testicular failure in sperm production and abnormal sperm function and quality. On the other hand, infertility in the female reproductive system is caused by any abnormalities present in the ovaries, uterus, fallopian tubes and/or endocrine system, among others (World Health Organization, 2018). The prevalence of these etiologies can considerably differ between regions and age groups,

and can be influenced by family history, environmental and lifestyle factors, and possible hazardous exposures (Rutstein and Shah, 2004; Gore *et al.*, 2015; Segal and Giudice, 2019).

In all cases, it is considered an essential human right to decide to have children, therefore addressing infertility is crucial assisting individuals or couples to found a family (Zegers-Hochschild *et al.*, 2013).

According to the 22nd annual report of the European IVF-monitoring Consortium (EIM) for the European Society of Human Reproduction and Embryology (ESHRE), including 39 participating European countries, the number of Assisted Reproductive Technology (ART) services addressing infertility has exponentially increased reaching over 1 million treatment cycles and over 200,000 babies born in the year 2018 (Wyns *et al*., 2022). **Figure 1** displays the same increasing trendline in the United States (US) in the past 34 years.

The knowledge and experience acquired in ART protocols and techniques since the first baby was born, back in 1978, has been reflected in the increase of its success rates. The US Society for Assisted Reproductive Technology (SART) observed in its National Summary Report of 2019 that patients undergoing a cycle with egg retrieval had a 28.1% chance of having a live birth. On the other hand, when using donated oocytes or embryos their chance increased up to a 43.4%.

Although increasingly higher success rates have been achieved, ART is still considered a young and imperfect field with several potential pathways with room for improvement. As any interdisciplinary field, ART, comprising gynaecology, obstetrics, endocrinology, immunology, urology, andrology, embryology, microbiology, and genetics, requires continuous learning and development towards a state-of-the-art medicine.

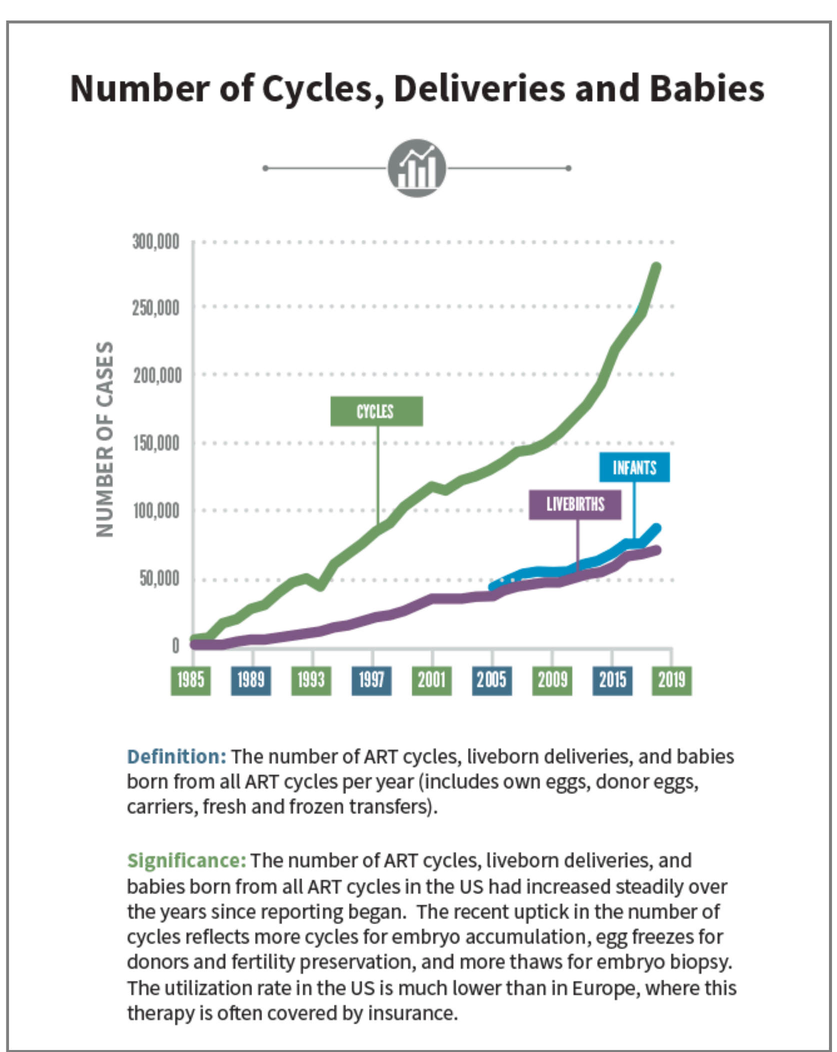


Figure 1. *National Summary Report performed by the Society for Assisted Reproductive Technology (SART) describing the number of ART services, in the form of treatment cycles, deliveries and babies, in the US from 1985 up to 2019 (image from SART).*

The below outlines potential areas that could be further improved in ART:

- Controlled Ovarian Stimulation (COS) Protocols & Ovarian Pick-Up (OPU) technique:

Optimization of stimulation protocols for oocyte growth, maturation, and collection; aiming to obtain a higher number and better-quality oocytes.

- Urology & Andrology:

Optimization of sperm production stimulation protocols, collection, maintenance, and selection techniques.

- Laboratory Techniques and Oocyte/Embryo Handling:

Optimization of oocyte denudation, gamete fertilization methods (*in vitro* fertilization, IVF; or intracytoplasmic sperm injection, ICSI), vitrification & thawing, embryo biopsy, and tubing.

- Gamete & Embryo Culture:

Optimization of gamete & embryo culture environment, mimicking *in vivo* conditions, to increase sperm, oocyte & embryo quality.

- Embryo Transfer, Fresh and Frozen Endometrial Protocols:

Optimization of endometrial preparation protocols, in both ovulatory and programmed cycles, for a suitable environment for embryo implantation.

- Embryo Selection:

Optimization of embryo selection methods, ideally non-invasive procedures, to select, within a cohort, the embryo with the highest implantation potential, or viability. Promote the implementation of single embryo transfers (SET) as a gold standard to avoid the risks of multiple pregnancy for the mother and fetus.

.

- Pre-Implantation Genetic Testing:

Optimization of genetic testing protocols, ideally towards non-invasive protocols, for aneuploidies (PGT-A), monogenic/single gene disorders (PGT-M) or chromosome structural rearrangements (PGT-SR).

A proper assessment of embryo viability and quality plays a critical role in determining the success of an IVF treatment. Current research mainly focuses on improving implantation rates (IRs), decreasing miscarriage rates, and improving the detection of aneuploid embryos. Implantation success depends to the same extent on both the endometrium's optimal development and the embryo's quality (Weimar *et al.*, 2013). However, current embryo selection techniques are not reliable enough leading to an inaccurate embryo ranking. As a direct consequence, multiple embryos have to be transferred to increase the probability of success, without considering the complications that may come from multiple pregnancies (Aparicio-Ruiz *et al.*, 2018). As single embryo transfer is the current desired standard, the need for optimization of embryo selection techniques increases. Identifying the variables determining which embryo, among a cohort, has the greatest potential to implant, in particular when there is a high number of good-quality ones, is essential to improve IVF techniques.

The present thesis major research target is the optimization of embryo selection through non-invasive protocols, aiming for the standardization of SET as gold standard. Special focus being given to embryo development, through the analysis of their physiological and metabolic mechanisms.

2. TIME-LAPSE SYSTEMS

The introduction of non-invasive time-lapse cinematography enabled us to track the timing of events and measure cell-cycle lengths at different stages in order to better understand embryo development. Moreover, it has introduced us to the possibility of identifying irregular cleavages. The vast amount of morphokinetic data collected, as a result, has been proposed in the form of embryo selection algorithms (ESAs) as possible predictors of IVF treatment outcomes. However, non has been sufficiently consolidated. Although numerous papers have pointed out the benefits of using time-lapse systems (TLSs) for embryo culture and evaluation, some publications have discussed their value due to their unproven clinical benefits (Kaser and Racowsky, 2014; Bhide *et al.*, 2017; Racowsky and Martins, 2017; Paulson *et al.*, 2018). An official publication from the ESHRE on recommendations provided on the use of time-lapse technology in the IVF laboratory was recently published (ESHRE Working group on Time-lapse technology *et al.*, 2020). The below focuses on the state-of-the-art, concerns, and possible future prospects in time-lapse technology.

2.1 Culture Conditions & Safety

"Are culture conditions improved by avoiding embryo manipulation outside the incubator?" (Meseguer, 2016).

The application of time-lapse technology in IVF laboratories has spread rapidly due to the non-invasive evaluation of embryos and the precise monitoring of optimal culture conditions. The evaluation of embryos through the analysis and annotation of the time-lapse video avoids the exposure of the embryo to changes in the environmental conditions, hence, comes closer to matching *in vivo* conditions. Hypothetically, this should lead to a better embryo development, but is this really happening?

Since 2008, a wide range of TLSs have become commercially available to IVF clinics. Current devices can be differentiated into two categories depending on whether the microscopic components need to be introduced in a conventional incubator, like Primo Vision (Vitrolife), or are integrated in the system, like Embryoscope (Vitrolife) and Geri (Genea Biomedx) (with their respective improved Plus versions), Miri TL (Esco Medical) and Astec CCM-IBIS (Astec). Other relevant differences between the devices need to be acknowledged: unlike the rest, the Eeva test (Early Embryo Viability Assessment; Merck-Serono), currently introduced in the Geri Plus version, uses dark field optics and has the ability to switch from bright to dark when necessary. The culture dish capacity of all of them ranges from 12–16 embryos. Furthermore, Geri Plus and Miri TL differ from others by using individualized chambers for single patient. Out of all devices, only Geri Plus and box incubators are able to control humidity conditions. Each individual TLS is extensively explained by Tejera *et al.* (2017) and (ESHRE Working group on Time-lapse technology *et al.*, 2020).

Continuing with the culture performance of the TLSs, various studies have been developed in order to check whether embryo culture conditions were improved in the non-invasive systems. Most studies have confirmed the safety of TLSs, finding no differences or detrimental effects on embryos cultured in these devices (Nakahara *et al.*, 2010; Cruz *et al.*, 2011; Kirkegaard *et al.*, 2012a; Park *et al.*, 2015). To observe the real effect of culture, it is crucial to use the exact same embryo selection criteria. In 2011, Cruz *et al.* (2011) proved that Embryoscope incubation up to day three or five was comparable to standard incubation with regards to clinical outcome. This finding was later supported by Park *et al.* (2015) in a randomized controlled trial (RCT) with short-term culture up to day two and in various studies collected by Swain (2014). This last review also highlights the importance of faster recovery times from environmental changes when using benchtop incubators with smaller chamber sizes. Recently, studies regarding obstetric and perinatal outcomes in embryos cultured in TLS have concluded that no detrimental effects were observed, thereby confirming the safety of the incubation (Zaninovic *et al.*, 2015; Insua *et al.*, 2017).

Few RCTs have been performed in order to prospectively validate TLSs according to their clinical outcome. A 2018 Cochrane review based on eight RCTs (Kahraman *et al.*, 2013; Kovacs *et al.*, 2013; Rubio *et al.*, 2014; Park *et al.*, 2015; Goodman *et al.*, 2016; Wu *et al.*, 2016;

Kaser *et al.*, 2017; Yang *et al.*, 2017) did not find enough evidence of benefits of using TLSs (Armstrong *et al.*, 2015). However, a meta-analysis performed by Pribenszky *et al.* (2017) regarding four of the Cochrane's included RCTs (Kahraman *et al.*, 2013; Kovacs *et al.*, 2013; Rubio *et al.*, 2014; Goodman *et al.*, 2016) and an additional one (Siristatidis *et al.*, 2015) demonstrated the use of TLS yielded higher significant ongoing pregnancy rates (OPR) and live birth rates (LBR). Not only early pregnancy loss was significantly reduced but no differences were found in stillbirth rate. Out of these five studies, Goodman's was recognized as the less biased RCT (Wu *et al.*, 2016). As stated by Meseguer (2016), it is not clear whether this improvement was due to the unchanged culture conditions or the use of the ESAs based on morphokinetics. However, if comparable reproductive results were obtained between TLS and standard incubator when using the same selection method (Cruz *et al.*, 2011; Park *et al.*, 2015), the morphokinetic selection might be the decisive factor. Despite this, a couple of studies comparing only selection methods using TLS in both, control and study groups, were not conclusive in terms of pregnancy (Goodman *et al.*, 2016; Kaser *et al.*, 2017).

2.2 Limitations of Conventional Embryo Scoring

"Are the variables related to timing of cleavages and dynamic changes in morphology useful markers of embryo viability (morphokinetics and morphology dynamics)?" (Meseguer, 2016).

Static evaluation of embryo morphology at predefined time points has proven to have clear limitations including the high inter- and intraobserver variability and the missing of decisive events potentially detrimental to embryo's viability (Ruiz de Assin *et al.*, 2009; Paternot *et al.*, 2011). Time-lapse imaging, on the other hand, has considerably increased the amount of reliable and detailed information regarding embryo development, in the form of qualitative and quantitative parameters. It aims at increasing the objectivity of the selection method and adding precision by detecting subtle unnoticed shifts. Morphokinetic parameters include the timings of embryonic morphological changes measured from the time of insemination and the duration of the cell cycles (**Table 1**). Numerous research groups have

joined the search for new useful markers determining embryo quality. Several have subsequently developed ESAs, used as prediction tools. ESAs work by combining certain qualitative and quantitative parameters, based on observed clinical outcomes, as selection or deselection criteria to pick the best embryo for transfer.

***Table 1**. Morphokinetic Parameters.*

Parameter	Formula	Definition
t0		time of insemination
tPB2		time to appearance of the second polar body
tPNa		time to pronuclei appearance
tPNf		time to pronuclei fading
P1		duration of the first citokinesis
t*n*		time to *n* cell division
tM		time to the morula stage
tSB		time to start of blastulation
tB		time to blastocyst formation
tEB		time to expanded blastocyst
tHBi		time to start of hatching
tHB		time to hatched blastocyst
cc1	t2-t0	duration of the first round of cleavage; equal to t2
cc2 or P2	t3-t2	duration of the second round of cleavage
cc3	t5-t3	duration of the third round of cleavage
cc4	t9-t5	duration of the fourth round of cleavage
s1	t2-t2	duration of the first synchronisation parameter, equal to 0
s2 or P3	t4-t3	duration of the second synchronisation parameter
s3	t8-t5	duration of the third synchronisation parameter

2.3 Qualitative parameters

Qualitative parameters are used as deselection criteria in ESAs, as many are associated with poor prognosis. These morphological phenomena include abnormal cleavage patterns, multinucleation, formation and internalization of fragments, appearance of the blastocoel and blastocyst collapse, and reexpansion (Zaninovic *et al.*, 2017; Coticchio *et al.*, 2018; ESHRE Working group on Time-lapse technology *et al.*, 2020). Embryos presenting these phenotypes have shown a concerningly high distinct quality on day three but a following sharp decrease in the correct formation of blastocysts (Athayde Wirka *et al.*, 2014) and in particular euploid blastocysts (Zhan *et al.*, 2016). Many other studies have linked them to lower IRs, clinical pregnancy rates (CPR), and LBRs (Rubio *et al.*, 2012; Desai *et al.*, 2014, 2016; Aguilar *et al.*, 2016; Goodman *et al.*, 2016; Zhan *et al.*, 2016; Azzarello *et al.*, 2017; Desch *et al.*, 2017; Ebner *et al.*, 2017).

Yang *et al.'s* (2015) study focusing on abnormal divisions of 345 *in vitro* matured metaphase I, concluded that developmental arrest, direct cleavage, disordered division, and fragmentation events were more detrimental to blastocyst formation than distorted cytoplasmic movement, uneven blastomeres, and the presence of big fragments. Barrie *et al.'s* (2017) retrospective study, including 15,819 embryos cultured in a TLS, examined five abnormal cleavage patterns: direct cleavage, reverse cleavage, absent cleavage, chaotic cleavage, and cell lysis. The prevalence of all abnormal phenotypes was 11.4% in total, constituting chaotic and direct cleavage the majority of them. Embryos showed a reduced developmental capacity resulting in reduced IR, suggesting this kind of embryos should be deselected for transfer when alternative embryos are available. A very recent study performed by Desai *et al.* (2018) found no association with aneuploidy when one of these anomalies was present but did when two or more dysmorphisms were observed in the embryo. As shown in **Table 2**, various embryo ESAs have implemented these abnormal phenotypes as exclusion criteria to deselect embryos.

Table 2. *Published Embryo Selection Algorithms (ESAs).*

Published ESAs	Sample size	Selection parameters: Qualitative	Selection parameters: Quantitative	AUC (95% CI)	AUC according to Petersen et al., 2016	AUC according to Barrie et al., 2017b	Predictive outcome	ESA day	Study type
Wong *et al*., 2010	100 emb	-	P1 P2 P3	-	-	-		2	Single-centre
Kirkegaard *et al*., 2012	571 emb	DC (1 to 3 Cells)	P1 s2	0.690 (0.650–0.740)	-	-		2	Single-centre
Conaghan *et al*., 2013	292 emb	-	P2 P3	-	0.620 (0.61–0.631)	-		2	Multicentre (5)
Chamayou *et al*., 2013	244 emb	-	t1 t2 t4 t7 t8 tPNf-tPNa s3	-	-	-	Blastocyst	3	Single-centre
Milewski *et al*., 2015	432 emb	-	t2 t5 cc2	0.806 (0.747–0.864)	0.688 (0.677–0.700)	-	Formation	3	Single-centre
Petersen *et al*., 2016	11,218 emb	-	t3-tPNf t3 (cc3)/(t5-t2)	0.745 (0.734–0.756)	0.745 (0.734–0.756)	-		3	Multicentre (24)
Goodman *et al*., 2016	1012 emb	MN AC	cc2 t5 s2 s3 tSB	-	-	-		5–6	Single-centre
Motato *et al*., 2016	7483 emb	-	tM s3	0.849 (0.835–0.854)	-	-		4	Single-centre
Meseguer *et al*., 2011	247 KID	BE (2C) DC MN (4C)	t5 s2 cc2	0.720 (0.645–0.795)	0.676 (0.665–0.688)	-		3	Single-centre
Chamayou *et al*., 2013	178 KID	-	cc3	-	-	0.552		3	Single-centre
VerMilyea *et al*., 2014	331 KID	-	P2 P3	-	0.685 (0.673–0.697)	-		2	Multicentre (6)
Basile *et al*., 2015	754 KID	BE (2C) DC MN (4C)	t3 cc2 t5	0.610 (0.574–0.638)	0.700 (0.688–0.712)	0.558	Implantation	3	Multicentre (4)
Liu *et al*., 2016	270 KID	DC RC < 6 ICCP (4C)	t5-PNf cc2	0.762 (0.701–0.824)	0.753 (0.743–0.764)	-	Potential	3	Single-centre
Petersen *et al*., 2016	3275 KID	-	t3-tPNf t3 (cc3)/(t5-t2)	0.650 (n/a)	0.650 (n/a)	-		3	Multicentre (24)
Goodman *et al*., 2016	94 KID	MN AC	cc2 t5 s2 s3 tSB	-	-	-		5–6	Single-centre
Motato *et al*., 2016	832 KID	-	tEB s3	0.591 (0.552–0.630)	-	-		5–6	Single-centre
Motato *et al*., 2016	832 KID	Morphology	tEB s3	0.602 (0.559–0.645)	-	-		5–6	Single-centre
Azzarello *et al*., 2012	159 KID	-	tPNf	-	-	0.584	Live birth rate	2	Single-centre
Campbell *et al*., 2013a	98 blas	-	tB tSB	0.720	-	0.579		5–6	Single-centre
Basile *et al*., 2014	504 emb	-	t5-t2 cc3	0.634 (0.581–0.687)	-	-	Aneuploidy risk	3	Single-centre
Del Carmen Nogales *et al*., 2017	485 emb	-	t3 t5-t2	-	-	-		3	Single-centre

Quantitative Parameters defined in Table 1. AUC = area under the curve; CI = confidence interval; emb = embryos; KID = known implantation data; blas = blastocysts; DC = direct cleavage; MN = multinucleation; AC = abnormal cleavage; BE = blastomeric evenness; 2C = 2-cell stage; 4C = 4-cell stage; RC = reverse cleavage; ICCP = intercellular cell count points.

3. EMBRYO SELECTION ALGORITHMS

The main ESAs reviewed below are depicted in **Table 2**. Another list with the parameters with the strongest biological or clinical significance has been published by ESHRE Working group on Time-lapse technology *et al.* (2020). A clear progressive evolution can be noted in the decision of the clinical outcome endpoint, from blastocyst formation to LBR with genetic testing. This has been highly dependent on the complementary evolution of the embryo transfer day from three to five/six. Blastocyst transfers is the current desired standard towards a most likely implantation. For this reason, ESAs predicting blastocyst formation might become obsolete, as they are only used by clinics still performing day three transfers.

3.1 Predictive Algorithms

3.1.1 Blastocyst Formation

Initially, day three embryo transfers promoted the forecasting of blastocyst formation as a decisive factor for selection. Morphokinetic events in the first blastomeric divisions therefore constituted the majority of ESA's parameters. The introduction of the first ESA was led by Wong *et al.* (2010), despite it was not clinically applied until years later by Conaghan *et al.* (2013). It was based on a previous publication by Pribenszky *et al.* (2010) which stated the strong predictive value of duration of the first citokinesis (P1) and duration of the second round of cleavage (P2) in mouse embryos. The chances for an embryo to result in a successful outcome were explained to be relying on its proper development before its genomic activation, happening close to the 4-cell stage. However, day five transfers introduced significant later stage parameters, time to the morula stage (tM) (Motato *et al.*, 2016) and time to start of blastulation (tSB) (Goodman *et al.*, 2016), to the predictive models. Being at such late stages in development, it is almost certain that embryos will expand to the blastocyst stage. This is why Motato *et al.* (2016) obtained quite a high area under the curve (AUC) value. Petersen *et al.* (2016) also obtained a high AUC supported by a massive

multicentre sample size. The most used quantitative parameters were duration of the second round of cleavage (cc2) or P2 and duration of the second synchronization parameter (s2 or P3) (Wong *et al.*, 2010; Kirkegaard *et al.*, 2012a; Conaghan *et al.*, 2013; Milewski *et al.*, 2015; Goodman *et al.*, 2016), followed by duration of the third synchronization parameter (s3), t2, and t5 (Chamayou *et al.*, 2013; Goodman *et al.*, 2016; Milewski *et al.*, 2016; Motato *et al.*, 2016; Petersen *et al.*, 2016).

3.1.2 Implantation Potential

Meseguer *et al.'s* (2011) study was the first one performed in a clinical setup with a more relevant endpoint, implantation potential. Meseguer introduced the concept of kinetic parameters having optimal ranges, where embryo's rushing or delaying out of range resulted in poor prognosis. The outstanding quantitative parameters predicting implantation potential were cc2 or P2 and t5 (Meseguer *et al.*, 2011; Vermilyea *et al.*, 2014; Basile *et al.*, 2015; Goodman *et al.*, 2016; Liu *et al.*, 2016; Petersen *et al.*, 2016) followed by s2 (Meseguer *et al.*, 2011; Vermilyea *et al.*, 2014; Goodman *et al.*, 2016). Time to pronuclei fading (tPNf) was also supported by different studies as a potential predictor (Azzarello *et al.*, 2012; Chamayou *et al.*, 2013; Desai *et al.*, 2016; Liu *et al.*, 2016; Petersen *et al.*, 2016). Posterior studies also gave importance to t4 and t8 with optimal ranges (Carrasco *et al.*, 2017).

Briefly summarizing the main ESAs, Wong's (Wong *et al.*, 2010) was further developed in a multicenter prospective trial (Conaghan *et al.*, 2013) and later correlated with IR and OPR in different retrospective and prospective studies performed by external groups (Vermilyea *et al.*, 2014; Diamond *et al.*, 2015; Adamson *et al.*, 2016; Aparicio-Ruiz *et al.*, 2016). Contrary to these findings, a prospective study found no differences between Eeva and conventional scoring groups (Kieslinger *et al.*, 2016). Meseguer *et al.'s* (2011) model has been widely validated by external groups (Kahraman *et al.*, 2013; Rubio *et al.*, 2014; Siristatidis *et al.*, 2015) and upgraded in a larger multicentre sample (Basile *et al.*, 2015). The extension of embryo culture up to the blastocyst stage led to the emergence of new ESAs in the same group (Motato *et al.*, 2016). It also brought the use of later stage parameters, time to expanded blastocyst (tEB) (Motato *et al.*, 2016) and tSB (Goodman *et al.*, 2016), which were closely

related to IR at day five transfers. A rather high AUC value was obtained by Liu *et al.* (2016), which was later supported by Petersen *et al.* (2016). However, further RCTs should be performed to validate this algorithm. A kinetic algorithm was developed by Petersen (Petersen *et al.*, 2016) as a result of the unsuccessful external validations of existing ESAs. The high number of dependent variables present in each laboratory makes it hard to find one ESA suitable for every clinic. Petersen *et al.'s* (2016) model claimed to work regardless of the culture conditions (culture media, oxygen concentration, etc.) and fertilization method (IVF or ICSI) as the retrospective evaluation data had been extracted from 24 different clinics. A validation with data collected from 31 clinics was performed strengthening the application of the algorithm. However, another validation might need to be performed without overlapping clinics used in the development of the algorithm. Goodman *et al.*'s (2016) parameters were picked as a combination from previous working ESAs, however, no significant differences were found in terms of pregnancy. Finally, it should be noted the progress towards forecasting LBR, pointed out by Azzarello *et al.*'s (2012), as the next predictive outcome to pursue in the validation of the existing models.

Although ESAs have increased the amount of information to make a better decision, we must bear in mind that there are other factors to be considered to predict implantation, such as embryo's euploidy and the endometrial status. Perhaps a higher predictive power is unattainable if these factors remain uncertain.

3.1.3 Aneuploidy Risk

"Can euploidy be forecast by a non-invasive morphokinetic test?" (Meseguer, 2016).

Not always seemingly good-quality embryos implant successfully. Chromosomal errors can translate into poor clinical outcomes. It has been shown that over half of the embryos in the cohort can carry aneuploidies once women reach the age of 35 (Fragouli and Wells, 2011). PGT-A has been consolidated as the best technique to identify aneuploid embryos. However, not only it requires a high degree of technical proficiency from a specialist, but it is also an invasive method that involves the performance of a biopsy with unknown consequences. For

this reason, development of non-invasive alternatives is considered. In light of the morphokinetic parameters predicting successful development to blastocyst, their ability to distinguish between euploid and aneuploid embryos was explored. It was therefore hypothesized that cell division length has to be within an optimum range in order to overcome all the cellular processes preceding cytokinesis.

One of the first studies examining the relationship between genetics and cell timings was performed by Chavez *et al.* (2012), in which euploid embryos demonstrated precise timings at the first cell divisions up to the four-cell stage. Continuing with further studies performing biopsies at day three analyzed using array-comparative genome hybridization (aCGH), Basile *et al.* (2014) proposed an ESA to detect aneuploid risk, depicted in **Table 2**. Vera-Rodriguez *et al.* (2015) identified P1 and s2 as being significantly longer in aneuploid embryos. In fact, s2 reached a three-fold difference. Overtime observed in P1 suggested an error in chromosome segregation with a subsequent negative impact on the first mitotic divisions. Del Carmen Nogales *et al.* (2017) went one step further by distinguishing between the specific aneuploidies, bearing in mind that not all types of chromosome abnormalities behave kinetically equal. They observed that kinetic patterns present in monosomic and trisomic embryos were not that far from euploid embryos compared to very complex aneuploidies, which differed to a larger extent.

The extension of culture to the blastocyst stage introduced new significant parameters in later stages of development. The ESA model proposed by Campbell *et al.* (2013) was based on the observed delay of tSB and time to blastocyst formation (tB) in aneuploid embryos diagnosed using aCGH. A validation followed, performed by the same group, in which significant differences were demonstrated between the risk categories (Campbell *et al.*, 2013b). However, two independent groups, one using the same diagnostic test (Kramer *et al.*, 2014) and another using single nucleotide polymorphism microarray (Rienzi *et al.*, 2015), tried to apply this model to their data but did not find significant differences in the proportion of euploid/aneuploid embryos in any of the categories. This could be due to the limited sample size (n = 98). The number of embryos considered to have a high risk of aneuploidy in these studies seemed too low (12 out of 97 and 4 out of 88), compared to what is expected according to literature (Fragouli and Wells, 2011). A posterior study with a much larger

sample size (n = 1730) supported the presumption that late stages are the decisive ones in aneuploidy search, finding t4, s2, tSB, tB, tEB, and time to hatched blastocyst (tHB) significant ($P < 0.05$ for early and $P < 0.0001$ for later stages of development) (Minasi *et al.*, 2016). Finally, a more recent study correlated chromosomal status, analyzed using a combination of both tests aCGH and next generation sequencing, with the morphokinetic parameters tSB, tEB and tEB-tSB (Desai *et al.*, 2018). The early arrival to the blastocyst stage was a sign of proper chromosomal development.

Nevertheless, aneuploidy risk models are only recommended when PGT-A is unavailable, when alternative embryos are available in the cohort, or when they are in conjunction with PGT-A, prioritizing embryos for the biopsy.

3.2 Embryo Selection Algorithms Concerns

3.2.1 Morphokinetic Parameters Variation

Other concerns regarding ESAs' performance include variation of morphokinetic values due to the uncertainty in the definition of time of insemination (t0) and the subjectivity of the manual annotations (Barrie *et al.*, 2017b). The use of relative parameters regarding the duration of events (P1, P2, P3, cc2, s2, etc.) may overcome the limitations of imprecise starting points, as they are more reliable than absolute time points (Kirkegaard *et al.*, 2013; Cetinkaya *et al.*, 2014). ESAs making use of these range parameters are depicted in **Table 2**. Alternative reference starting points, such as tPNf, have also been proposed as a solution (Liu *et al.*, 2015a).

Consistency in the TLS annotations is also essential to objectify embryo scoring and selection. Publications related to annotation guidelines (Ciray *et al.*, 2014) have been unable to completely eliminate variation. Moreover, accuracy is highly dependent on the skill and expertise of the individual in charge. Sundvall *et al.'s* (2013) study noticed a high level of intra- and interobserver agreement, but there were still some parameters where observers

struggled: time to pronuclei appearance (tPNa), multinucleation, evenness of blastomeres, late-stage timings, and number of collapses during the blastocyst expansion (**Table 3**). A following study by Martínez-Granados *et al.* (2017) strengthened this conclusion with a validation of 24 laboratories adding time to appearance of the second polar body (tPB2), t8, and tM to the struggling parameters. This subjectivity needs to be corrected and might be overcome by automation.

Table 3. *Inter-observer and intra-observer agreement for assessed parameters (image from Sundvall et al., 2013).*

Parameter	Inter ICC	*n*	95% CI	Intra ICC	*n*	95% CI
PN appearance	0.678	113	0.568–0.764	0.469	118	0.1–0.684
PN abuttal	0.773	90	0.692–0.838	0.685	108	0.569–0.775
PN breakdown	0.999	130	0.998–0.999	1	136	0.999–1
First cytokinesis	0.92	138	0.895–0.94	0.909	142	0.876–0.934
First nucleus appearance	0.984	90	0.977–0.989	0.982	123	0.974–0.987
Second nucleus appearance	0.967	78	0.953–0.978	0.965	117	0.948–0.976
First nucleus disappearance	0.989	86	0.985–0.993	0.966	115	0.951–0.976
Second nucleus disappearance	0.75	75	0.66–0.824	0.744	109	0.646–0.817
First cleavage division	0.958	134	0.945–0.969	0.953	150	0.935–0.965
Second cleavage division	0.885	133	0.85–0.913	0.954	145	0.937–0.967
Third cleavage division	0.809	141	0.756–0.853	0.87	145	0.824–0.905
Fourth cleavage division	0.863	125	0.821–0.897	0.929	138	0.902–0.949
Fifth cleavage division	0.868	123	0.826–0.901	0.899	129	0.86–0.928
Sixth cleavage division	0.906	99	0.872–0.933	0.923	118	0.889–0.946
Seventh cleavage division	0.9	68	0.852–0.935	0.975	93	0.962–0.983
Final division	0.599	82	0.336–0.757	0.708	74	0.573–0.805
Compaction	0.686	114	0.581–0.769	0.817	122	0.745–0.87
Morula	0.737	111	0.642–0.81	0.908	116	0.87–0.936
Early blastocyst	0.894	114	0.8–0.939	0.932	114	0.899–0.954
Full expanded blastocyst	0.913	116	0.869–0.941	0.937	118	0.91–0.956
Hatching blastocyst	0.872	75	0.816–0.914	0.968	84	0.952–0.979
Hatched blastocyst	0.99	11	0.972–0.997	0.973	12	0.912–0.992
Blastocoel contractions	0.653	75	0.536–0.752	0.771	114	0.685–0.836
Evenness	0.272	115	0.146–0.4	0.514	141	0.382–0.625
Multinucleation	0.268	153	0.168–0.372	0.472	155	0.339–0.586
Median ICC	0.87			0.92		
Mean ICC	0.81			0.85		

Note: inter-observer and intra-observer agreement expressed by ICC and 95% confidence interval (CI) of ICC values. N, number of embryos annotated by all three observers (inter-observer assessment) or by repeated annotations (intra-observer assessment).
ICC, intra-class correlation coefficient; PN, pronuclei.

3.2.2 Universal Algorithm

In the past few years, diverse algorithms determining embryo quality have been developed based on morphokinetic markers. Although promising results have been observed, none of them has yet found a universal acceptance. A recent study performed by Storr *et al.* (2018) determined the level of agreement between seven different published algorithms (Meseguer *et al.*, 2011; Conaghan *et al.*, 2013; Vermilyea *et al.*, 2014; Basile *et al.*, 2015; Goodman *et al.*, 2016; Liu *et al.*, 2016; Petersen *et al.*, 2016). As expected, Conaghan *et al.* (2013) and Vermilyea *et al.* (2014) showed a "very good agreement" as well as Meseguer *et al.* (2011) and Basile *et al.* (2015) with a "good agreement", as they are upgrades of their preceding algorithms. Validation of ESAs has mostly only been successful internally (Meseguer *et al.*, 2012; Campbell *et al.*, 2013b; Rubio *et al.*, 2014; Vermilyea *et al.*, 2014; Basile *et al.*, 2015; Diamond *et al.*, 2015; Adamson *et al.*, 2016; Liu *et al.*, 2016; Petersen *et al.*, 2016).

On the other hand, validation by other independent external groups has usually been rather unsuccessful (Fréour *et al.*, 2015; Liu *et al.*, 2015b; Rienzi *et al.*, 2015; Kieslinger *et al.*, 2016; Minasi *et al.*, 2016). Some have blamed the lack of resources or difficulties with logistics (Meseguer, 2016). Many others have blamed differential patient, treatment, and environmental conditions between laboratories as the main cause (Barrie *et al.*, 2017b; Aparicio-Ruiz *et al.*, 2018). However, the lack of agreement with and within the rest of ESAs raises concerns regarding the parameters being used, the uncertainty in the implantation potential and the unclear possibility of clinically applying ESAs in different settings.

Barrie *et al.* (2017b) examined the efficacy of six published ESAs regarding their predictive value, expressed as IR, to assess their clinical applicability (Azzarello *et al.*, 2012; Cruz *et al.*, 2012; Dal Canto *et al.*, 2012; Campbell *et al.*, 2013a; Chamayou *et al.*, 2013; Basile *et al.*, 2015). All ESAs showed a limited predictive ability and poor diagnostic value (**Table 2**). Despite Campbell *et al.*'s (2013a) ESA not being considered robust enough to be clinically applied, it was yet found to be the most effective overall. Basile *et al.* (2015) was also pointed out as the ESA with higher significant IR differences between categories in its hierarchy. Overall, they concluded that the ESAs available so far do not provide a significant aid,

encouraging the development of personalized in-house ESAs. Milewski *et al.* (2016) even stated the need of customized clinic-specific algorithms. This goes against the will of others who pursue the development of a universally accepted ESA. Anyway, future strong prospective studies will be crucial to determine the way.

Transferability of ESAs between laboratories using quantitative parameters is reduced according to observations. Qualitative parameters, however, seem to be independent of cleavage timings, ensuring a higher reproducibility between clinics Liu *et al.* (2015b). A study performed by Martínez-Granados *et al.* (2017) depicted almost perfect inter-laboratory agreement in most of the qualitative variables, which subsequently increased the level of agreement on the clinical decision when using TLS. Time-lapse undoubtedly standsout due to the witnessing of previously ignored phenomena significant to embryo viability. Maybe we should move forward towards the development of ESAs only based on qualitative parameters (Barrie *et al.*, 2017a). A study performed by Yang *et al.* (2015) already developed a hierarchical day three ESA based solely on qualitative parameters. Although significantly different IR were obtained in different categories, a subsequent validation performed by the same group registered worse OPR with the developed ESA than with a conventional morphological score (Yang *et al.*, 2018). Highlights from this ESA, it was developed from *in vitro* maturation metaphase I oocytes and validated in a reduced sample size of 144 KID (Known Implantation Data) embryos. Further studies with larger sample sizes and matured metaphase II embryos should be performed to confirm this hypothesis.

The high prevalence of aneuploid embryos (Fragouli and Wells, 2011; Campbell *et al.*, 2013a; Franasiak *et al.*, 2014) makes it unlikely for an ESA to be robust enough, as many of these embryos reaching the blastocyst stage subsequently fail to become pregnancies, either not implanting or resulting in a miscarriage. A genetic test should complement kinetic data to achieve a precise embryo-specific diagnosis. Although certain morphokinetic parameters — as previously described — have shown a correlation differentiating euploid and aneuploid embryos, their accuracy is still not sufficient to replace PGT-A (Zaninovic *et al.*, 2017). Moreover, no validation studies have confirmed the relationship observed between ploidy and morphokinetics (Aparicio-Ruiz *et al.*, 2018). Further studies are needed to broaden the range of morphokinetic markers, possibly more related to aneuploidy, check the specific

impact of certain aneuploidies and prospectively validate these models. A comparison of the different aneuploid risk ESAs should be performed in order to determine their specificity, sensitivity, and predictive value, similar to the one performed by Barrie *et al.* (2017b). Even though it is unlikely ever to be as accurate as PGT-A with a biopsy, this method could, nevertheless, be further developed into an important tool increasing the chances of a live birth by non-invasive means.

Many studies stress the importance of time-lapse in embryo selection as a combination of the morphokinetic parameters with a conventional morphology grading classification. Thus, time-lapse and morphology-based selection are not in competition with each other but rather complementary (Meseguer *et al.*, 2011, 2012; Cruz *et al.*, 2012; Cetinkaya *et al.*, 2014; Motato *et al.*, 2016). In all cases, laboratories should take a cautious approach performing an in-house validation before integrating any ESA into their clinical routine.

4. AUTOMATION

Since the beginning of computer vision, including acquisition, processing, and analysis of digital images, software programs are being explored as automatic alternatives in order to standardize TLS annotations. Automation in TLS is mostly based on machine learning, described by Yeung *et al.* (2018) as the capacity of a developed system for autonomous learning recognizing patterns in data. Exposing the system to a wide variety of labeled embryo development images in a "training" phase would later allow it to identify the different stages of embryo development. The system continues building knowledge from new unlabeled images refining its analysis. This can be helpful to eliminate human error, especially as the complexity of clinical practices continues to increase. The additional exponential growth in the number of cycles incubated with TLSs in the past few years would be benefited by automation as embryologists would not spend that much time performing manual annotations, allowing them to fulfill the many other tasks the IVF laboratory encloses. However, a clear limitation is that human vision easily identifies and interprets relevant events in complex scenarios associating the observed images with previous experiences, a task currently unmatched by computer vision (Danuser, 2011). **Figure 2** better summarizes the strengths of computer or human vision in image analysis. Moreover, poor embryos with complex morphologies or artifacts present in the well hampers the image analysis. For these reasons, embryologists must keep making the final call checking automation's performance. The establishment of automation needs a gradual transition with proper validations controlled by laboratory professionals.

As previously described, Eeva was developed as the first automatic algorithm-based embryo analysis platform (Wong *et al.*, 2010). Aparicio-Ruiz *et al.* (2016) performed a retrospective study with the purpose of validating its automatic cell-division tracking and autoclassification of embryos. Their results suggested it was found to be superior to conventional scoring using morphology. A limitation in the system was its use of dark field optics, as the low light level hampered the interpretation of the images. This could be problematic if the development of ESAs move towards the use of qualitative parameters, such

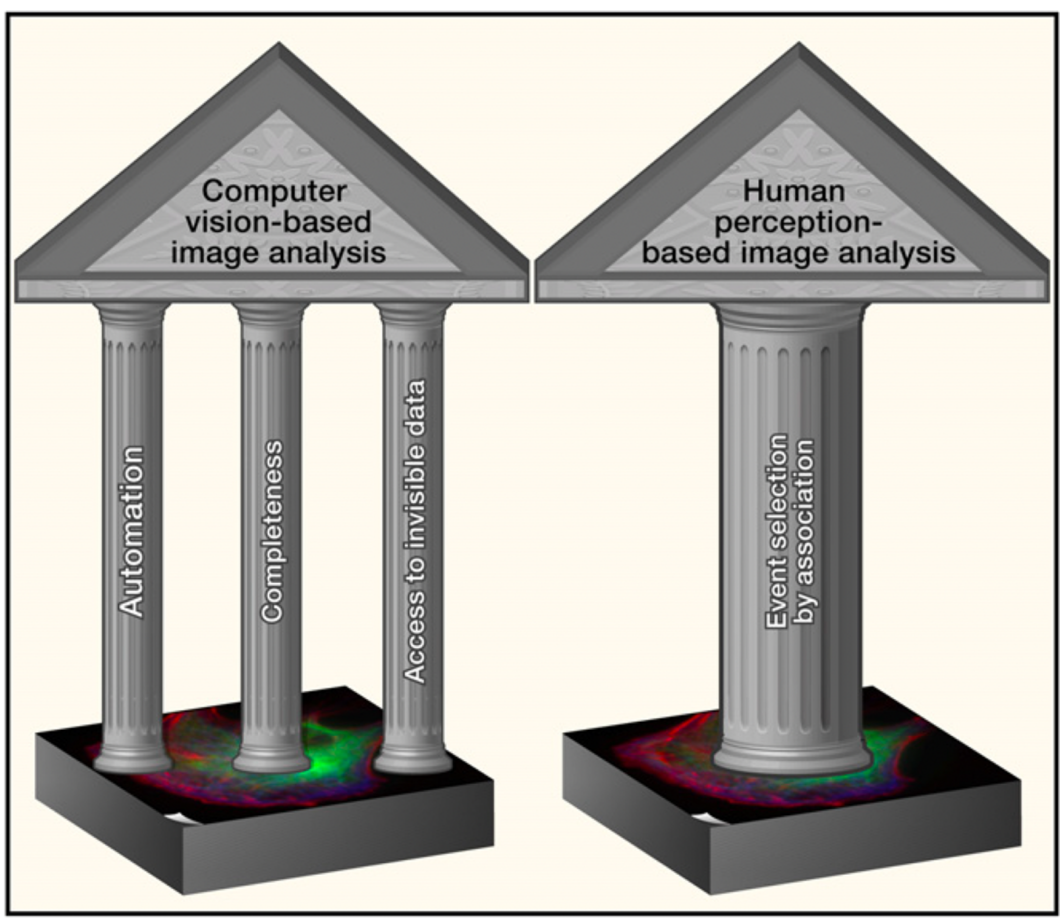


Figure 2. *The Strengths of Computer Vision and Human Vision in Image Analysis* (image from Danuser, 2011).

***Computer vision** rests on three pillars: it provides partial or complete automation of the analysis pipeline; it generates completeness in the data in that every image event fulfilling set selection criteria is considered by the analysis; and it can give access to processes underlying the image content that are not visible. Together, these pillars build a framework for solving complex image analysis tasks that require integration of a large number of well-defined yet multidimensional and possibly indirectly accessed image events. Computer vision systems generate quantitative and reproducible models of image content.*

***Human vision**, in contrast, rests on one strong pillar, that is the association of observed image signals with previous visual experiences. Because the memory of visual scenes stored by the human brain is huge, the association strategy permits a fast and adaptive interpretation of everchanging scenes that may consist of weakly defined image events, a performance currently unmatched by computer vision systems. However, human vision analysis results in a qualitative description of perceived image content that matches the best interpretation of the scene. The description may vary between individuals, it may be incomplete, and it may miss subtle but significant differences between distinct scenes.*

as multinucleation, as suggested by Liu *et al.* (2015b) and Barrie *et al.* (2017a). However, this was overcome, by Geri Plus where embryologists can move from dark to bright field. Another predictive model of implantation was recently developed by Milewski *et al.* (2017) based on principal component analysis and artificial neural networks. A training set of 428 embryos' quantitative and qualitative parameters was selected. Patient's age was also included as it is well described to strongly bias predictive outcomes (Milewski *et al.*, 2017; Liu *et al.*, 2018). The subsequent predictive model achieved a promising AUC value of 0.75 with a 95% confidence interval (0.7–0.8). Further validation studies need to be performed. Our group has been working on a different embryo automatic morphokinetic evaluation software. This strategy was assessed through cohort quality and prediction of embryo transfer success. The system uses UAD (units of average distance) values to check embryo's proximity to the implanting cloud-like dataset of KID embryos. The implantation potential of all the cohort embryos is subsequently ranked (Aparicio-Ruiz *et al.*, 2018). Recent studies showed significantly higher CPRs when embryos were selected according to the software's ranking (Alegre *et al.*, 2017, 2021).

As described by Yeung *et al.* (2018), major obstacles to machine learning include the poor quality of the databases, the complexity of the development of the clinical determination through the "training" phase and the failure to adapt to the routine clinical practice. The use of big data technology as a prediction tool is a promising approach to improve the effectiveness of IVF treatments.

5. ALTERNATIVE NON-INVASIVE TECHNIQUES

5.1 Introduction of Metabolomics

As previously described, morphological assessment is the current standard method of embryo evaluation, despite being subjective and associated with considerable interobserver variability (Baxter Bendus *et al.*, 2006). Specifically, the blastocyst stage has become established as the preferred time for assessing embryo quality before transfer (Glujovsky *et al.*, 2016). Morphological embryo evaluation has been enhanced by the development of TLSs, which allow non-invasive embryo assessment, and the use of cellular kinetics to analyse embryo quality (Meseguer *et al.*, 2011). Unfortunately, morphology is not always indicative of high-quality embryos and cannot predict euploidy, but until new and more refined methods of embryo evaluation are developed, morphological assessment remains the method of choice (Gardner and Schoolcraft, 1999; Gardner and Balaban, 2016).

An alternative method of assessing embryo quality is preimplantation genetic screening. This is an objective technique identifying chromosomally normal embryos, but the invasiveness of the biopsy procedure and the risk of undetected mosaicism are the main reasons why specialists are considering other options (The Practice Committee of the Society for Assisted Reproductive Technology and Practice Committee of the American Society for Reproductive Medicine., 2008; Kirkegaard *et al.*, 2012b; Cohen *et al.*, 2013; Juneau *et al.*, 2016).

Non-invasive techniques could overcome these limitations. For example, nuclear magnetic resonance imaging, mass spectrometry, or high-performance liquid chromatography can be used to evaluate the metabolism of pyruvate, glucose, lactate, or amino acid levels (Gardner *et al.*, 2001, 2011; Bromer and Seli, 2008; Sallam *et al.*, 2016). Other techniques assess oxygen consumption during cleavage (Tejera *et al.*, 2011, 2016) or conduct genomic and proteomic profiling (Dominguez *et al.*, 2015; Bori *et al.*, 2021). However, these technologies are usually not suitable for routine clinical practice because of the cost of

equipment, the time required to complete the process, the complexity of procedures, and the need for highly trained specialists, or because they add no substantial value to embryo selection relative to current methods.

During development, embryos undergo extensive metabolic processes that change throughout the different stages, which are reflected in the surrounding culture medium (Paszkowski and Clarke, 1996; Bromer and Seli, 2008). Some studies have suggested that the metabolic profile of the spent culture medium or ‘‘secretome’’ may be a better marker than morphological assessment for predicting successful implantation (Seli *et al.*, 2007, 2008; Scott *et al.*, 2008; Vergouw *et al.*, 2008; Sallam *et al.*, 2016). In particular, the metabolites involved in oxidative processes have been found to be the most predictive of implantation success (Seli *et al.*, 2007; Bromer and Seli, 2008).

One of the assays being investigated for culture assessment is thermochemiluminescence (TCL), which provides a rapid and accurate view of a sample's susceptibility to oxidation (i.e. oxidizability) (Wiener-Megnazi *et al.*, 2002; Shnizer *et al.*, 2003). Analysis by TCL of follicular fluid (Wiener-Megnazi *et al.*, 2002; Wiener-megnazi *et al.*, 2004) or seminal plasma (Lissak *et al.*, 2004) has been previously validated as a predictive tool in IVF. In addition, a prospective study performed by Wiener-Megnazi *et al.* (2011) showed that the oxidative status of the embryo assessed by TCL was indicative of successful implantation. Combining a TCL assay of the spent culture medium with a robust morphology and morphokinetic analysis using the TLS technique may be the key for a valid embryo selection algorithm.

JUSTIFICATION

JUSTIFICATION

The past few years have been marked by controversy surrounding the use of TLSs in routine clinical practice. Non-invasive culture conditions have proven to be at least equally effective when compared to standard incubators. However, whether the continuous observation of human embryos increases, clinical outcome success remains an open question, as some RCTs support this statement whilst others reject it.

Similarly, available evidence regarding ESAs is currently contentious, as, although validating studies performed by their creators have proven their efficiency, other external groups have pointed out their limited predictive value, questioning their application under different conditions. Concerns regarding the reproducibility of ESAs call the prospect of developing a universal ESA applicable to all different time-lapse devices and clinics. However, standardizing patient, treatment, and environmental laboratory conditions is hardly feasible.

Regardless of ESAs performance and considering the proven safety of TLSs, this technology is definitely the answer to get new insights into key stages of embryo development. Detection of detrimental abnormal embryo phenotypes has proven to be crucial for the process of deselecting embryos with a poor prognosis for transfer, preventing the negative implications that a negative pregnancy test can have on the patients.

Changing our approach towards a personalized in-house or qualitative-parameter based ESAs could pave the way for a new morphokinetic era. A combination of **morphology**, **kinetics**, **metabolomics**, and **genetics** could come close to the theoretical ideal for a comprehensive embryo selection. However, the optimization and standardization of each individual method is crucial for the correct application and clinical use.

Focusing specifically on morphology & kinetics, publications related to morphokinetic annotation have observed the high intra- and interobserver variability (Sundvall *et al.*, 2013; Martínez-Granados *et al.*, 2017; Adolfsson and Andershed, 2018). Accuracy has been proven

to be highly dependent on the skill and expertise of the individual in charge. This subjectivity hampers the elimination of variation and subsequently could modify the embryo being selected for transfer. Automation, in the form of image analysis through machine learning, can constitute an objective alternative to perform morphokinetic annotation, potentially increasing their predictive power.

On the other hand, metabolomics could be the perfect addition to complete a comprehensive embryo selection. Previous studies have suggested the use of the spent culture medium or ''secretome'' of the embryo to describe its metabolic profile, in particular metabolites involved in oxidative processes (Wiener-Megnazi *et al.*, 2002; Lissak *et al.*, 2004; Wiener-megnazi *et al.*, 2004; Botros *et al.*, 2008; Bromer and Seli, 2008). One of the assays being investigated for culture assessment, TCL, showed that the oxidative status of the embryo was indicative of successful implantation (Wiener-Megnazi *et al.*, 2011).

Combining a TCL assay of the spent culture medium with a robust automated morphokinetic analysis could provide the foundation for a more objective embryo selection method before transfer during routine clinical practice.

OBJECTIVES

OBJECTIVES

PRIMARY OBJECTIVES

- Validation of an automated annotation software as a standardized via to perform the morphokinetic annotations of the embryo developmental events during culture in the Time Lapse System.

- Validation of the thermochemiluminescence (TCL) assay, as a non-invasive tool to perform the analysis of the oxidative stress of the spent culture media of the embryo.

- Optimize embryo selection for transfer with non-invasive & objective techniques.

SECONDARY OBJECTIVES

- Evaluate the clinical use of the software in embryo grading and outcome prediction when using the automated annotations with previously published Embryo Selection Algorithms (ESAs).
- Study whether there is a correlation between the results of the TCL assay and the embryo quality and viability, to be used as a potential assessment method to select the embryo with the best chance of ensuring a successful pregnancy.
- Combining the TCL assay of the spent culture medium with a robust morphokinetic analysis to provide the foundation for a more objective embryo selection method before transfer during routine clinical practice.

MATERIAL & METHODS

MATERIAL & METHODS

The present thesis complies with the Spanish law governing assisted reproductive technologies (14/2006) and has the approval of the clinical research ethics committee of IVIRMA Valencia (IBR codes: 1611-VLC-079-MM and 1308-C-127-FD).

This project consists of two different parts. The fist one is carried out exclusively at IVIRMA Valencia and second one is carried out in two clinics of the IVIRMA group: IVIRMA Valencia and IVIRMA Alicante.

The population subject to the first retrospective study includes cycles from a total of 284 patients undergoing IVF treatments between February 2016 and January 2019. The second prospective study included a total of 292 patients undergoing cycles in the oocyte donation program between January 2016 and January 2017. Detailed cycle selection criteria is explained in the subsequent **Specific Objectives**.

1. *IN VITRO* FERTILIZATION STANDARD PROCEDURES

1.1 Controlled Ovarian Stimulation (COS)

Patients & donors were stimulated using the conventional controlled ovarian stimulation protocol with gonadotropin-releasing hormone (GnRH) agonist treatment, as described by Muñoz *et al.* (2012). Briefly, gonadotropins and GnRH agonist (Decapeptyl® 1, Ipsen Pharma, Barcelona, Spain) were administered by intramuscular injection until the mean diameter of the leading follicle(s) was ≥ 18 mm, when the hCG injection was administered. Transvaginal oocyte retrieval was scheduled 36 hours later.

1.2 Oocyte Pick-up & Fertilization Method

Oocyte preparation, ovum retrieval, and fertilization were performed using the standard protocols employed by the Instituto Valenciano de Infertilidad (IVI). Transvaginal oocyte retrieval was performed, where follicles were aspirated and the oocytes washed in gamete medium (Cook Medical; Sydney, Australia). Oocytes were then cultured in a fertilization medium (Cook Medical; Sydney, Australia) at 5% CO_2 and 37 °C. Four hours after retrieval, denudation was carried out by mechanical pipetting in 40 IU/mL of hyaluronidase in the same medium. Oocytes underwent exhaustive denudation, as time-lapse systems need a very clear image of the embryo. Immediately after this denudation, ICSI was performed in a HEPES-buffered gamete medium at x400 magnification using an Olympus IX7 microscope.

1.3 Time-lapse System Embryo Culture

Immediately after ICSI, oocytes were placed in preequilibrated dishes with cleavage medium (Cook Medical) in a time-lapse incubator. The tri-gas TLSs were set to 5% O_2, 6.0% CO_2, and 37°C. Individual embryos were cultured in EmbryoSlides (Vitrolife), composed of 12 straight-sided cylindrical wells, each containing 28 µL of cleavage medium. An overlay of 1.5 mL of mineral oil was used to prevent evaporation. Group cultured embryos were placed in the Geri dishes (Genea Biomedx), composed of 16 micro-wells with shared 80 µl cleavage medium and an overlay of 4 mL of mineral oil. The slides were pre-equilibrated to the incubator's conditions overnight, and air bubbles were removed to enable correct image acquisition. Transfers were performed at the blastocyst stage, so the medium was changed to continuous culture medium (CCM) (Vitrolife) on day 3.

1.4 Embryo Evaluation (Manual Morphokinetic Annotations)

Fertilization was assessed at 16 to 19 hours after ICSI by checking for the appearance of two distinct pronuclei and two polar bodies. Embryo evaluation was performed analysing the TLS images without opening the incubator or disturbing the embryos. On day 3, embryos underwent morphological evaluation using the 2015 Asociación para el Estudio de la Biología de la Reproducción (ASEBIR) morphology criteria (**Supplemental Table 1**) (ASEBIR, 2015). The morphological assessment on day 5 (approximately 120 hours after ICSI) evaluated each blastocyst according to the day 5 categories defined by ASEBIR in 2015 (**Supplemental Figure 1**).

Optimal or good morphology blastocysts were defined as having a cohesive trophectoderm (TE) composed of numerous sickle-shaped cells as well as a tightly packed inner cell mass (ICM). Blastocysts of optimal or good morphology were those at grade 2–4 stage of development, and with grade 1 or 2 ICM and TE, as described in the Alpha/ESHRE consensus guidelines (ALPHA Scientists in Reproduction medicine and ESHRE SIG of Embryology, 2011). A range of cell cycle durations and embryo maturation time parameters were also measured using TLS, as described on **Table 1** or defined by Motato *et al.* (2016) (**Supplemental Table 2**). Blastocysts were selected by applying a hierarchical classification procedure based on a combination of standard ASEBIR's morphological grading at day 5 (**Supplemental Figure 1**).

1.5 Embryo Transfer & Pregnancy Status

Embryo transfer was scheduled for day 5 by default in all patients, based on the date of oocyte capture and the time usually required to reach blastocyst stage. However, in a small number of patients the embryos that had not reached the blastocyst stage by day 5 were kept

under observation for one more day and transferred on day 6. All embryos were transferred fresh. Only single embryo transfers were performed in this study. The protocol for endometrial preparation of the patients was as described by Muñoz *et al.* (2012). Transfers could be cancelled due to concerns about endometrial receptivity. Supernumerary embryos with a quality equal to or greater than ASEBIR's grade D (**Supplemental Figure 1**) were frozen using the standard vitrification technique (Cobo *et al.*, 2010). After embryo transfer, oocyte recipients received a daily dose of 400 mg of vaginal micronized progesterone (Progeffik®, Lab. Effik, Madrid, Spain) as luteal phase support every 12 hours. To confirm successful implantation, the β human chorionic gonadotropin (β-hCG) value was determined 13 days after embryo transfer. Ongoing pregnancy was further verified after 12 weeks of pregnancy when a gestational sac with fetal heartbeat was visible on ultrasound examination.

2. AUTOMATED EMBRYO ASSESSMENT SOFTWARE

Geri Assess® 2.0 is an embryo development detection tool that automatically annotates key morphokinetic events and observations as the images are captured by the TLSs. The software was developed to assist embryologists in their daily clinical practice, reducing the time spent performing the morphokinetic annotations and increasing the information available to optimize the embryo ranking process. Its application aims to standardize the embryo development evaluation across and within laboratories.

This software is integrated in the TLS Geri® incubator, which captures high-resolution images of embryos at specified frames or time-points and at different focal planes, capturing all the image information coming from the cultured embryo. The software then selects the sharpest image at the most focused plane for the embryo and performs the standardised automated annotations for the occurring developmental event.

2.1 Software Development

The Geri Assess® 2.0, initially called Dana, was developed with the use of Artificial Intelligence (AI). The deep learning technique used is called Convolutional Neural Networks (CNNs) inspired by biological brain neural networks. This technique is commonly applied to pattern detection for image analysis. The first version of the software Dana was validated by our group in a 3 phase "training" study, described in Alegre *et al.* (2020). As summarized in **Figure 3,** in *Phase 1* we created a training cloud dataset of known implantation embryos with their manual annotations performed by an experienced clinical embryologist team in IVI Valencia. Essentially, the software was taught to recognize and classify embryo images as the different developmental events such as 2-cell, expanding blastocyst, etc. Later, it compared consequent images to track developmental patterns in the correct order. In *Phase 2,* the automated annotations performed by the software were checked manually looking for any false positives of false negatives. The software was re-trained to correct its mistakes and

improve its performance. The last *Phase 3* evaluated whether the embryo selection made by Dana correlated with its implantation rate.

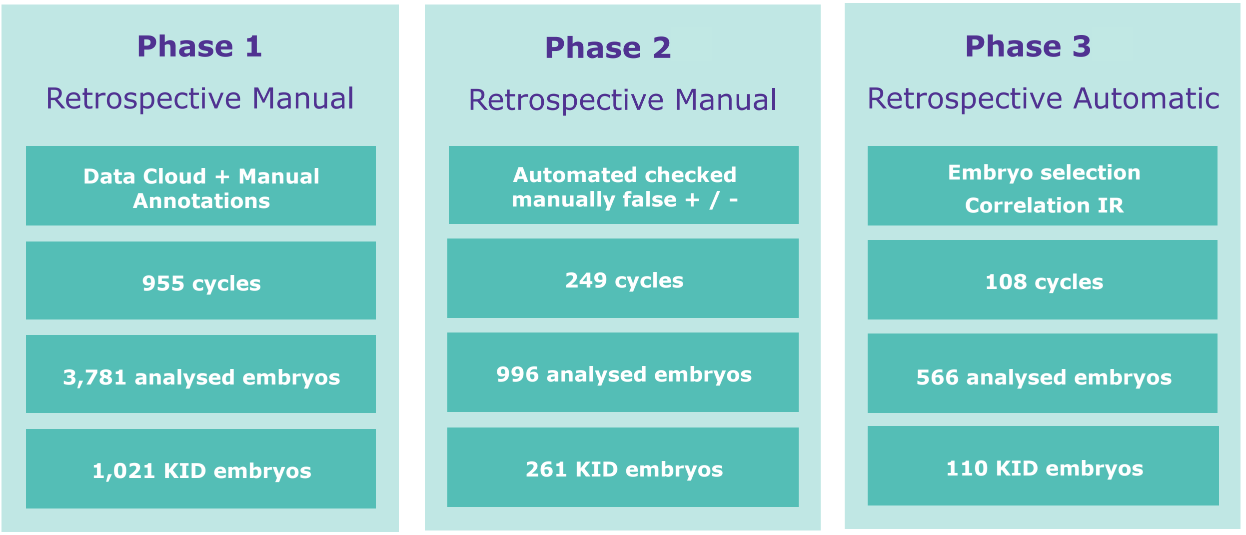


Figure 3. *Description of the initial 3-phase validation of the software Dana.*

2.2 Functionality

The integrated automated annotations software starts performing several image adaptations steps for a better analysis. First, it identifies the best focal plane and crops the video to position the embryo in the centre of the image and omit any external artifact, including the well of the culture dish.

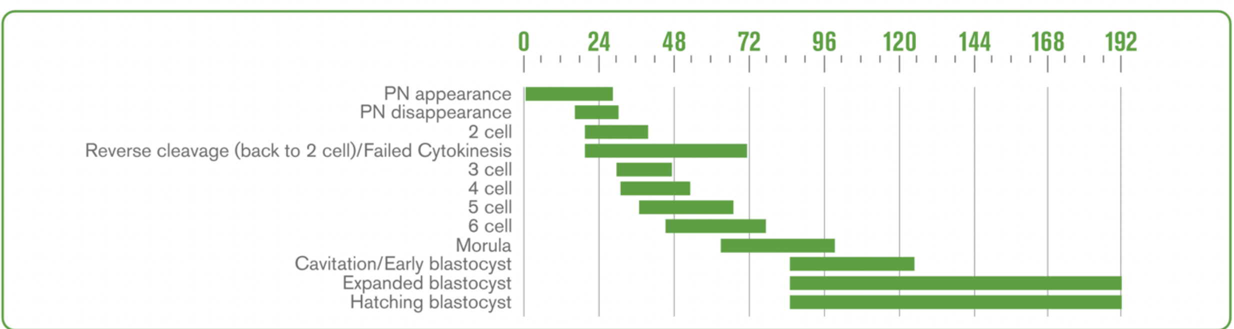
Observations of the different events which can occur during embryo development can be categorized as transient, like for example reverse cleavage, or ongoing, like fragmentation. The developmental events annotated by the software includes pronucleus or pronuclei appearance (tPNa), pronucleus or pronuclei disappearance (tPNf), divisions to 2-, 3-, 4-, 5- and 6-cell stages (t2, t3, t4, t5 and t6), transitions to morula (tM), early blastocyst (tSB), expanded (tEB) and hatching blastocyst stages (tHBi). The latest version also includes the detection of reverse cleavage, failed cytokinesis back to the 2-cell stage, and embryo fragmentation. **Table 4** summarizes and defines the full list. **Supplemental Figure 2** includes an explanation regarding the fragmentation detection performed by the software.

Table 4. *Summary of the key developmental events and observations annotated by Geri Assess 2.0. Observations marked with *. (Geri Assess 2.0 Tech Note, 2019)*

Icon	Event or Observation*	Definition of the Event or Observation
	PN Appearance	The first frame where appearance of at least one visible PN is identified
	PN Disappearance	The first frame where disappearance of all visible PNs is identified
	2-cell division	The first frame where 2 discrete membrane-separated blastomeres is identified
	Reverse Cleavage*	The first frame where the embryo has returned back to 2 discrete membrane-separated blastomeres after a cleavage or a failed cytokinesis event (applicable for 2-cell stage only)
	3-cell division	The first frame where 3 discrete membrane-separated blastomeres is identified
	4-cell division	The first frame where 4 discrete membrane-separated blastomeres is identified
	5-cell division	The first frame where 5 discrete membrane-separated blastomeres is identified
	6-cell division	The first frame where 6 discrete membrane-separated blastomeres is identified
	Morula Transition	The first frame where appearance of cellular compaction and blurring of distinctive individual cell membranes is identified
	Early Blastocyst Transition	The first frame where appearance of blastocyst cavitation is identified
	Expanded Blastocyst Transition	The first frame where appearance of clearly expanded blastocyst with a minimum diameter of 167 µm is identified
	Hatching Blastocyst Transition	The first frame where appearance of cellular hatching, shown as trophoblast cells extruding from the zona pellucida, is identified
N/A	Fragmentation*	A set of five consecutive images where the embryo is identified as having cellular fragmentation present in ≥15% of its volume (defined as extracellular membrane-bound cytoplasmic structures <45 µm diameter in a Day-2 embryo and <40 µm diameter in a Day-3 embryo). This event could appear several times or persist on the development timeline.

To ensure the reliability of the annotations performed by the software, a control feature was introduced as part of its functionality. Each embryo developmental event detected must fall within a specified time range to be included in the annotations. For example, it is well known that the time of cellular cleavage from one to two cells happens around 24 hours after insemination, therefore, if the software was annotating this event at 72 hours, falling outside the specified range for **t2**, a yellow triangle will mark the time, which will turn red (**Figure 4, A**). This feature pretends to alert the embryologist from two main scenarios, either the software has wrongly detected the event, or the detection was right, but the timing is so odd this embryo development should be thoroughly followed. The out-of-range annotations will require an embryologist closer inspection and possibly a manual modification of the annotation. The pre-defined time ranges for each event, as shown in **Figure 4 (B)**, were established based on published recommended review times (ALPHA Scientists in Reproduction medicine and ESHRE SIG of Embryology, 2011; Ciray *et al.*, 2014). They are meant as a guide in the daily clinical practice and should not be conclusive of the embryo's viability. This control feature will be referred as filtering in the Study Phases of the present thesis.

A



B

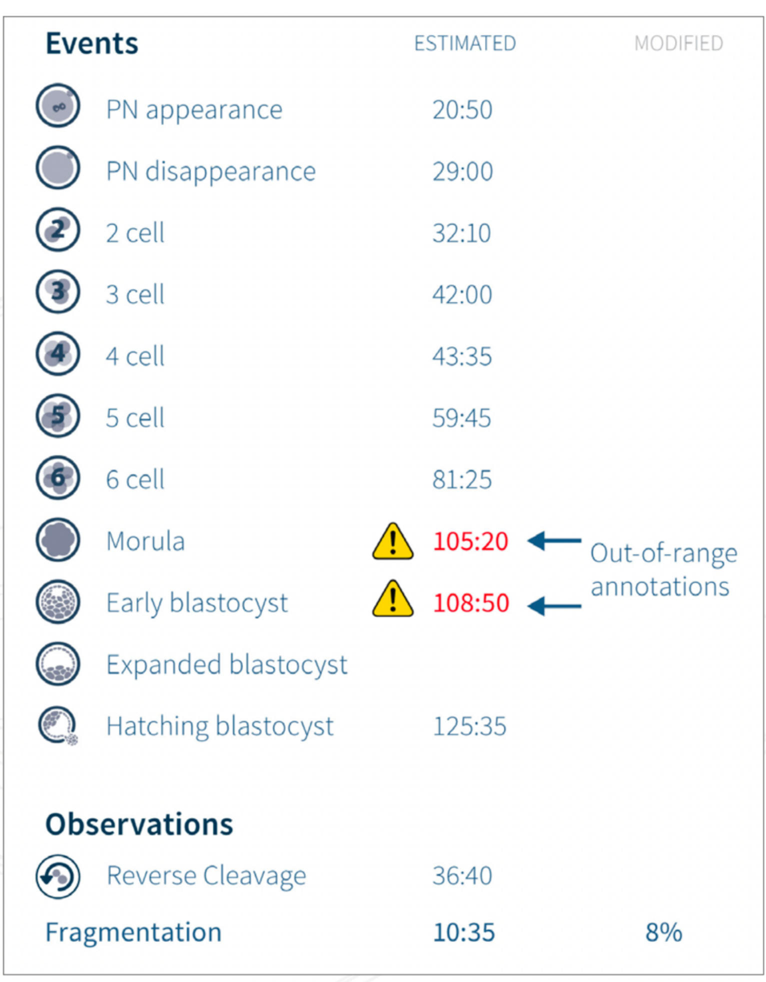


***Figure 4. (A)** Pre-defined time ranges for each developmental event determining its inclusion in the annotations performed by Geri Assess 2.0. **(B)** Geri Assess 2.0 tab showing the automated annotations, including the out-of-range, in red with the warning sign* (image from Geri Assess 2.0 Tech Note, 2019).

3. OXIDATIVE STATUS OF THE SPENT CULTURE MEDIA

Oxidizability of the embryo media samples was measured with the TCL Analyzer^{TM} (Luminest, Haifa), described in **Figure 5**. A sample preparation block and an analysis block compose the TCL Analyzer (**Figure 5, B**). A sample of 15 µl of culture CCM medium (Vitrolife) is positioned in the sample preparation block and distributed over the surface of a cylindrical aluminium tray (dimensions: diameter – 19mm, height – 3mm, thickness – 0.12mm). After being tightly sealed and vacuum dried, the sample is transferred to the analysis block where it is heated to a constant temperature of 80 +- 0.5ºC. As shown in **Figure 5 (C)**, biological molecules, such as proteins and lipids, undergo an induced oxidation generating unstable electronically excited species (EES). EES will further decompose into stable carbonyl end products and light energy, counted as photons emitted per second (cps). Sequential photon counting is performed in a spectral range of 350-600nm wavelength using a TCL Photometer incorporating bandpass interference filters and an electronic cathode tube, R265P photomultiplier (Hamamatsu Photonics Co. LTD.), for 300 seconds. TCL will therefore indicate the rate of formation of unstable carbonyls in the sample reflecting the initial amount of ROS, that is, the total oxidants.

A computerized TCL curve (**Figure 5, D**) is obtained as a trend line from individual cps points, mathematically described as the amplitude of the kinetic curve. Significant variation of photon readings is measured in the first 55 seconds, considering it as a stabilisation phase. Beyond 255 seconds no additional meaningful information is obtained. Therefore, the period between 55 and 255 seconds was selected due to its high reproducibility. Specifically, three time points were empirically selected: H1 (55 seconds), H2 (155 seconds) and H3 (255 seconds). Parameters such as average media (AVE) of the three time points and the slope between H1 and H3 (Ratio) were also considered.

$$\text{TCL Ratio (\%)} = \frac{\text{TCL amplitude 255 seconds (H3)} \times 100}{\text{TCL amplitude 55 seconds (H1)}}$$

A smoothing algorithm (sm) was applied to overcome the photomultiplier fluctuations, achieving a more continuous line. Moreover, the number of counts created by dark current, measured by the photomultiplier when it is not exposed to any light, was added to the measured values as dark noise (dn).

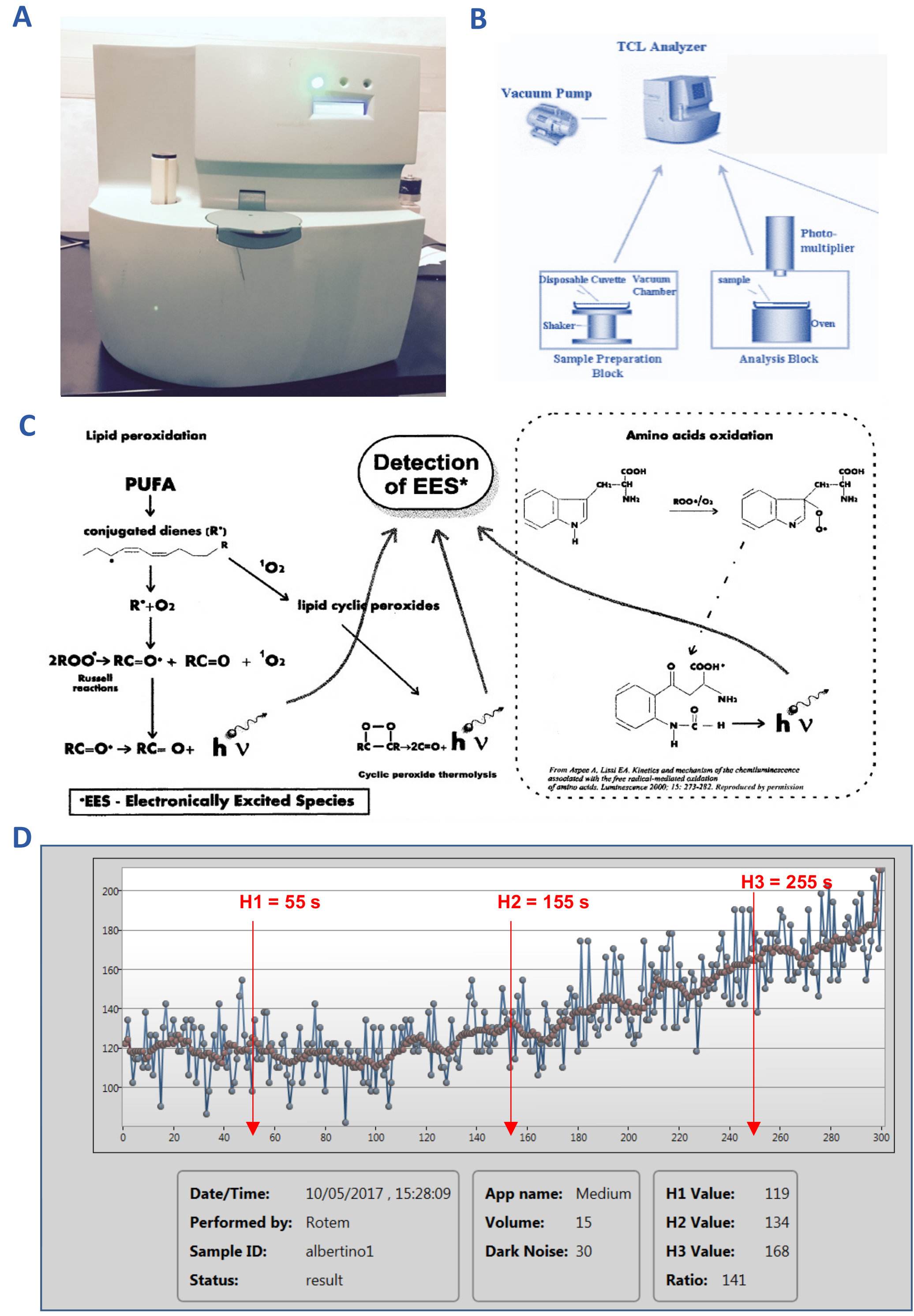


Figure 5. (A) *TCL device.* ***(B)*** *Sample preparation block and Analysis block.* ***(C)*** *Photochemical principles of thermochemiluminescence.* ***(D)*** *Example of a TCL kinetic curve. Parameters measured were amplitudes at 55 seconds, 155 seconds, and 255 seconds (**H1**, **H2** and **H3**, respectively).*

STUDY PHASES

STUDY PHASES

SPECIFIC OBJECTIVE I.

Characterization of the detection sensitivity & accuracy when performing the annotation of the human preimplantation embryo developmental events with an automated software.

1.1 Purpose

The main purpose of the development of an automated annotations software in IVF laboratories is to ease the clinical routine workflow of embryologists, standardize morphokinetic annotations, and ultimately improve embryo selection. The first version of our software Dana was developed in a 3 phase “training” study where a data cloud of KID embryos with their manual annotations was created. The automated annotations were checked manually for the software to learn if there were any false positives of false negatives. Finally, it was evaluated whether the embryo selection made by Dana correlated with a good implantation rate (Aparicio-Ruiz *et al.*, 2018; Alegre *et al.*, 2021).

The purpose of the first **Specific Objective I** of the present thesis was to validate the automated software for clinical use, as it is now integrated in the Geri system. In this regard, we compared the detection rate and accuracy of the embryo morphokinetic events annotations performed by embryologists in their routine clinical practice as compared to those performed by the automated embryo assessment software.

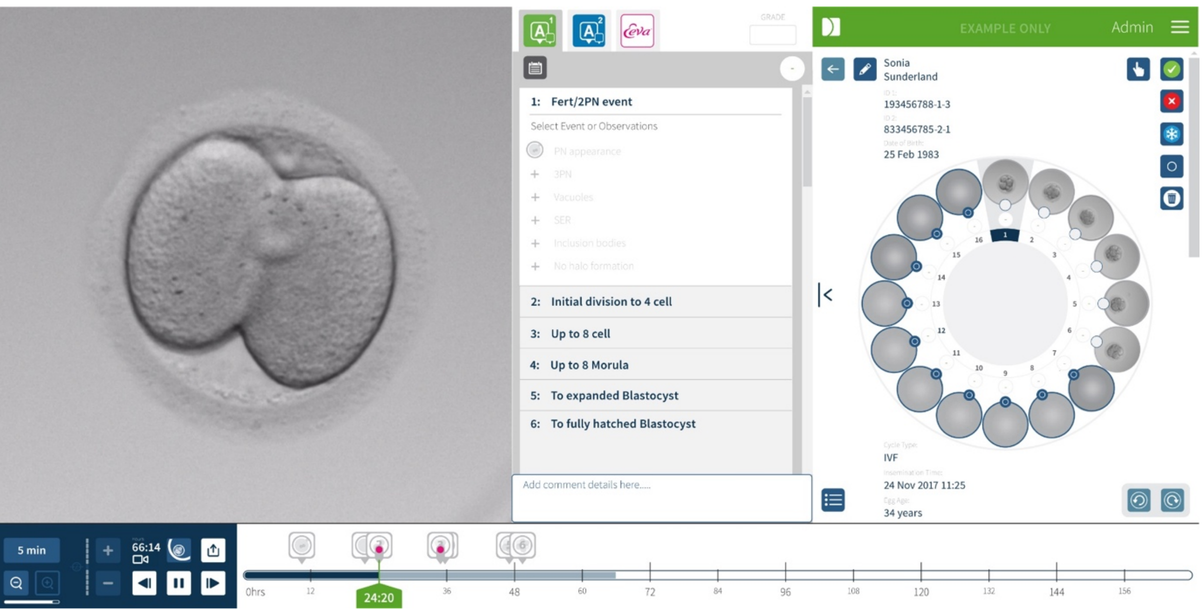
1.2 Analysis

This retrospective study was performed with data collected between February 2016 and January 2019 in IVIRMA Valencia clinic. Cycles from a total of 284 patients undergoing IVF treatments were cultured in the Geri TLS. Embryos were analysed manually by the embryologist team and automatically by Geri® Assess 2.0 annotation software. A busy embryologist team at IVIRMA clinics annotated embryos for 10 developmental events as per their normal clinical practice using Geri® Assess 1.3 software (**Figure 6, A**). The same videos were then analysed retrospectively with the stand-alone Geri® Assess 2.0 software (**Figure 6, B**). All videos of embryos developing up to the blastocyst stage were manually downloaded from the Geri platform and uploaded to the automated platform to obtain their annotations.

For the **Detection Rate Analysis**, the percentage of events detected by either the embryologist team, the software or both were calculated. Detection rate refers to annotation rate, or the percentage of events that are annotated when they actually happen. Next, the software applied the pre-determined filtering tool for the out-of-range annotations incorporated in the Geri® Assess 2.0 system. Both annotations, filtered and unfiltered were kept for the descriptive analysis.

For the **Accuracy Analysis**, we subtracted both annotations to calculate the difference of the embryologist's times minus the Geri Assess 2.0 times. For this purpose, if a specific event was annotated at the same time by both methods, finding no difference, we could be very confident the event took place at that time. However, if the annotated times were far apart, we would not be able to tell which one was the correct one, or if none did a good job. For this accuracy, we calculated the mean and standard deviation for each parameter. Datasets were also compared before and after the filtering was applied.

A



B

Figure 6. *Patient Review page in Geri incubator.* ***(A)****. Geri Assess 1.3 tab, to manually annotate events & observations.* ***(B)****. Geri Assess 2.0, automatically annotates events & observations.*

1.3 Results

Cycles from a total of 284 patients undergoing IVF treatments were cultured in the Geri TLS. Following embryology standard practice, only those embryos arriving to the blastocyst stage, therefore having real utilisation, were selected for the initial analysis. 1,402 blastocysts were identified by the embryologist team in their routine clinical practice using Gardner and Schoolcraft's (1999) blastocyst evaluation criteria. Both, manual and automated annotations for these embryos were analysed. Specifically, the annotations for the following developmental events: tPNa, tPNf, t2, t3, t4, t5, t6, tM, tSB, tHBi; described in **Table 1**.

The initial analysis of the missing data, described as non-annotated events, revealed missing annotations were not evenly distributed between stages. **Figure 7** represent the result of the missing data by event (**A**, missing data by stage and method) and by embryo analyzed (**B**, missing data by Link).

When focusing on the missing data by stage (**Figure 7, A**), we realized a large proportion of tPNa and tHBi were not annotated. At the time of the study, our group was working on a parallel study validating the EEVA system following the study of Aparicio-Ruiz *et al.* (2016), therefore some videos of those embryo were starting only after the time of pronuclei appearance. In fairness to the study results, we decided to remove tPNa from our analysis. A similar phenomenon occurred with tHBi. Not all the blastocyst analyzed were arriving to the initiation of hatching stage, which is natural as most are vitrified or transferred before this stage. For this reason, the number of non-annotations in this stage was not realistic as we cannot assume it was missing as an annotation, maybe it never occurred. Therefore, we decided to remove tHBi from the analysis as well. Likewise in the rest of events, where the software presented fewer missing data overall in all stages.

Figure 7. *Missing Data Analysis described as non-annotated events (NA).* ***(A)*** *Missing Data by stage and method represents the amount of non-annotated events per stage.* ***(B)*** *Missing data by Link represents the amount of non-annotated events per embryo analyzed.*

On the other hand, when focusing on the Missing data by embryo (**Figure 7, B**) we could observe that most of the embryos, in both annotation groups, had only one missing annotation. When we compare both groups, we can observe that Geri 2.0 annotations had double the number of embryos without missing any annotation as compared to the embryologist group. As the amount of non-annotated events per embryo were increasing, equal or higher than 1, the embryologist team had more missing annotations than the automated software.

In the **Detection Rate Analysis**, the final developmental events analysed were: tPNf, t2, t3, t4, t5, t6, tM and tSB. As the total number of embryos analysed were 1,402, we could possibly detect a total of 11,216 events. **Table 5** summarizes the result of the Detection Rate Analysis for both groups, Geri Assess 2.0 automated software vs. a busy team of embryologists. Both datasets were depicted before and after applying the filtering tool for the out-of-range annotations, described in **Table 6**.

From 11,216 possible developmental events, the embryologist detected an 89.6% (10,054), and Geri® Assess 2.0 a 98.4% (11,035). Time ranges for filtering the annotations excluded data points which represented a 13.5% of the Geri Assess 2.0 annotated events and 8.1% of the manually annotated events. Before filtering, Geri Assess 2.0 presented a higher detection rate in all developmental events, however, after filtering, stages t4, t5 and t6 showed a higher detection rate in the Embryologist's annotations. The lowest concordance rates in detection between groups were found in the pronuclei disappearance, time of morula and time for start of blastulation. After the filtering tSB became very similar between groups.

Table 5. *Detection Rate Analysis of the morphokinetic developmental events by the Geri Assess 2.0 vs. a busy embryologist team in their routine clinical practice.*

		tPNf	t2	t3	t4	t5	t6	tM	tSB	All events
Unfiltered	Geri Assess 2.0	1399	1398	1398	1399	1396	1393	1348	1304	11035
		99.8%	99.7%	99.7%	99.8%	99.6%	99.4%	96.1%	93.0%	98.4%
	IVIRMA Embryologists	1208	1321	1280	1337	1287	1286	1184	1151	10054
		86.2%	94.2%	91.3%	95.4%	91.8%	91.7%	84.5%	82.1%	89.6%
Filtered	Geri Assess 2.0	1325	1362	1182	1212	1153	1055	1109	1148	9546
		94.5%	97.1%	84.3%	86.4%	82.2%	75.2%	79.1%	81.9%	85.1%
	IVIRMA Embryologists	1175	1310	1151	1233	1183	1068	1016	1108	9244
		83.8%	93.4%	82.1%	87.9%	84.4%	76.2%	72.5%	79.0%	82.4%

Table 6. *Summary of filtered out-of-range annotations by embryo developmental event.*

		tPNf	t2	t3	t4	t5	t6	tM	tSB	All events
Filtered Data	**Geri Assess 2.0**	74	36	216	187	243	338	239	156	1489
		5.3%	2.6%	15.5%	13.4%	17.4%	24.3%	17.7%	12.0%	13.5%
	IVIRMA Embryologists	33	11	129	104	104	218	168	43	810
		2.7%	0.8%	10.1%	7.8%	8.1%	17.0%	14.2%	3.7%	8.1%

Having a closer look to the data after applying the filter (**Table 6**), we could observe that the detection rate by Geri Assess 2.0 maintained a higher rates in the early events, but decreased substantially, especially in time to 5 cells, 6 cells and time to morula. Manual annotations were also modified by the filtering, finding t6 and tM as the most filtered events. Hence, it was decided to examine how the filter was modifying the distribution of our data, and time to 6 cells was chosen as it was highly filtered in both annotation techniques. Data distribution for t6 in both techniques and superposed data is represented in **Figure 8.**

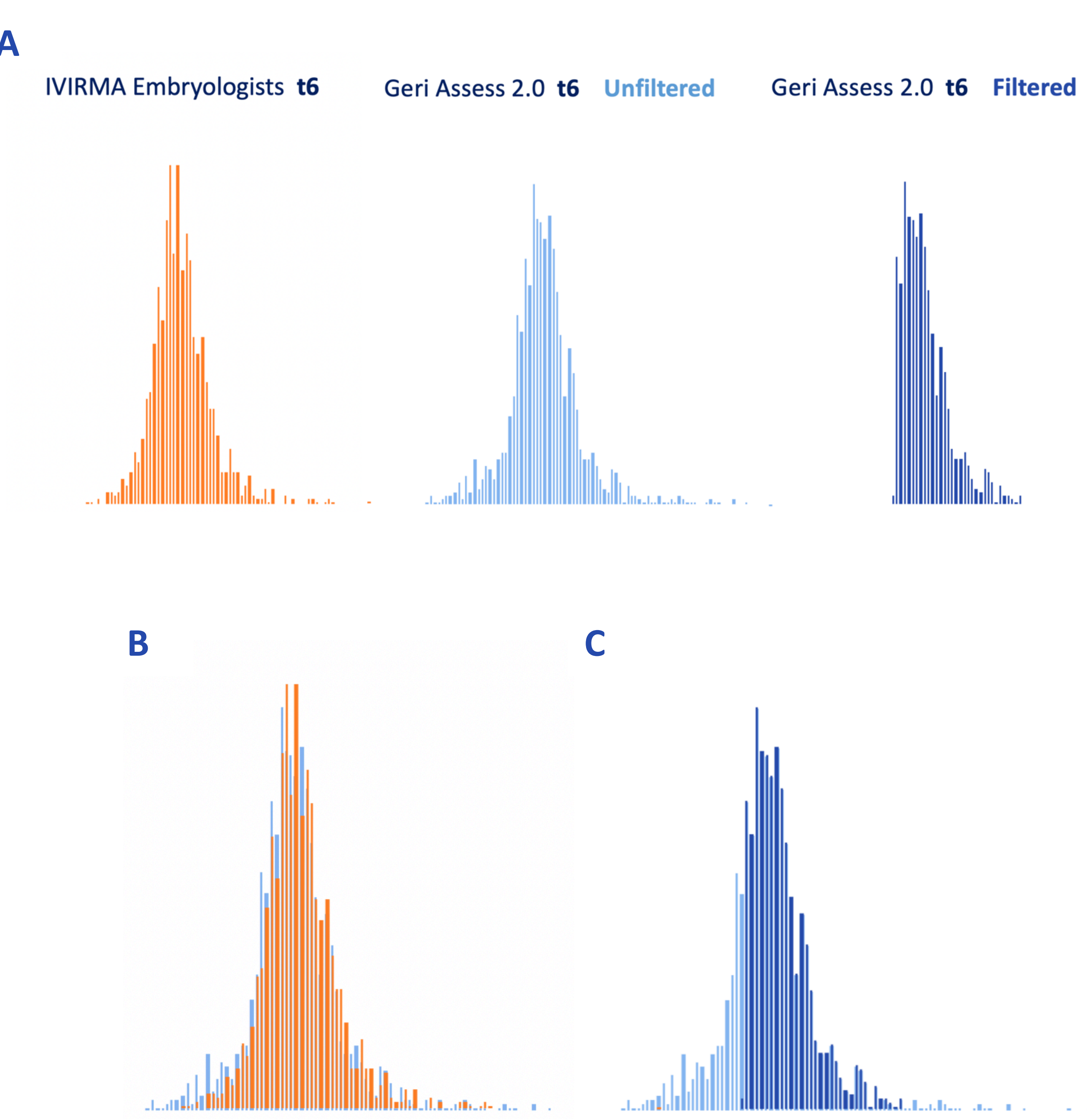


Figure 8. *(**A**) Data distribution of the annotations performed for time to 6 cells (t6) by the IVIRMA embryologists' team and the Geri Assess 2.0 before and after applying the filtering tool for the out-of-range annotations. (**B**) Superposed data distribution for both annotation techniques before filtering. (**C**) Superposed data distribution for Geri Assess 2.0 annotation before and after filtering.*

Data distribution for t6 showed annotations between both techniques were similarly distributed. Times were slightly more spread out in Geri Assess 2.0 to the extremes, especially towards earlier timings (**Figure 8**, **B**). When applying the filter (**Figure 8**, **C**), a large number of datapoints were excluded in the earlier annotated t6 values. As described in **Table 6**, they represented nearly a quarter of the annotations. These datapoints were also excluded in the manual annotations, representing a 17%. The data distribution comparison between annotations for the rest of developmental events can be observed in **Supplemental Figure 3**. The reduction of the dispersion of timepoints into the extremes was meant to increase the data accuracy, taking us to our next analysis.

Moving on to the **Accuracy Analysis**, **Table 7** summarizes the results formatted as mean in hours (± standard deviation), with the unfiltered and filtered datasets. Differences in timings varied according to the developmental event. The largest difference was detected in the later events, time to morula (tM) and time to early blastocyst (tSB). Nevertheless, the majority of events didn't show highly significant differences, especially early cleavage divisions, where the match rate was very high. When focusing on the adjustment of data after the filtering, we could observe the most altered accuracy mean was tSB.

Table 7. *Accuracy of the manual and automated annotations at the different developmental events, with the unfiltered and filtered dataset.*

		tPNf	t2	t3	t4	t5	t6	tM	tSB
Unfiltered	Mean (± Std Dev) (hrs)	1.1 (± 3.5)	0.7 (± 2.8)	1.0 (± 3.5)	1.2 (± 5.2)	2.2 (± 5.8)	1.0 (± 6.3)	5.0 (± 11.6)	3.0 (± 9.0)
Filtered	Mean (± Std Dev) (hrs)	0.6 (± 1.1)	0.6 (± 1.3)	0.5 (± 2.1)	0.7 (± 4.1)	1.1 (± 4.3)	0.1 (± 4.4)	4.1 (± 7.4)	1.1 (± 5.4)

Another important aspect to consider in the **Accuracy Analysis** was the standard deviation of the different developmental events. Our aim was to measure the data dispersion in relation to the optimum value of zero. This is numerically described in **Table 7** and graphically depicted in **Figure 9**, for each developmental event. The highest dispersion was observed in the later developmental events, tM and tSB, followed by t4, t5 and t6. The data standard deviation comparison before and after filtering the data can be also observed in **Supplemental figure 4**.

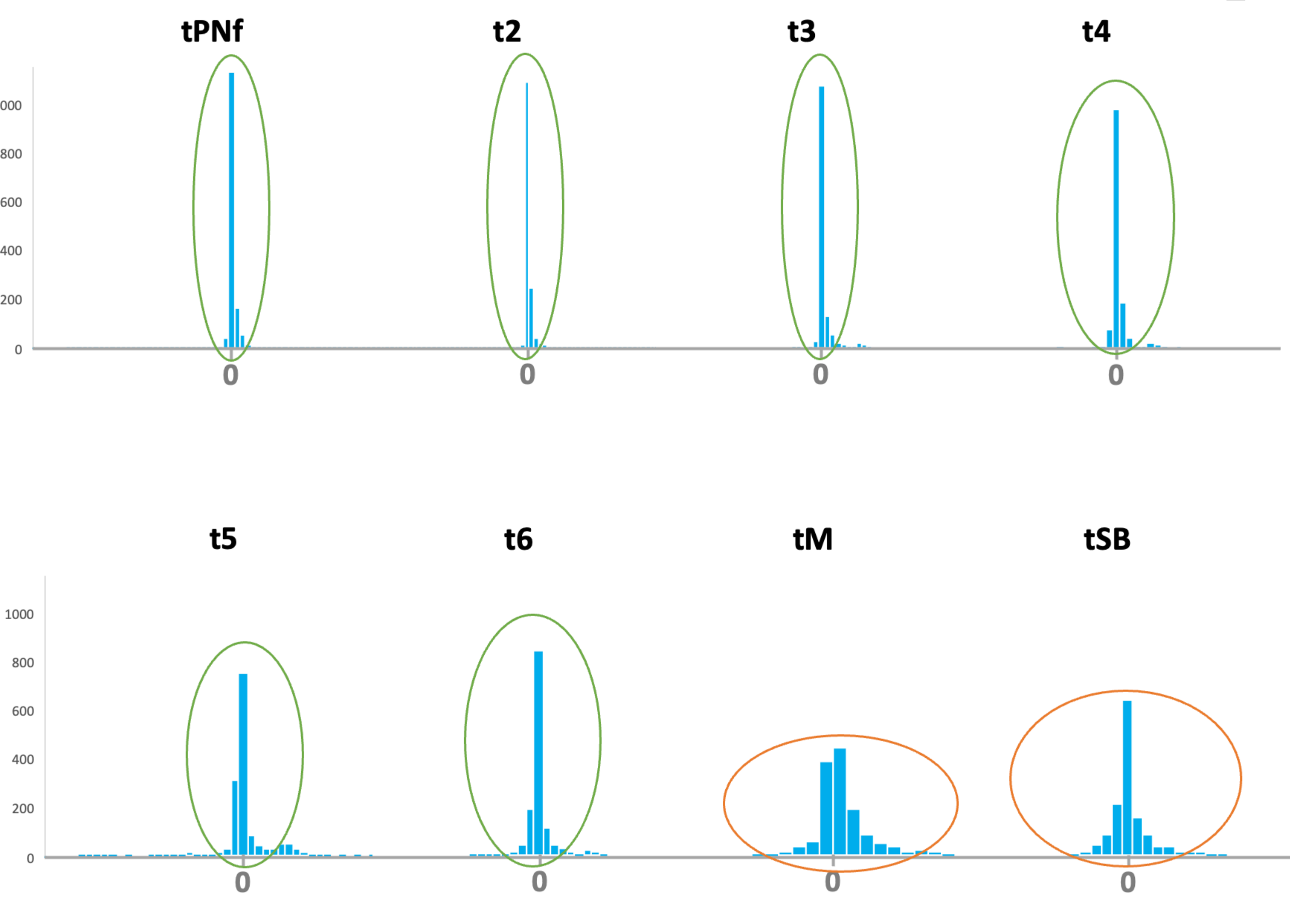


Figure 9. *Standard deviation calculated between the manual and automated annotations at the different embryo developmental stages.*

1.4 Discussion

The development of automated annotation software is essential in assisted reproduction clinics to facilitate, standardize, and improve the annotation process of the morphokinetic parameters. The main objective for its development consists in assisting the embryologists in their routine clinical practice. Subsequently, after a rigorous validation and correct deployment, the second objective consists in the improvement of embryo selection process for transfer.

The **Specific Objective I** aimed to compare the manual annotations performed by a busy embryologist team with the software Geri Assess 1.3 versus the automated annotations performed by Geri Assess 2.0. To reach this goal a missing data analysis, detection rate analysis and an accuracy analysis of both annotations were performed. 1,402 embryos, whose development continued up to the blastocyst stage, were assessed for the comparison of both morphokinetic annotations. A point worth highlighting was the comparison of the software with a busy embryologist team, aiming for a real validation, but different from previously published studies. Our results support the use of the Geri Assess 2.0 software as comparable to the embryologists' performance.

A major defect in our study was to drop from our analysis tPNa and tHBi. From tPNa we can assume the detection rate was very similar between groups, as a very similar number of missing annotations are represented in **Figure 7**, constituting the videos starting only after the appearance of the pronuclei. On the contrary, if we observe the tHBi, we can see a big difference between the number of automated and manual annotations missing, assuming the initiation of hatching may have happened in many embryos, but its annotation was missed by the embryologists. To properly perform this analysis, we should identify the subgroup of embryos with confirmed arrival to the tHBi stage and proceed with the comparison analysis. A positive conclusion of the missing data analysis was that it was rare to have embryos with more than 2 missing datapoints, being 1 the most common.

Continuing with missing datapoints, Geri Assess 2.0 software presents a major weakness, which is the inability to detect other morphokinetic events like t7 and t8. Especially

t8 has evidenced its importance becoming part of many ESAs as a predictive timepoint for blastocyst formation and implantation potential in several publications (Chamayou *et al.*, 2013; Goodman *et al.*, 2016; Motato *et al.*, 2016). This flaw is due to the challenge of a proper image analysis as the number of cells in the embryo increases. To overcome this, the image analysis should be performed in a 3D model considering the different focal planes of the embryo, and not with the best centered focal plane image, as it is currently performed. In addition, the time to expanded blastocyst (tEB) was under development at the time of the study and should be further analyzed.

As described in the **Detection Analysis**, the automated annotations software had a comparable detection rate to the embryologists' annotations in all the developmental events, outperforming in the later stages of development. Detection rates remained constant, never descending under 90%, as opposed to the embryologist team where lower than 90% rates were found in tPNf, tM and tSB. This difference could be explained not because the embryologists were not able to detect a developmental event, but due to the busy environment with a heavy workload experienced in IVF Laboratories (Paternot *et al.*, 2011). The lowest accordance rates in detection were found in the more subjective events: PN disappearance, time to morula and time to starting blastocyst, similar to the previous results presented by Sundvall *et al.* (2013) and Martínez-Granados *et al.* (2017). Clinically speaking, these points can be defined as events of great clinical importance, where we could assume the embryologists are focusing more on the count of the number of pronuclei and the assessment of the blastocyst for its vitrification, transfer, or biopsy, as a priority over the morphokinetic annotations.

A similar difference in the accordance rates between annotations was kept after the filtering of the events according to their optimal time-ranges, although attenuated. Time-ranges for filtering of annotations were: tPNf = 17-30 h, t2 = 20-40 h, t3 = 30-48 h, t4 = 32-54 h, t5 = 38-68 h, t6 = 46-78 h, tM = 64-100 h and tSB = 86-126 h, established based on published recommended reviewed times (ALPHA Scientists in Reproduction medicine and ESHRE SIG of

Embryology, 2011; Ciray *et al.*, 2014). Excluded data points represented a 13.5% of the Geri Assess 2.0 annotated events and 8.1% of the manually annotated events. Although the functionality of the filter is to exclude annotations that are biologically impossible for the given event, it was surprising than an 8.1% of the manual annotations were excluded, as they were performed by an experienced senior embryologist team. As described in **Table 6**, over a 10% of the t3, t6 and tM annotations were excluded in the embryologist team.

When the data distributions were closely analyzed before and after the filtering (**Figure 8**), we could observe the distribution of the unfiltered data was quite similar to the embryologist annotations. However, when the filter was applied it seemed a bit too strict at the early annotations, thereby eliminating numerous datapoints in both groups. This raised doubts on the accuracy of the time ranges selected for the filter. Maybe the filters should be readjusted, not to eliminate useful annotations.

Regarding the **Accuracy Analysis** of the software, values obtained were all pretty close to 0 and the standard deviations were not too spread out. Time to morula and time to starting blastulation showed the most different times between groups. This makes sense as they are the most subjective parameters, so it could be due to the variability between observers annotating them.

In the present thesis, the retrospective analysis of 1,402 embryos led us to the conclusion that the performance of the Geri Assess 2.0 software is comparable to the embryologists' manual annotations. Automated annotations can help the embryologists in their daily workload, especially early events were proven to be annotated with high accuracy. Furthermore, the filtering tool would highlight and eliminate out-of-range events to be thoroughly assessed, increasing the accuracy of the data. However, as previously suggested, the filtering time ranges need a further examination for a readjustment, not to eliminate useful datapoints.

A prospective validation would be useful to corroborate the results of the present thesis, including all kinds of embryos, not only the ones arriving to the blastocyst stage, as the ones used in the current study. Towards this end, we have observed chaotic embryos with aberrant divisions and fragmentation makes exact annotations very difficult to both, manual and automated annotations. However, it is unlikely that those embryos will be up for selection in the final stage of development. In any case, for those wrongly or non-annotated events we can always add them manually, giving the experienced embryologist the final word. Moving into the use of automated annotations is a natural progression for clinics utilizing time-lapse systems, such as ours.

SPECIFIC OBJECTIVE II.

Clinical result prediction comparison applying manual and automated development annotations of Known Implantation Data (KID) embryos in previously published Embryo Selection Algorithms (ESAs).

2.1 Purpose

The **Specific Objective II** aimed to compare embryo grading and clinical result prediction obtained with a morphokinetic algorithm using an automated system for embryo developmental events annotations compared to the manual annotations performed by an embryologist team. **Figure 10** summarizes the objective of the study: Would we select the same embryo for transfer using both annotation methods?

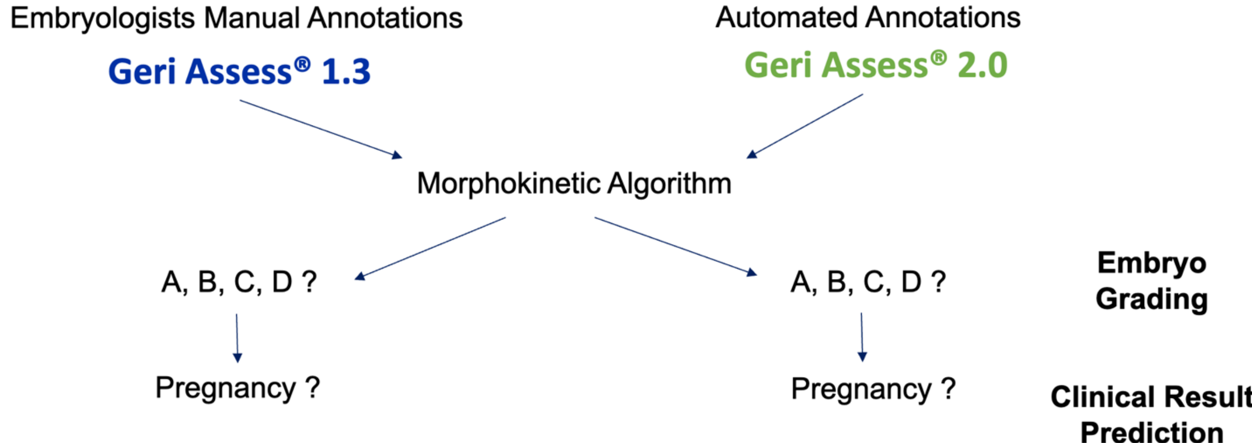


Figure 10. *Summary of Specific Objective II.*

2.2 Analysis

This study consists in the retrospective examination of the morphokinetic manual and automated data from 1,485 embryos transferred between February 2016 and May 2019 at IVIRMA Valencia clinic. All embryos were normally fertilized embryos cultured up to day 5 or 6 of development. All were included regardless of their cycle type (egg donation program or patient's autologous cycle), oocyte origin (fresh or frozen) or patient age (range: 27-44 years).

All embryos analyzed were annotated manually by a busy embryologist team in the routine clinical practice using Geri Assess® 1.3 software. The same videos were retrospectively assessed by the stand-alone Geri Assess® 2.0 software, including filtration of events falling outside the pre-defined time-ranges, as is done in the full Geri system. Both morphokinetic manual and automated annotations went through Embryo Selection Algorithms previously published. The reason behind the selection of specific ESAs is explained in the results & discussion. Embryos were graded and the accuracy in the prediction was assessed between both groups in terms of embryo outcome, hCG test, and fetal heartbeat. Data was statistically analyzed with chi-squared and binomial proportion tests. Statistical analysis was performed using the Statistical Package for the Social Sciences 22 (SPSS Inc.).

2.3 Results

In order to analyze the **Specific Objective II**, we decided to start with one of the latest ESA developed by our IVIRMA team (Basile *et al.*, 2015) predictive of Implantation Potential. **Figure 11** represents the hierarchical classification performed by the ESA, considering the morphokinetic parameters t3, cc2 (t3-t2) and t5. Qualitative parameters were not considered as they required a subjective analysis.

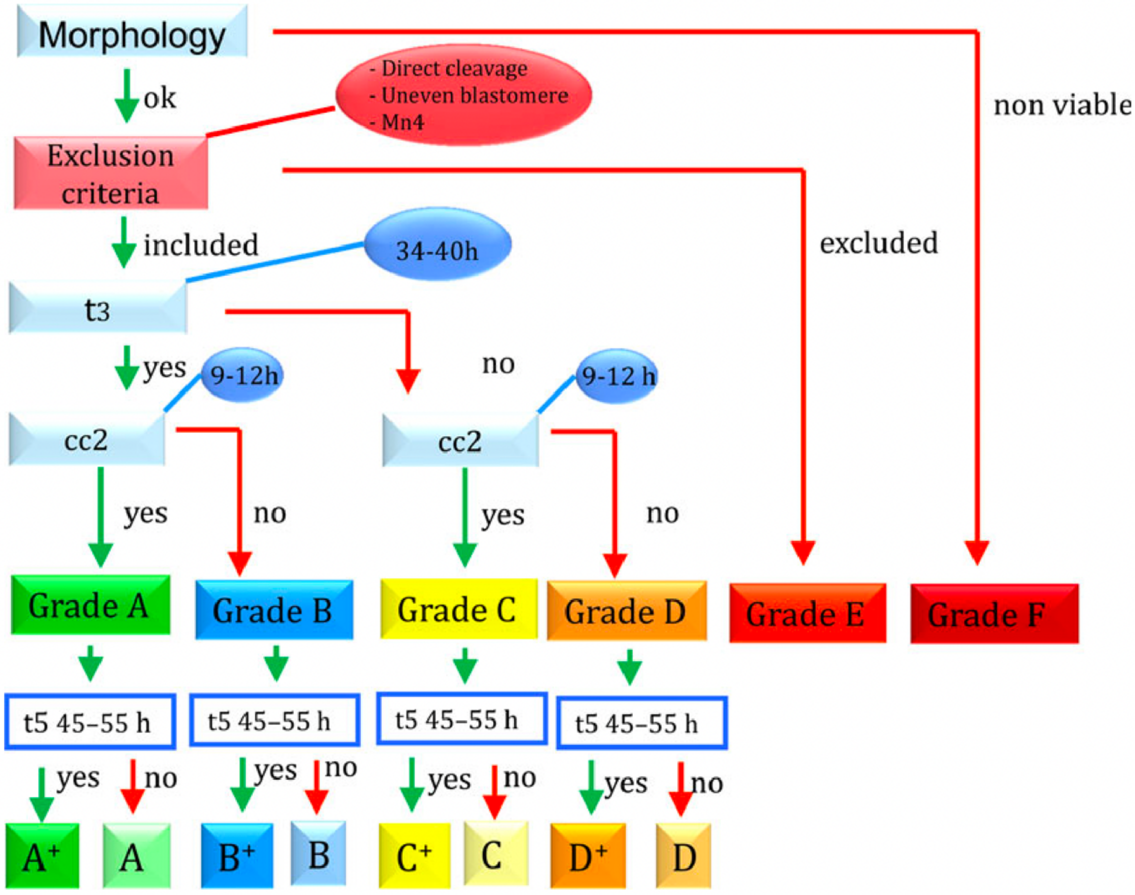


Figure 11. *Hierarchical classification of embryos based on: (i) Morphological screening; (ii) the new morphological criteria; (iii) timing of cell division to three cells (t3); (iv) duration of second cell cycle (cc2), i.e., the time from division to a two blastomere until division to a three blastomere embryo; (v) timing of cell division to five cells (t5). The classification generates 10 categories of embryos with increasing expected implantation potential (right to left) and almost equal number of embryos in each (image from Basile* et al.*, 2015).*

High accordance was found between the embryo grading using both annotation sets through Basile *et al.*'s (2015) algorithm. Out of the 1,485 embryos included in the analysis, blastocysts utilized for transfer or vitrification, there were no statistically significant differences between groups in all grades: A+, A, B+, B, C+, C, D+ and D; except for No Grade ($p < 0.05$), as shown in **Figure 12**. More ungraded embryos were found in the automated group, as Geri Assess® 2.0 was eliminating more events falling outside of pre-defined time-ranges, through the filtering tool. t3 was the most unavailable parameter in the automated annotation's data. Regarding only the KID transferred embryos, β-hCG test and fetal heartbeat data also did not show statistically significant results between both groups.

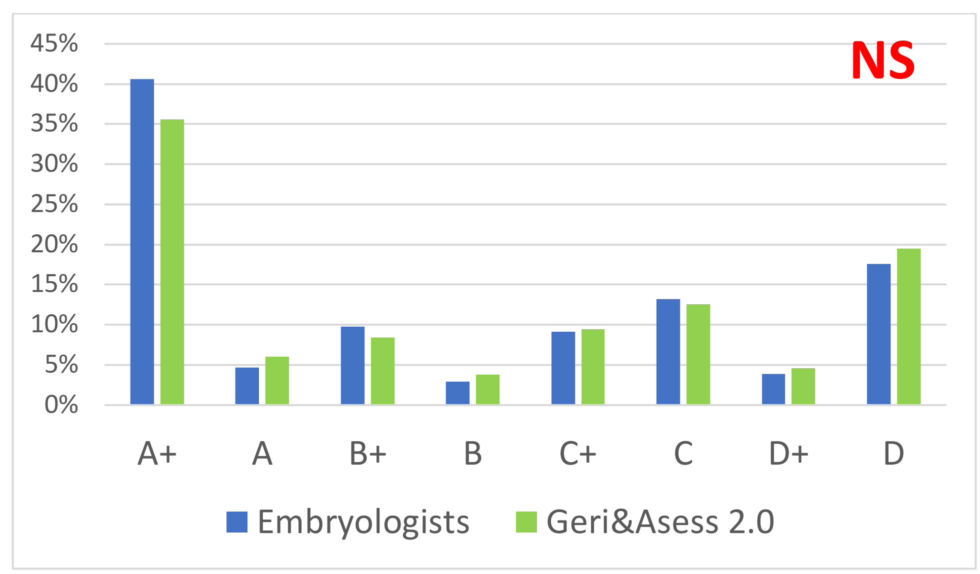


Figure 12. *Embryo grading classification after using the morphokinetic manual annotations performed by the embryologist team and the automated ones performed by the software. NS = Not Significant.*

The lack of significance between the groups with the ESA by Basile *et al.* (2015), made us move on to ESAs whose grading is performed later in development, at the blastocyst stage. However, the technical restrictions of the automated software presented a limitation, as not all the morphokinetic events were detected by the software. Table 8 summarizes the published ESA whose application happens later in development. As depicted in the Table, most of the ESAs required t8 as a quantitative parameter, as the parameter s3 equals t8 minus t5 (Goodman *et al.*, 2016; Motato *et al.*, 2016). Neither t8 nor tB, required by Campbell *et al.'s* (2013a) algorithm, can be automatically detected by the software, as it has been not trained to do so. Therefore, the only ESA we could apply for the purpose of the objective was Rienzi *et al.* (2019).

***Table 8**. Embryo Selection Algorithms (ESAs) with application at the blastocyst stage.*

		Selection parameters				
Published ESAs	Sample size	Qualitative	Quantitative	Predictive outcome	ESA day	Study type
Motato *et al*., 2016	7483 emb	-	tM s3	Blastocyst Formation	4	Single-centre
Goodman *et al*., 2016	94 KID	MN AC	cc2 t5 s2 s3 tSB	Implantation Potential	5–6	Single-centre
Motato *et al*., 2016	832 KID	-	tEB s3		5–6	Single-centre
Campbell *et al*., 2013a	98 blas	-	tB tSB		5–6	Single-centre
Rienzi *et al*., 2019	830 blas	TE quality	tM	Live birth rate	5–6	Multi-centre

*Note. Modification of **Table 2**.*

As described in **Figure 13**, Rienzi *et al.* (2019) ESA is based on the trophectoderm (TE) quality at time of blastulation (tB) and whether the time to morula (tM) is lower or higher/equal to 80 hours post-insemination. The endpoint used for the training of this algorithm was Live Birth Rate, in euploid single embryo transfers (SET). A tM lower than 80 hours was a good predictor for live birth.

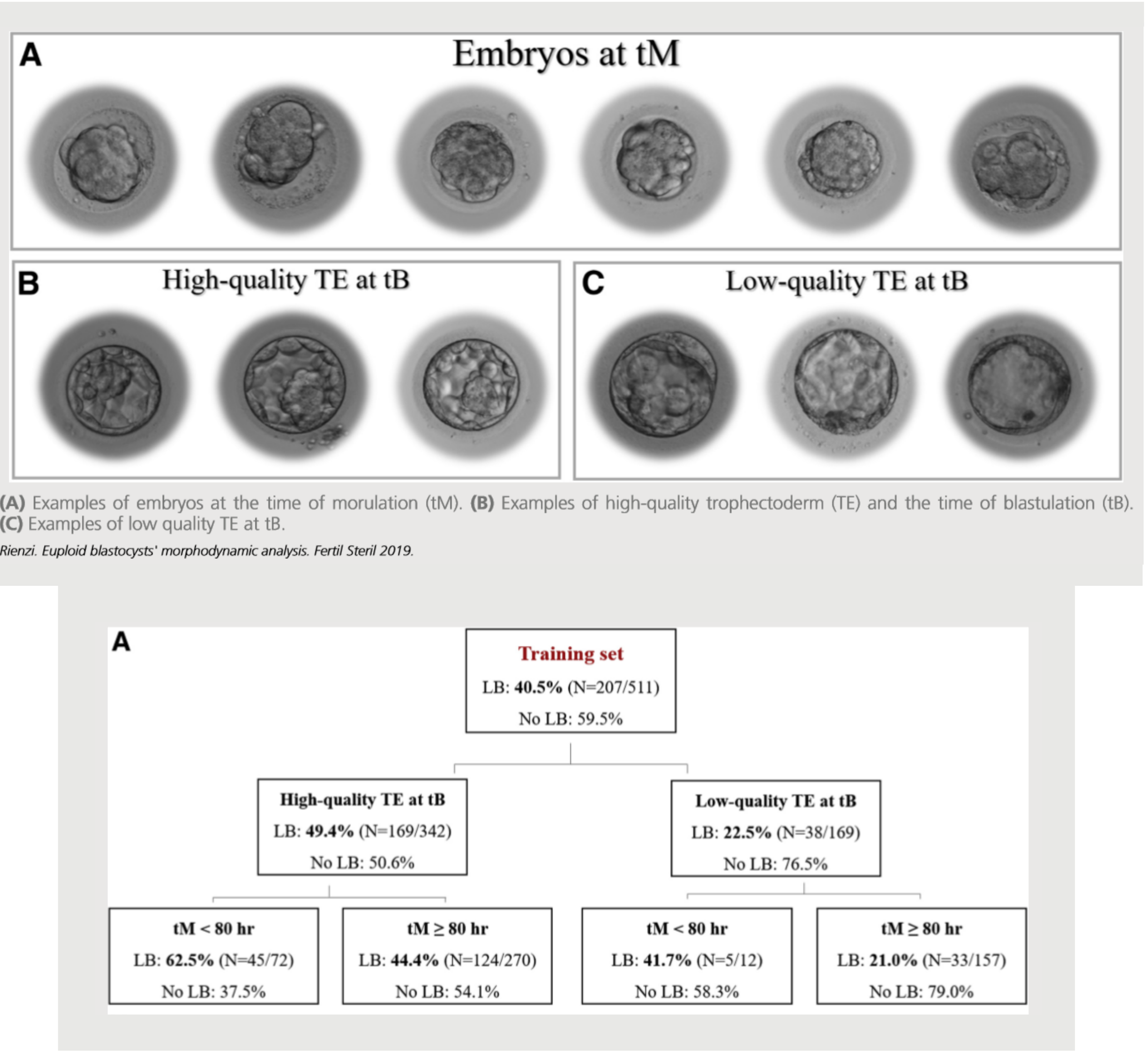


Figure 13. *Classification of embryos according to Rienzi et al. (2019) publication, based on the trophectoderm (TE) quality at time of blastulation (tB) and the time to morula (tM) (image from* Rienzi *et al.*, 2019)*.*

The application of both annotations datasets, manual *vs.* automated, of 1,493 blastocysts through Rienzi *et al.*'s (2019) algorithm revealed both groups were annotating time to morula very differently (**Figure 14**). Overall, the embryologists were detecting tM a bit later than the automated system, categorizing more embryos in the tM equal or over 80 hours group. The assessment continued to observe if this grading difference was affecting the clinical result prediction after transfer. Implantation Rate and Ongoing Pregnancy Rate were the endpoints selected for our analysis (**Figure 15**). All the blastocysts underwent SET with KID outcome.

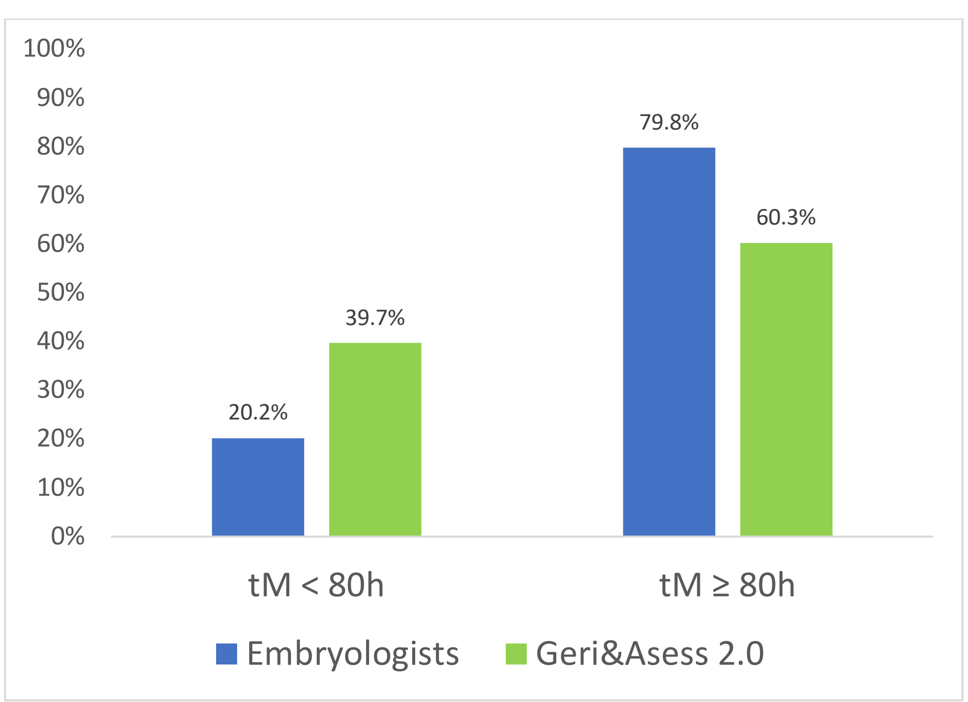
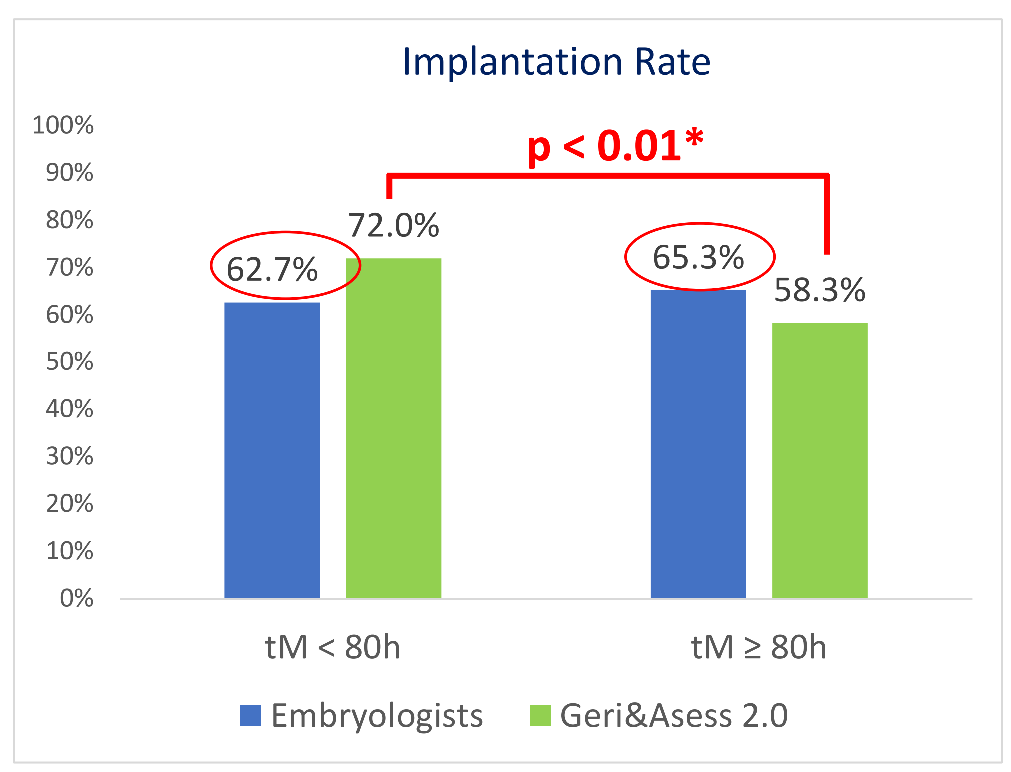


Figure 14. *Blastocyst classification using* Rienzi *et al.*'s (2019) *tM cut-off value with both, the embryologist's manual, and the software's automated annotations.*



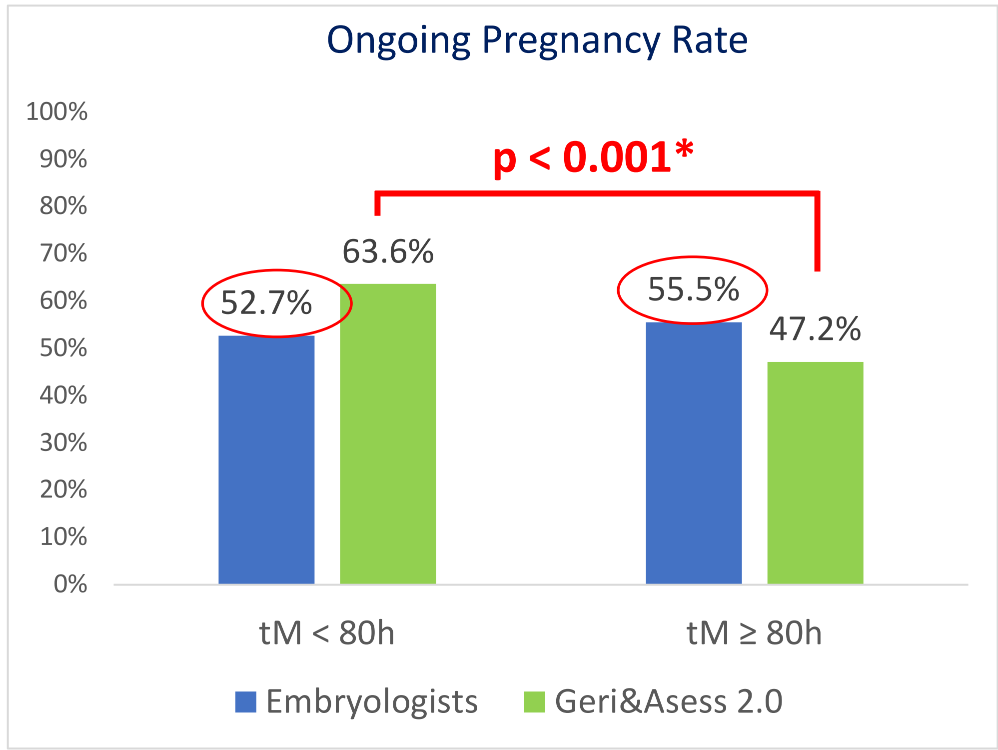


Figure 15. *Blastocyst clinical rate prediction using* Rienzi *et al.*'s (2019) *tM cut-off value with both, the embryologist's manual, and the software's automated annotations.*
** = Significant difference.*

Figure 15 shows a significancy, with over 10% and 15% difference, when analyzing Implantation Rate and Ongoing Pregnancy Rate respectively, when the automated annotations were classified using Rienzi *et al.*'s (2019) algorithm. On the other hand, neither of the success rates were significantly different when using the embryologist's manual annotations.

2.4 Discussion

In the **Specific Objective II**, we applied the same morphokinetic algorithm to both manual and automated annotations to check if embryos would obtain the same grade, and whether this difference in grading affected the clinical result prediction. Our objective was to evaluate whether the automated annotations could affect our clinical results.

Starting off with the discussion of the ESAs selected, our main goal was to perform the embryo grading and prediction as objective as possible. For this reason, we left the qualitative selection parameters out of the question, as they can still only be subjectively analyzed by experienced embryologist. One of the reasons behind the start of the analysis with Basile *et al.'s* (2015) algorithm was because the quantitative parameters used were common significant parameters in many other published ESAs: the duration of the second round of cleavage (cc2 = t3-t2), also called P2 (Meseguer *et al.*, 2011; Vermilyea *et al.*, 2014; Goodman *et al.*, 2016; Liu *et al.*, 2016; Petersen *et al.*, 2016), the time to 3 cells (Meseguer *et al.*, 2011; Chamayou *et al.*, 2013; Vermilyea *et al.*, 2014; Goodman *et al.*, 2016; Liu *et al.*, 2016; Petersen *et al.*, 2016), and the time to 5 cells (Meseguer *et al.*, 2011; Chamayou *et al.*, 2013; Goodman *et al.*, 2016; Liu *et al.*, 2016; Motato *et al.*, 2016; Petersen *et al.*, 2016).

The high accordance found between the embryo grading using both annotation sets through Basile *et al.*'s (2015) algorithm was expected as the parameters used in this ESA were early events in development. The previous study in **Specific Objective I** proved a high detection rate and accuracy in the great majority of the events detected by the automated software, especially in early cleavage divisions. The parameters t3, cc2 and t5 must have been annotated similarly by both groups and consequently show a similar performance with the algorithm. However, this model only assesses embryos until day 3 of development and therefore not representative of the reality of culture reaching the formation of the blastocyst. For this reason, we decided to focus on ESAs that were developed with data up to the blastocyst stage. Advances in embryo culture has led to a shift in the IVF practice from early cleavage to blastocyst stage. The rationale is to improve uterine and embryonic synchronicity and to improve the selection of those embryos surviving up to a later developmental stage (Glujovsky *et al.*, 2016).

The decision on the blastocyst stage ESA to use was limited to the ones using morphokinetic parameters detected by the software. As explained by **Table 8**, four ESAs were discarded as they were using t8 and tB, parameters not available in the software's automated annotations. In the case of t8, the higher the number of blastomeres in the embryo, the more difficult it is for the system to detect them without moving the focal plane, as embryologists do. This is a clear limitation of the software; the automated detection can only happen up to 6 blastomeres. After an extensive search of algorithms deployed at the blastocyst stage, the only ESA we could apply for the purpose of the objective was Rienzi et al. (2019).

Time to morula, with a cut-off value of 80 hours, demonstrated to be a good predictive parameter when using the automated annotations of Geri Assess 2.0. The automated annotations were more helpful at a clinical level when used together with this morphokinetic algorithm than the embryologist manual annotations. These findings correlated with our previous **Accuracy Analysis** finding tM as the most differently annotated between groups. This might be influencing the higher proportion of blastocyst in the tM>80 group by the embryologist team. A possible explanation could be the subjective nature of the parameter, being differently annotated between individuals or even by the same embryologist (Sundvall *et al.*, 2013; Martínez-Granados *et al.*, 2017; Adolfsson and Andershed, 2018). tM has been described to be difficult to assess its real time frame, as there are cases where some blastomeres are excluded from compaction (ESHRE Working group on Time-lapse technology *et al.*, 2020). This may influence the hesitation in our human nature, where in order to be certain about the event happening, we may delay a few frames its annotation. The software, in the other hand, with an objective image analysis, may be standardizing its annotation.

When we investigate the presence of time to morula in the literature it is surprising to observe it has not been a parameter commonly used in algorithms. Early studies such as Chamayou *et al.* (2013) and Kirkegaard *et al.* (2013) did not find a significant difference with this parameter in blastocyst formation or implantation. However, later, Kramer *et al.* (2014) stated the only parameter that yielded any predictive value towards euploidy was the duration of compaction (tSB-tSC), although the AUC was not that high (0.674). Next, Storr *et*

al. (2015) described a shorter tM to be correlated with top-quality blastocyst morphology. The optimal cutoff value for this parameter was <97.33. Motato *et al.* (2016) also found a higher correlation with blastocyst formation, when the time was falling in its optimal range from 81.28 to 96.00 hours. A study by Mizobe *et al.*, (2017) outlined those embryos that completed compaction within 79.93 hours had a high implantation potential. And finally Rienzi *et al.* (2019) study, used in the present validation, which strength was the use of only euploid blastocyst followed by single embryo transfers.

If we look deeper into the morula stage, you realize how important this stage is in development. **Figure 16** shows the time-lapse progression from the cellular stage through compaction, up to the morula stage. Compaction starts after the flattening and formation of adherens and tight junctions between the blastomeres and the organization of filopodia (Iwata *et al.*, 2014; Coticchio *et al.*, 2019). Throughout the development, there is an increase in the transcription and translation processes, resulting from the genome activation happening in the 4 to 8 cell transition (Braude *et al.*, 1988). Cell fate definition is coordinated and there is an activation of gene expression pathways committing cells into different developmental destinies (Jedrusik, 2015). This differentiation ends up dividing the embryo into the polarized trophectoderm (TE), which will form the placenta, and the apolar inner cell mass (ICM), later the fetus and the yolk sac (Elder and Dale, 2011; Meng *et al.*, 2020). All these crucial processes, together with our study, should be further investigated to ultimately understand the way embryos acquire their reproductive competence during compaction.

Regardless of the parameter to be used for the prediction, this can only be performed in the transferred embryos that present the morphokinetic annotations. If the morphokinetic event is missing, as it has not been annotated, there is no possible prediction to be made, and as shown in **Specific Objective I**, there is an 8.8% difference of events annotated, being less in the manual annotation group. If we are going to choose tM as our predictive parameter, this difference increases to a 11.6% of cases not annotated manually in the routine practice. Therefore, even with no time difference between the annotations, we can at least obtain predictions on a higher number of embryos in the software group.

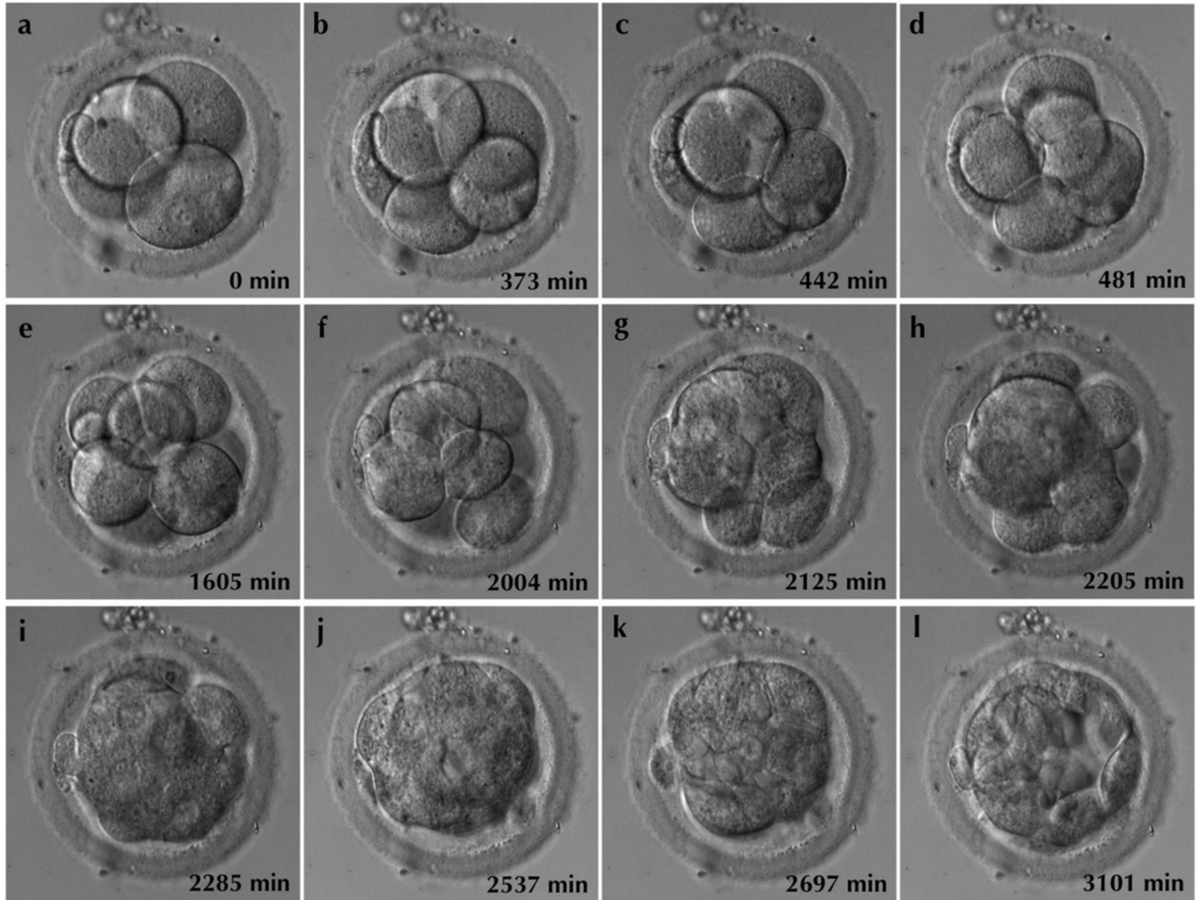


Figure 16. *Compaction in the human embryo. After several cell divisions (a–e), the blastomeres became flattened (f), and the intercellular boundaries became obscured (g–i), until they finally unified in one cluster (j,k). These morphological changes are called "compaction", and blastulation occurs only after complete compaction of the embryo (l)* (image from Iwata *et al.*, 2014).

A clear limitation of our study is the fact that we only took into consideration morphokinetics as predictive factor of pregnancy. Our intention was merely to describe the difference automation can bring to annotations in all perspectives. Therefore, we do not

recommend the only use of morphokinetics as a predictive factor of pregnancy. In fact, we believe the combination of kinetics, morphology, deselection through atypical phenotypes, qualitative parameters, genetical testing, and any factor influencing the treatment, could help us refine the diagnosis and prediction of a particular embryo. This big data analysis combining the wide number of factors affecting outcome can only be performed through Artificial Intelligence (AI). The introduction of automated annotations into Time-Lapse Systems is the first step towards the use of AI in the IVF lab, and other similar systems have also been developed and validated by other groups, Eeva and KIDScore™ D5 model (Wong *et al.*, 2010; Kaser and Racowsky, 2014; Feyeux *et al.*, 2020; Kato *et al.*, 2021; Ueno *et al.*, 2022). Eeva continues to be validated by its latest publication (Aparicio-Ruiz *et al.*, 2018), although an RCT would be recommended to confirm this, especially in embryos developed up to the blastocyst stage, as most of Eeva's publications analyzed embryos on day 3 of development. About the KIDScore™ D5, a very recently published RCT showed it was not improving the prediction on ongoing pregnancy rate as compared to morphology alone (Ueno *et al.*, 2022).

Soon these technological developments would not only allow computers to be able to process millions of data but to obtain results at a very fast pace. Knowledge is power more than ever and the ability computers now have to massively process information has led to a society and, specially, an economy of which data has become the epicenter. All this objective and standardized morphokinetic data together with more input information regarding the culture conditions, patient details, etc., can lead to the development of precise outputs helping the reproductive specialists in individual decision making.

In conclusion, the results of the study support the use of automated systems for embryo morphokinetic annotations and embryo selection, as being more discriminative highlighting embryos resulting in a viable pregnancy. This non-invasive and objective tool standardizes the annotating process avoiding inter- and intra-observer variability. From a clinical perspective, our study highlights the benefits in the implementation of an automated annotation system into the routine clinical practice. Embryologist will gain the time spent performing the

annotations to conduct crucial tasks where they are indispensable, and to continue developing and improving our field. The establishment of automation would need a gradual transition controlled by lab professionals, as chaotic embryos with aberrant divisions and artifacts present in the well still makes the detection difficult for both the embryologist and the software. However, for those non-annotated events there is always the option to adding them manually, giving the embryologist the final word. In order to further progress with the applicability of the software, the rest of the morphokinetic parameters currently not detected should be further developed, to be able to assess other ESAs available. However, we can already conclude the benefit of the automated annotations in pregnancy prediction when using already published morphokinetic algorithms. The help of AI techniques in IVF offers a beneficial future perspective to further improve the standardization and objectivity of each evaluation performed in the laboratory between operators and clinics.

SPECIFIC OBJECTIVE III.

Non-invasive oxidative status analysis of the spent embryo culture medium in combination with Time-Lapse morphokinetics.

3.1 Purpose

The purpose of **Specific Objective III** was to develop a non-invasive embryo selection algorithm consisting of time-lapse Morphokinetics and the oxidative status of the spent embryo culture medium determined using the Thermochemiluminescence (TCL) Analyzer.

Following the ultimate goal evaluating embryo's viability through non-invasive methods has created the need for new techniques of analysis. Usually, these technologies are not suitable for routine clinical practice because of the cost of equipment, the time required to complete the process, the complexity of procedures, and the need for highly trained specialists, or because they add no substantial value to embryo selection relative to current methods. However, The TCL Analyzer has been previously validated in several publications assisting the performance of the IVF process though the analysis of oxidative stress (Wiener-Megnazi *et al.*, 2002, 2011; Lissak *et al.*, 2004; Wiener-megnazi *et al.*, 2004).

Thus, combining an oxidative TCL assay with a robust morphology and morphokinetic analysis using the TLS technique may form the basis of a more accurate embryo selection algorithm.

3.2 Analysis

3.2.1. Design

This prospective study included a total of 505 samples of spent embryo culture media analysed from 292 ICSI cycles between January 2016 and January 2017 in IVIRMA Valencia and Alicante. Only one cycle per patient was included in the study. The ICSI cycles were selected from women undergoing fertility treatment only using donor oocytes. The analysis was confined to single-embryo transfers resulting in a singleton pregnancy. In this study, none of the recipients shared the same oocyte donor.

3.2.2. Oxidative Status Analysis

The continuous culture media (CCM) samples were collected from the culture dishes on day 5. The oxidizability of the embryo medium samples was measured using the TCL Analyzer (Carmel Diagnostics). A 15 µL sample of CCM culture medium from each well was positioned in an analysis cuvette. After being tightly sealed and vacuum dried, the samples were heated to a constant temperature of 80 ± 0.5 °C, during which biological molecules, such as proteins and lipids, underwent induced oxidation generating unstable electronically excited species. These electronically excited species were further decomposed to stable carbonyl end products and light energy and counted as photons emitted per second (cps). Sequential photon counting was performed in a spectral range of 350–600 nm wavelength using 9913 Photon Counting Head (Hamamatsu Photonics) for 300 seconds. Thus, TCL indicated the rate of formation of unstable carbonyls in the sample, reflecting the total oxidants.

3.2.3 Algorithm Development

As previously published by our group, the following morphokinetic variables related to the duration of cell cycles were considered as predictive of embryo implantation: the duration of the second cell cycle defined as the time from division to a two-blastomere embryo until division to a three-blastomere embryo or t3–t2 (range: 9–12 hours); the duration of the

transition from a three-blastomere embryo to a four-blastomere embryo or t4–t3 (range: 0–0.75 hours); the time of three-blastomere embryo (34–40 hours); the time of the five-blastomere embryo (46.6–58.8), the time of the expanded blastocyst (tEB) ≤112.9 hours after ICSI; and the transition time from a five-blastomere embryo to an eight-blastomere embryo (t8–t5) of ≤5.67 hours after ICSI (Meseguer *et al.*, 2011; Motato *et al.*, 2016).

Logistic regression analysis was also applied to the TCL parameters described in the previous section to determine which was most predictive of implantation. We also included blastocyst morphology based on the ASEBIR categories and included donor oocyte age as a potential confounding factor. An algorithm was then developed by combining optimal morphokinetic time ranges, blastocyst morphology, and donor age together with TCL parameters. We then determined the ability of this algorithm to achieve more accurate embryo selection compared with morphokinetic criteria alone.

<u>3.2.4 Statistical Analysis</u>

Statistical analysis was performed using the Statistical Package for the Social Sciences 22 (SPSS Inc.). $P<0.05$ was considered statistically significant. Oxidative parameters were statistically analysed using a multifactorial analysis of variance ANOVA model. Embryos were classified into two groups based on day-5 outcomes: transferred and vitrified (n=426) or discarded (n=79). Oxidative status was compared between groups using chi-square tests for categorical data.

Logistic regression analysis was performed using the previous variables to determine how they are related to blastocyst implantation potential. The binary response parameter was the presence of a gestational sac (‘‘100” for implanted embryo or ‘‘0’’ for no gestational sac). Each of the binary variables was submitted to computer analysis by the forward step method (likelihood method), and those with $P > 0.05$ were not considered in the final model. After this statistical analysis, the variables in question were classified as those which statistically significantly predicted implantation potential.

The odds ratio (OR) of the effect of all binary variables that were associated with implantation potential were expressed in terms of 95% confidence interval (95% CI) and statistical significance. Receiver operating characteristic (ROC) curves were employed to test the predictive value of all the variables included in the model with respect to blastocyst stage and implantation. A ROC curve analysis provides the area under the curve (AUC) values between 0.5 and 1 and provides a measurement of the global classification ability of the model.

3.3 Results

3.3.1 Characteristics of Included Embryos

The demographic characteristics of the women and donors are shown in **Table 9**. Embryos cultured individually in the Embryoscope incubator (n=505) were evaluated on whether or not they developed to blastocyst stage on day 5. Based on their morphokinetic characteristics, 426 embryos were transferred or vitrified (84.4%), and 79 were discarded (15.6%). A total of 205 successful single-embryo transfers were performed, and out of these, 111 successfully implanted. There was only one case of monozygotic twins, and it was excluded from the analysis. Morphokinetic data were only available in 201 embryos transferred, where 107 successfully implanted. Successful implantation was determined by the presence of a gestational sac after 12 weeks of pregnancy.

Table 9. *Patient demographic characteristics.*

Characteristic	Value	95% CI
Fresh donors		
No. donation cycles	292	—
Age (y)	27.2	26.7–27.7
Recipients of fresh oocytes		
No. of patients	292	—
Age (y)	40.8	40.2–41.3
BMI (kg/m^2)	23.0	22.2–23.8
Days of endometrial preparation	14.7	14.1–15.3
MII inseminated by ICSI (mean/recipient)	10.8	10.3–11.3
Fresh sperm concentration	42.7	39.2–46.3
Fresh sperm motility	33.1	31.9–34.3
Fertilization rate (2PN)	77.0%	76.3–77.6
Blastocyst formation rate (by day 5)	64.0%	62.2–65.7
Blastocyst formation rate (by day 6)	71.1%	69.4–72.7

Note: Data are provided as number or percentage, as specified. BMI = body mass index; CI = confidence interval; ICSI = intracytoplasmic sperm injection; MII = metaphase 2; PN = pronuclei.

3.3.2 Oxidative Status Analysis

We examined the oxidative status of the culture medium on day 5 to assess a potential relationship between oxidative stress and implantation potential. The oxidative stress parameters H1sm, H2sm, H3sm, and the average of the three values (Hsm) were statistically significantly higher in culture media from embryos that were transferred and vitrified compared with those from discarded embryos ($P<0.05$; **Figure 17, A**). However, the Ratiosm parameter was not statistically significantly different between media from transferred and vitrified embryos versus discarded embryos. Transferred blastocysts that successfully implanted (n=111) also showed statistically significantly higher levels of H1sm, H2sm, H3sm, and average Hsm compared with those that did not implant (n=94) ($P<0.05$; **Figure 17, B**).

3.3.3 Combined Embryo Selection Algorithm

Logistic regression identified H2sm as the most predictive parameter of successful implantation outcomes. A threshold H2sm level of ≥92.96 cps was associated with a statistically significantly increased likelihood of successful implantation (OR 1.854; 95% CI, 1.023–3.357; P=0.042).

The H2sm parameter was introduced into the hierarchical classification tree model based on morphokinetic parameters to derive seven different categories of embryo potential: A, B, C, D, E, F, or discarded (**Figure 18**). As observed in **Table 10**, our TLS morphokinetic categories B and C were very similar in relation to implantation rates (60% and 57.1%); as a result, we combined both in one, as is seen in **Figure 18**. The six categories of the combined TCL & morphokinetics were created by combining three morphokinetic categories with one TCL value, thereby creating two new categories for each existing morphokinetic category. The algorithm provided a more discriminatory classification of implantation success compared with the ASEBIR morphological categories of embryo quality (**Table 10**).

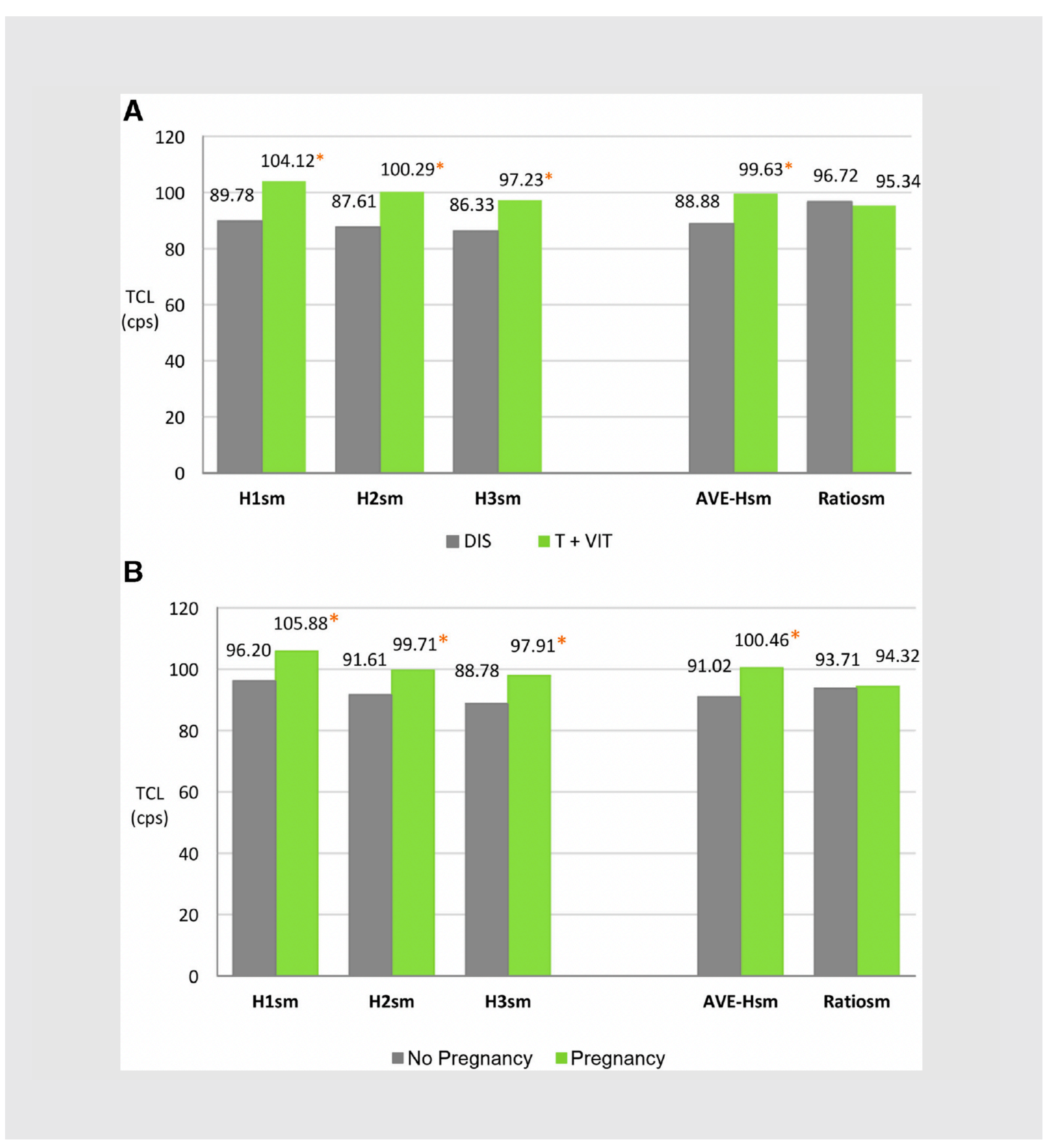


Figure 17. (A) *Thermochemiluminescence (TCL) parameters (H1sm, H2sm, H3sm, AVE-Hsm, Ratiosm) in relation to embryo outcome.* ***(B)*** *TCL parameters (H1sm, H2sm, H3sm, AVE-Hsm, Ratiosm) in relation to implantation.*

Note. DIS = Discarded; T = Transferred; V = Vitrified.

P<0.05 versus discarded embryos* *(A)*** *or versus embryos not resulting in pregnancy* ***(B)****.*

A more favourable combination of morphokinetic and TCL parameters was associated with a higher likelihood of successful implantation. The ROC curve for the predictive potential of the model showed an AUC value of 0.656 (95% CI, 0.579–0.733). The age of the oocyte donor was considered as a confounding factor.

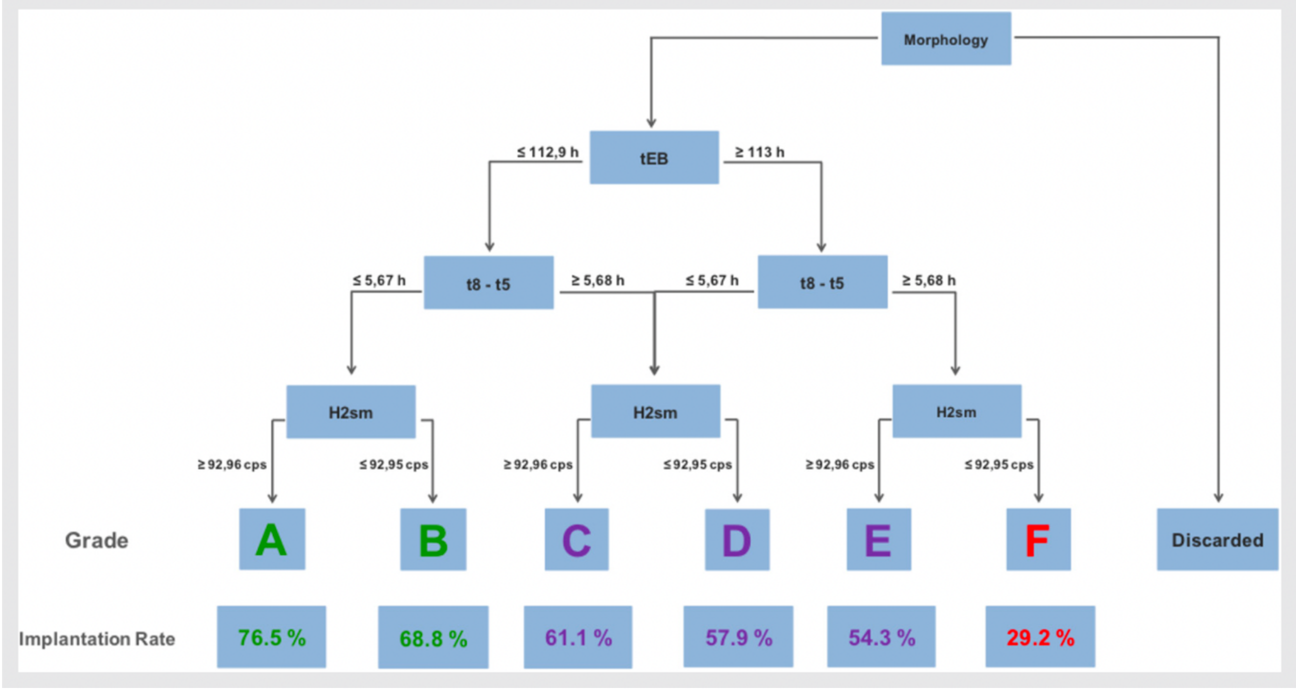


Figure 18*. Hierarchical classification of embryos based on morphokinetic screening parameters (tEB and t8–t5)* (Motato *et al.*, 2016) *and thermochemiluminescence (TCL) parameter H2sm. The classification generates seven categories of embryos with differential implantation rates.*

Table 10. *Implantation success rates using our combined TCL and morphokinetic algorithm compared with implantation rates using the ASEBIR morphologic classification.*

Category	Total no. of transferred blastocysts (n = 201)	Blastocyst implantations (n = 107)[a]	Implantation rate (%)	OR (95%CI)[b]	*P* value
ASEBIR category					
A	44	29	65.9	—	—
B	130	69	53.1	—	—
C	18	7	38.9	—	—
D	9	2	22.2	—	—
Morphokinetic category					
A	33	24	72.7	—	—
B	60	36	60.0	—	—
C	14	8	57.1	—	—
D	94	39	41.5	—	—
Category based on algorithm combining TCL and morphokinetics					
A	17	13	76.5	7.893 (2.19–28.44)	.002
B	16	11	68.8	5.343 (1.57–18.22)	.007
C	36	22	61.1	3.816 (1.53–9.52)	.004
D	38	22	57.9	3.339 (1.36–8.18)	.008
E	46	25	54.3	2.891 (1.24–6.77)	.014
F	48	14	29.2	—	—

Note: For ASEBIR (Asociación para el Estudio de la Biología de la Reproducción) classification, see Motato et al. (28). CI = confidence intervals; OR = odds ratio; TCL = thermochemiluminescence analyser.

[a] Successful implantation was determined by the presence of a gestational sac after 12 wk of pregnancy.

[b] For successful implantation for each category relative to class F embryos.

3.4 Discussion

Results of this study strongly suggest that the oxidative status of the spent embryo culture medium may correlate with its ability to develop a healthy-looking blastocyst and subsequently succeed at implantation. This confirms previous data showing the embryo's secretome is altered depending on the embryo's reproductive potential (Paszkowski and Clarke, 1996; Bromer and Seli, 2008; Lipari *et al.*, 2009). The techniques and algorithm we describe may represent an additional predictor to enhance current techniques or provide new insights into embryo quality markers.

In our study, statistically significantly higher TCL values from the spent embryo culture medium were associated with an increased likelihood of achieving a successful pregnancy. Findings with a similar direction were shown in the preliminary study of Wiener-Megnazi *et al.* (2011) with a previous generation TCL device in which the oxidative status of day 3 embryos, as defined by a different TCL parameter H1, correlated with their ability to implant. In that study, a positive predictive value of 70.6% for the occurrence of pregnancy was obtained when the maximal intracohort H1 amplitude was >210 cps and TCL ratio was ≥80% ($P<0.0001$). Unfortunately, although a similar trend is observed, pregnancy rates cannot be compared between our study and Wiener-Megnazi *et al.'s* (2011) because we used a different TCL parameter (H2) and we focused on combining the TCL data with morphokinetics in an algorithm.

Contradicting the usual understanding, the findings of our study suggest that oxidative stress is not a negative marker. Oxidative stress is a major contributor to infertility (Agarwal *et al.*, 2016), and even previous studies validating the TCL assay for the assessment of sperm quality (Lissak *et al.*, 2004) or myocardial infarction diagnosis (Shnizer *et al.*, 2003) concluded that less oxidative stress was a positive indicator, our results imply an opposite conclusion. A explanation could be that high-quality embryos display a more extensive oxidative metabolism, exerting an ''oxidative load'' on the surrounding medium (Wiener-Megnazi *et al.*, 2011). In fact, previous studies about oxygen consumption in embryos support this statement. Tejera *et al.* (2016) used a Clark O_2 sensor embedded in a time-lapse system to

measure the oxygen concentration of culture medium and estimate the embryo's respiration rate. Oxygen consumption was higher in the embryos that resulted in a pregnancy compared with the embryos that did not. This suggests that embryos with a higher implantation potential have a more active metabolism, performing more oxidative processes. In addition, a prospective study using Raman and near-infrared spectroscopic (NIRS) analysis of spent culture media concluded that the most predictive markers of implantation success were metabolites related to oxidative stress (–CH, –NH, and–OH groups). The metabolic profile of embryos on day 3 with successful pregnancy results presented a relative increase in the amounts of –OH with a decrease in –CH and –NH (Seli *et al.*, 2007).

By contrast, a subsequent patient-level meta-analysis of data from randomized studies found that adding NIRS to the usual morphological assessment of embryo quality did not result in an increase in live births compared with the standard morphological assessment alone (Vergouw *et al.*, 2014). There may be a number of reasons why the NIRS analysis of culture medium failed to provide additional information on embryo quality whereas our study does. First, the NIRS technology uses a broad band of the electromagnetic spectrum (780 to 2,500 nm), making it difficult to assign features of the complex spectra to specific molecules. In contrast, TCL uses a relatively narrow range of wavelengths (350 to 600 nm), making it easier to discern specific compounds. Second, the algorithm in our study combined TCL data with morphokinetic parameters, whereas the NIRS data were used in combination with standard morphological criteria.

We propose that measuring oxidative products in the culture medium may not be sufficient on its own to derive information on embryo quality. However, when combined algorithmically with morphokinetic criteria it can show a greater discriminatory power than morphological assessment alone identifying high-quality embryos. Our selection algorithm classifies those embryos achieving higher categories (A and B) when both morphokinetic parameters lie within the optimal ranges, as well as those that are in the lower categories (E and F) when the parameters are suboptimal. Categories C and D represent embryos with

optimal times in either one of the two parameters. This shows similar decisive power of tEB and t8–t5 as little difference was found in their implantation rates.

The oxidative process probably plays an important role in embryo–endometrium communication, with a certain threshold to be reached for an embryo to successfully implant. This is supported by the results of Lipari *et al.* (2009) where a correlation was found between the production of a pro-oxidant (nitric oxide) in the insemination medium and the embryo's potential to progress to the blastocyst stage in day 5. In addition, a specific range of TCL amplitude values in follicular fluid was positively correlated with pregnancy rate (Wiener-megnazi *et al.*, 2004). Therefore, optimal embryo maturation may require a certain level of oxidative stress, but excessive levels of oxidation may have a detrimental effect, causing fragmentation, degeneration, or developmental arrest. This is consistent with the ‘‘Goldilocks zone’’ hypothesis by (Leese *et al.*, 2016), in which viable embryos can moderate their metabolic activity between a minimum threshold required to undergo development, and a maximum level that does not exhaust available nutrient resources. Embryos with a metabolic rate outside of this range may not be viable because they either cannot generate sufficient energy for development or cannot maintain high rates of metabolism over a prolonged period.

The production of reactive oxygen species (ROS) is one of the most influential factors affecting the embryo's viability during its manipulation. Reactive oxygen species are by-products of cellular metabolism and need to be maintained at physiologic levels for the optimal function of signal transduction pathways (Sharma *et al.*, 1999). The oxidative status of the embryo reflects a delicate balance between ROS production and the embryo's ability to detoxify ROS (i.e., its antioxidant activity). Higher ROS levels may have a negative impact on cell development and function, altering the balance of crucial molecules such as lipids, proteins, or nucleic acids (Agarwal *et al.*, 2014). High oxidation potential could be counterbalanced by a high total antioxidant capacity (TAC) in high-quality embryos, while ROS generated by the embryo could be neutralizing exogenous antioxidants leading to a decline

in the TAC of the embryo medium (Paszkowski and Clarke, 1996). Hence, the embryos that achieve this delicate balance between pro- and antioxidant activities would be of higher quality than those that do not.

Conventional techniques for analysis of oxidative stress markers are not easily applicable in the routine IVF clinical practice. However, new technologies like TCL open the possibility for the routine measurement of these parameters. TCL is simple and has an intuitive use, hence there is no need for specialist training. The assay takes only 10 minutes to complete, thus predictive assessments of embryo quality can be made in the limited period before embryo transfer. The parameters of TCL also provide an objective assessment of oxidative stress, limiting the potential for interobserver variability. Last but most important, the analysis of the spent medium at the end of embryo culture ensures an undisturbed development and an optimal non-invasive technique. A very recent alternative technique, non-invasive PGT (niPGT), analyzing the genetic composition of the embryo by sampling its culture media, is gaining strength although still under comprehensive examination (Leaver and Wells, 2020; Rubio *et al.*, 2020; Hanson *et al.*, 2021).

It is important to remember that embryo selection is only one determinant of implantation success, as having an optimal endometrium is the other key element in the equation. A displaced window of implantation, the presence of uterine abnormalities, or any systemic disease could hamper endometrial receptivity, leading to a negative result (Garrido-Gómez *et al.*, 2013). Consequently, embryo outcome at day 5, rather than implantation outcome, is frequently used as an end point in studies of embryo selection techniques. It is noteworthy that the TCL parameters of H1sm, H2sm, H3sm, and average Hsm in our study were statistically significantly different in both analyses that examined day 5 embryo outcome or implantation, confirmed with gestational sac and fetal heartbeat, as the end point. Embryos selected as the best quality ones for transfer based on current techniques also showed statistically significantly elevated levels of TCL parameters, supporting the correlation between embryo quality and successful implantation. This encouraged us to combine TCL

oxidative parameters with TLS morphokinetic parameters in an algorithm for embryo selection based on variables that can be measured non-invasively. Understanding the characteristics of embryos with high reproductive potential may help identify those with a high likelihood of successful pregnancy in an attempt to reduce the number of embryos to be transferred, thus reducing the likelihood of multiple pregnancies.

Our study has several limitations that deserve mention. We used embryo outcome and ongoing pregnancy rate as measures of embryonic quality, whereas they should be complemented with livebirth rate to provide a fuller picture. Additional TAC assays should be performed to calculate a ROS-TAC score in addition to TCL (Sharma *et al.*, 1999; Mahfouz *et al.*, 2009). The ROC of our algorithm was 0.656, indicating that further refinement may be needed to increase its discriminatory power. The oxidative status results of this study cannot be considered physiologically factual. Much of the oxidative stress found in the culture medium is due to the pathophysiology of impaired embryo development *in vitro* (Paszkowski and Clarke, 1996). However, by conducting the study at a single centre, we could ensure that culture conditions were standardized across all samples, so the process would have affected all embryos to the same extent. In addition, extensive oocyte denudation was undertaken to ensure that the results were not affected by the presence of cumulus cells, which may be an additional source of ROS. Therefore, the study was designed to limit variability for the development of a valid predictive intracohort model. As this was a development study on a small number of embryos, further validation in an independent dataset would be required to confirm the results.

CONCLUSION

CONCLUSION

1. ACHIVEMENT OF OBJECTIVES & CLINICAL RELEVANCE

Time-lapse technology has helped to elucidate key events of embryo development. The relevance of this methodology has increased in the last years since it eases the embryologist workload and it helps us perform a more objective embryo development analysis, with the ultimate goal of selecting the best embryo for a successful pregnancy. The use of deep learning algorithms to analyze developmental events automatically is a step towards the implementation of artificial intelligence into embryo assessment, which is becoming a significant trend in the future.

Our study validated Geri Assess 2.0 as an automated annotation software for clinical use in the IVF laboratories. A high detection rate accordance was obtained between the embryologists' manual annotations and the software Geri Assess 2.0 automated annotations. Early events were detected very similarly in both groups, showing a high accuracy in the software's performance. Late events showed a higher difference, but the automated annotations found out to be more helpful in clinical outcome prediction, with the use of a previously published ESA. This may be due to the subjective nature of later events, where the software helps on the standardization of the annotations.

Hence, the results of the study support the use of automated systems for embryo morphokinetic annotations in the clinical practice. The non-invasive and objective nature of this tool standardizes the annotating process avoiding inter- and intra-observer variability, in addition to facilitating the routine clinical practice. The use of the software will ease the embryologists' workload as they gain the time spent performing the annotations. The establishment of automation would need a gradual transition controlled by lab professionals. The main limitations where development is still required are (1) the filtering time-ranges

should be readjusted, as they seem a bit too tight. Not all the morphokinetic parameters are detected by the software, meaning we cannot try some of the embryo selection algorithms available. (2) Chaotic embryos with aberrant divisions and artifacts present in the well will continue make it difficult for both annotating techniques, (3) but for those non-annotated events we can always add them manually, giving the embryologist the final word.

On another note, continuing with our main objective, we considered the use of a practical and validated technology to predict implantation potential of embryos in a rapid non-invasive way. Spent culture media metabolomics was selected as the perfect addition for a comprehensive embryo selection. Even if usually these technologies are not suitable for routine clinical practice because of the cost of equipment, the complexity of procedures, the need for highly trained specialists, or because they add no substantial value to embryo selection relative to current methods; these limitations were overcome in our TCL Analyzer validation.

Our preliminary results show that the oxidative stress TCL parameters are higher in transferred and vitrified embryos than in discarded embryos. Likewise, successfully implanted embryos obtained higher oxidative values compared with those that did not implant. These results suggest that high-quality embryos have a more extensive oxidative metabolism, exerting an oxidative load on its surrounding medium. This, together with its simple non-invasive processing could easily be compatibilized in the clinical daily practice.

Including the oxidative stress parameters into a decision-making algorithm along with the morphokinetic parameters, we developed a decision tree determining embryo quality for a successful implantation. The combination of these robust non-invasive parameters provides the foundation for a more objective embryo selection method before transfer in the routine clinical practice. This together with the help of AI techniques offers a good prospect on further improving the standardization and objectivity of each evaluation performed in the laboratory between operators and clinics.

2. FUTURE PERSPECTIVES

Although the main objective of validating non-invasive methods of automated morphokinetic annotations and oxidative stress analysis was achieved, a further validation would be useful to corroborate the results or missing analysis of the present thesis. First, all kinds of embryos should be included in the new study, not only the ones arriving to the blastocyst stage, but also the previously blocked ones. Even though these embryos will never be considered for transfer, an objective analysis of their morphokinetic parameters might elucidate information of their cellular cycle times and possibly correlate them to their developmental arrest. Another missing validation should be performed for the morphokinetic parameters tPNa and tHBi, as they had to be dropped from our analysis, and tEB, which was under development at the time of the study.

Even though the functionality of the filter is to exclude annotations that are biologically impossible for the given event, the filter applied was found to be too strict, thereby eliminating numerous datapoints in both groups. This raised doubts on the accuracy of the time ranges selected for the filter. A further examination should be performed on the filters to readjust them to clinical feasible ranges.

When analyzing the clinical benefit that automated annotations may provide in predicting outcome success, we found out the relevance time to morula has in the embryo development. This should incite future investigations about this crucial stage in the preimplantation development, entailing massive morphological, cellular, and molecular changes, to understand the way embryos acquire their reproductive competence during compaction. Moreover, the rest of morphokinetic parameters should be further analyzed to assess their possible added prediction potential as standardized values. Maybe not in a complete ESA, as some parameters may not be detected by the software but selecting their cut-off or optimal time ranges.

About the oxidative stress metabolomics, further data is required to validate the accuracy and reliability of the technique for assessing embryo quality. Analysis of spent embryo culture medium with the TCL assay should undergo a prospective blinded evaluation of its predictive potential compared with current selection methods. A large-scale prospective randomized study is therefore proposed to confirm the correlation of TCL parameters with implantation potential, ideally live birth rate, before this technique can be incorporated into wider clinical practice.

As any interdisciplinary field, a more holistic approach should be aimed where gynaecology, obstetrics, endocrinology, immunology, urology, andrology, embryology, and genetics, are considered and be more closely interconnected towards a better medicine for our ART patients. In fact, we believe the combination of morphology, kinetics, metabolomics, genetics, culture parameters, and treatment features could help us refine the diagnosis and prediction of a particular embryo (**Figure 19**). This big data analysis combining the wide number of factors affecting outcome can only be performed through Artificial Intelligence (AI). The implementation of automated annotations into Time-Lapse Systems is the first step towards the introduction of AI in the IVF lab, through the automatization of morphokinetics. However, this opens the door towards the development of deep learning algorithms for embryo selection. There is still a long way to go as this entails high risks if not implemented cautiously with comprehensive in-house validations before integration into clinical routine.

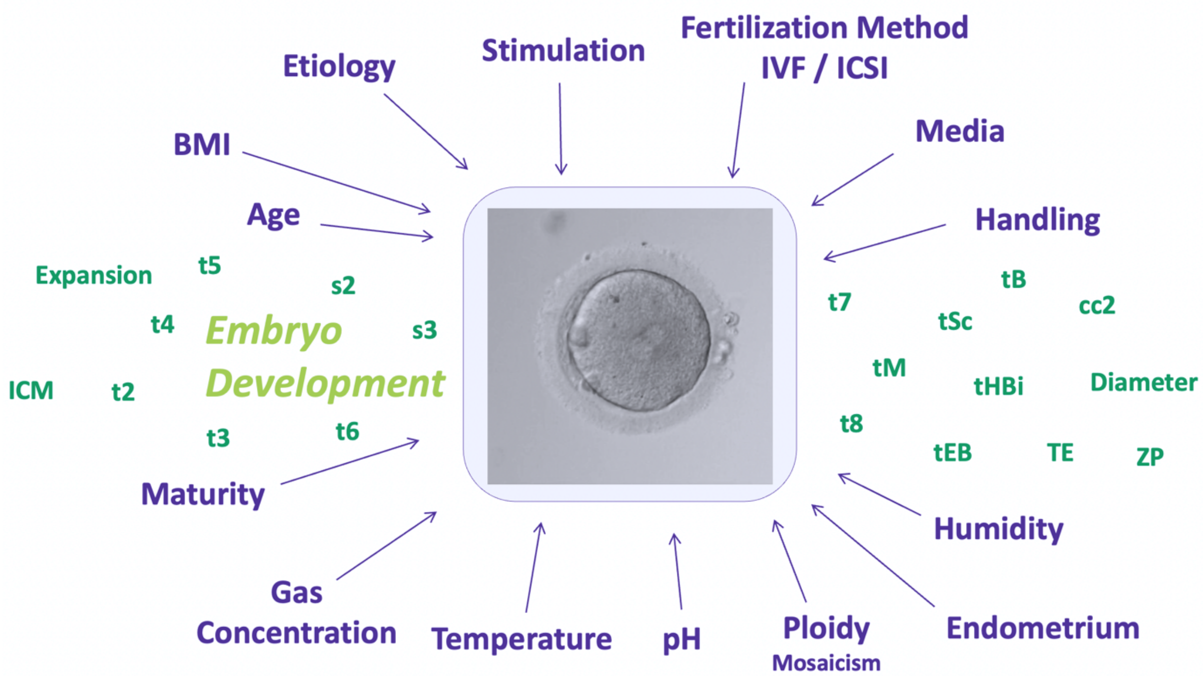


Figure 19. *Factors affecting in vitro embryo development.*

3. FINAL CONCLUSIONS

1. Geri Assess® 2.0 was successfully validated as an automated annotation software to guide embryologists to accurately detect key embryo developmental events, during culture in Time Lapse Systems, and standardise its evaluation within and across laboratories.

2. The automated annotations software presented a high detection rate and a comparable accuracy to the embryologists' manual annotations in all the developmental events. Although earlier events obtained more similar annotations between groups, later events were more helpful in clinical outcome prediction.

3. Filtering time ranges of the software require further examination for a readjustment, as excluded data points represented a 13.5% as of the Geri Assess® 2.0 annotated events and 8.1% of the manually annotated events.

4. Chaotic embryos with aberrant divisions and or artifacts in the well hinder the annotations process in both groups. These embryos rarely arrive to later stages of development for selection. Wrong or non-annotated events can always be corrected manually, giving the embryologists the final word.

5. High accordance was found in embryo grading when using both annotation sets in an ESAs developed with early embryo development events.

6. When analyzing blastocyst stage-developed ESAs, time to morula, with a cut-off value of 80 hours, demonstrated to be a better outcome predictive parameter when using the automated annotations than the manual annotations

7. The rest of the morphokinetic parameters, currently not detected by the software, should be further developed, to be able to assess other ESAs available.

8. The thermochemiluminescence (TCL) assay was successfully validated as a non-invasive tool to perform the analysis of the oxidative stress of the spent culture media of the embryo.

9. Statistically significantly higher TCL values obtained from the spent embryo culture medium were associated with an increased likelihood of achieving a successful pregnancy.

10. The combination of the TCL oxidative parameters with TLS morphokinetic criteria presented a greater discriminatory power than morphological assessment in the identification of high-quality embryos, providing the foundation for a more objective and non-invasive embryo selection method to reduce multiple embryo transfers.

3. CONCLUSIONES FINALES

1. Geri Assess® 2.0 fue validado con éxito como un software de anotaciones automatizadas para guiar a los embriólogos en la detección de eventos clave en el desarrollo embrionario, durante su cultivo en sistemas de lapso de tiempo, y estandarizar su evaluación dentro y entre laboratorios.

2. El software de anotaciones automatizadas presentó una alta tasa de detección y una precisión comparable a las anotaciones manuales de los embriólogos en todos los eventos de desarrollo. Aunque los eventos tempranos obtuvieron anotaciones más similares entre los grupos, los eventos tardíos fueron más útiles en la predicción del desenlace clínico.

3. Los rangos de tiempo preseleccionados para el filtrado de las anotaciones requieren un análisis detallado para su reajuste, ya que los datos excluidos representaron un 13,5 % de los eventos anotados por Geri Assess® 2.0 y un 8,1 % de los eventos anotados manualmente.

4. Los embriones caóticos con divisiones aberrantes y/o artefactos presentes en el pocillo dificultan el proceso de anotación en ambos grupos. Estos embriones rara vez llegan a etapas posteriores de desarrollo para su selección. Para los eventos incorrectos o no anotados siempre se pueden corregir manualmente, dando a los embriólogos la decisión final.

5. Se halló una alta concordancia en la clasificación de embriones cuando se usaron ambos conjuntos de anotaciones en un Algoritmo de Selección Embrionaria desarrollado con eventos de desarrollo embrionario temprano.

6. Al utilizar Algoritmos de Selección Embrionaria desarrollados en el estadio de blastocisto, el tiempo de desarrollo hasta el estadio de mórula, con un punto de corte de 80 horas, demostró ser un mejor parámetro predictivo de embarazo con las anotaciones automáticas que con las anotaciones manuales.

7. Es necesario el desarrollo del resto de los parámetros morfocinéticos, actualmente no detectados por el software, para la futura evaluación de otros Algoritmos de Selección Embrionaria disponibles.

8. La Termoquimioluminiscencia (TCL) se validó con éxito como una herramienta no invasiva para el análisis del estrés oxidativo de los medios de cultivo empleados por el embrión.

9. Los valores de TCL significativamente más altos se asociaron con una mayor probabilidad de lograr un embarazo con éxito.

10. La combinación de los parámetros oxidativos de TCL con los criterios morfocinéticos de los sistemas de lapso de tiempo presentaron un mayor poder discriminatorio que la evaluación morfológica en la identificación de embriones de alta calidad, lo cual estableció las bases para un método de selección de embriones más objetivo y no invasivo para reducir la transferencia múltiple de embriones.

SUPPLEMENTAL MATERIAL

SUPPLEMENTAL MATERIAL

1. SUPPLEMENTAL FIGURES

Supplemental Figure 1.

Inner cell mass (ICM) quality

	ICM size (μm²)	Cohesion
A	3800–1900	Compact
B	3800–1900	Noncompact
C	1900	Indifferent
D	Degenerating signs	
Discarded	Degenerated or nonexistent	

Trophectoderm quality

	Description
A	Many cells, homogeneous, forming a cohesive layer
B	Homogeneous, but few cells
C	Very few cells
D	Degenerating signs
Discarded	Degenerated

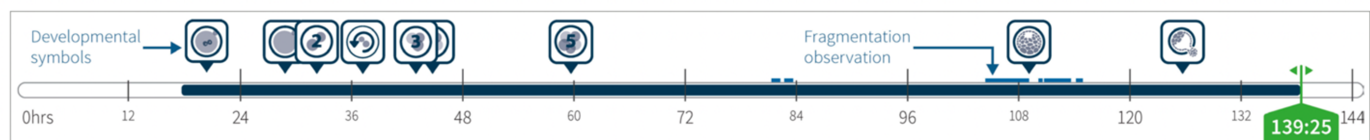
ASEBIR criteria

Expansion grade	ICM quality grade	Trophoectoderm quality grade	ASEBIR grade
BEi BE BH BHi	A	A	A
		B	B
		C	C
		D	D
	B	A	A
		B	B
		C	C
		D	D
	C	A	A
		B	B
		C	C
		D	D
	D	A,B,C or D	D
Early or cavitating blastocyst (thick zp)			C
Non-compact morula			D
Degenerated			Discarded

BE, blastocyst expansion; BEi, initiating blastocyst expansion; BH, hatching blastocyst; BHi, initiating blastocyst hatching.

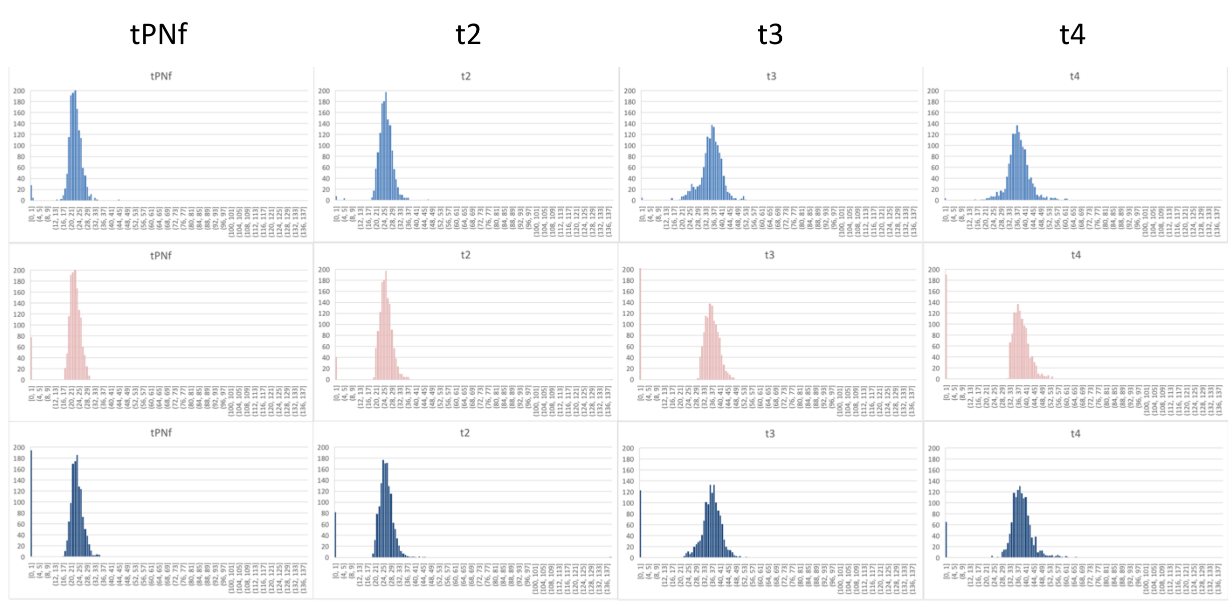
ASEBIR criteria (2015) for morphological assessment of blastocysts on day 5, between 112 and 120 hours.

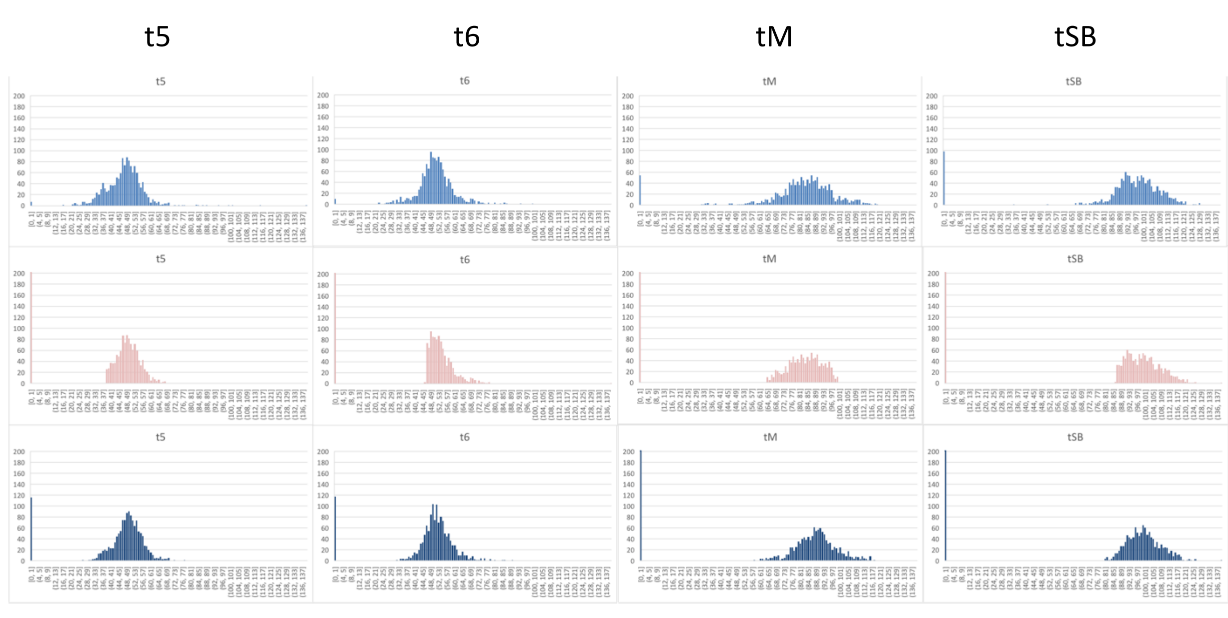
Supplemental Figure 2.



Geri Assess 2.0 Timeline Bar with annotated developmental events and observations per embryo. For fragmentation observations, as it can occur numerous times or during a continuous period of time, a blue bar is placed on the timeline to reflect the event's presence.

Supplemental Figure 3.

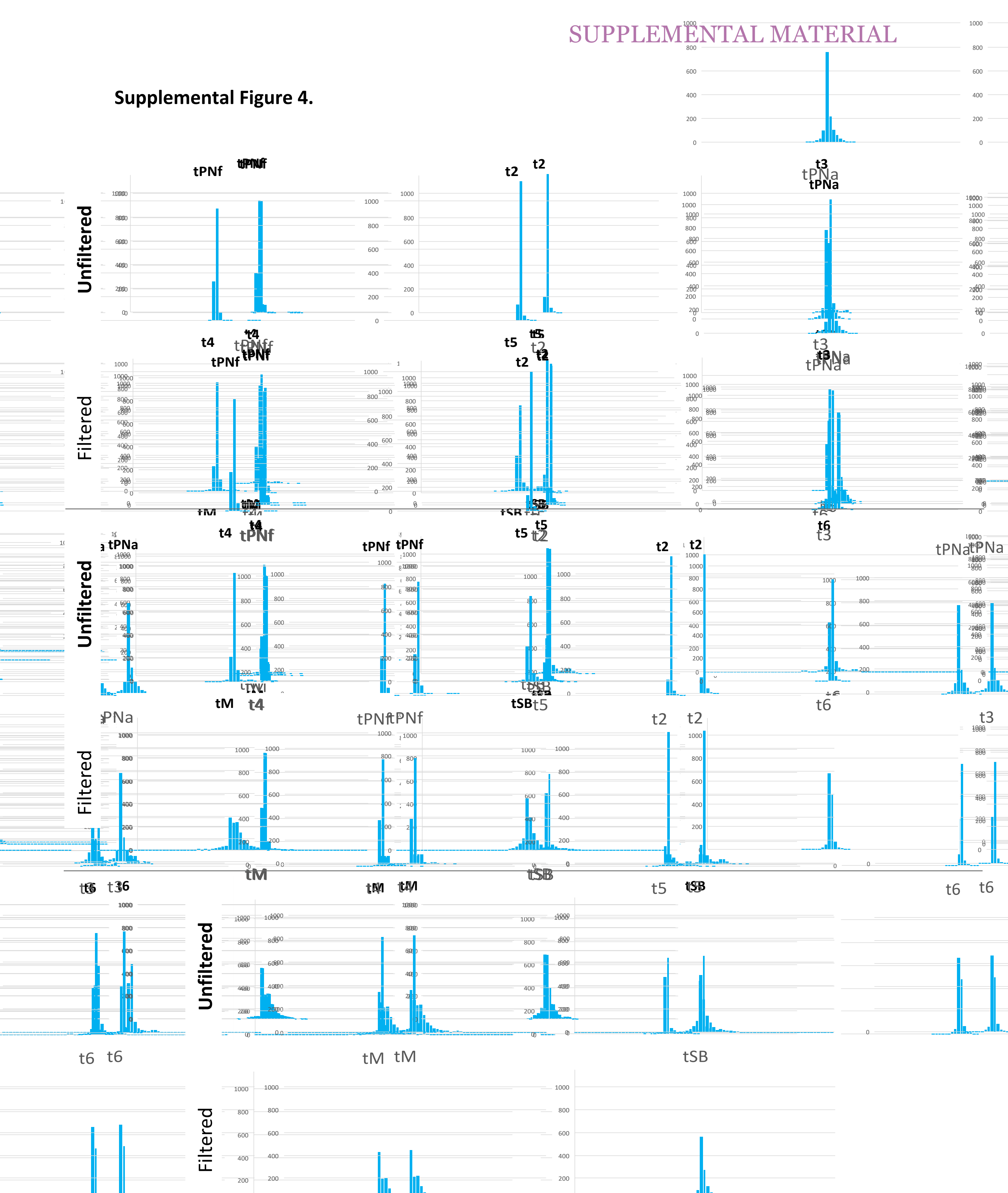




Geri Assess 2.0 Unfiltered Geri Assess 2.0 Filtered Embryologist Team

Data distribution of the annotations performed, for all embryo developmental events analyzed, by the IVIRMA embryologist team and the Geri Assess 2.0 before and after applying the filtering tool for the out-of-range annotations.

Supplemental Figure 4.



Standard deviation between the manual and automated annotations with the filtered and unfiltered dataset.

2. SUPPLEMENTAL TABLES

Supplemental Table 1.

ASEBIR criteria (2015) for morphological assessment of embryos on days 2-4.

Category	Definition
A	The two pronuclei (2PN) embryo consists of 4 cells at Day 2 and 7-8 cells at Day 3. Blastomeres are a consistent size at the 2-, 4- and 8-cell stage. No multinucleation is observed at any time and fragmentation is less than 10%.
B	The 2PN embryo consists of 2 or 5 cells at Day 2, and 7 or more cells at Day 3. Blastomeres are a consistent size at the 2-, 4- and 8-cell stage. No multinucleation is observed at any time and fragmentation is less than 20%.
C	The 2PN embryo consists of 3, 6 or more cells at Day 2, and 6–8 cells (or morula) at Day 3. The embryo may have asymmetric blastomeres and multinucleation can be observed in no more than one blastomere at each stage. The degree of fragmentation is less than 20%.
D	The 2PN embryo consists of 1–2 cells at 27 hours, 2–6 cells at Day 2 and, four to >8 cells or morula at Day 3. The embryo may have asymmetric blastomeres and be multinucleated. The degree of fragmentation is 50%.

Supplemental Table 2.

Time parameters assessed during TLS (Motato *et al.*, 2016).

Parameter	Definition
Embryo maturation	
t2, t3, t4, t5, t6, t7, t8, and t9	Time from ISCI to cleavage from two to nine cells
tM	Timing of morula – time from ISCI until the stage at which the individual blastomere membranes are no longer visible and full embryo compaction has been achieved
tB	Timing of blastocyst - the time from insemination to the formation of a "full blastocyst" when the blastocoele cavity is filled the embryo and the ICM and TE tissues are distinguishable from each other (30)
tEB	Timing of expanded blastocyst – time from ISCI until expanded blastocyst formation
tHB	Timing of hatched blastocyst – time from ISCI until blastocyst hatching
Cell cycle duration	
t3-t2	Duration of second cell cycle was defined as the time from division to a 2-blastomere embryo until division to a 3-blastomere embryo
t5-t3	Duration of the third cell cycle
t4-t3	Time for transition from a 3-blastomere embryo to a 4-blastomere embryo
t8-t5	Time for transition of a 5-blastomere embryo to an 8-blastomere embryo
t5-t2	The interval between two and five cells, which combines the concepts of cell cycle and synchrony

3. SCIENTIFIC CONTRIBUTIONS

3.1 Scientific Publications

1. Time-Lapse Imaging: The State of the Art.

 Biology of Reproduction (2019) – ioz035.

 doi: 10.1093/biolre/ioz035.

 Del Gallego R, Remohí J and Meseguer M.

2. Novel noninvasive embryo selection algorithm combining time-lapse morphokinetics and oxidative status of the spent embryo culture medium.

 Fertility & Sterility (2019) – 111(5):918-927.e3.

 doi: 10.1016/j.fertnstert.2019.01.022.

 Alegre L, Del Gallego R, Arrones S, Hernández P, Muñoz M, Meseguer M.

3. Time of morulation and trophectoderm quality are associated with live birth after euploid blastocyst transfer: a multicenter study

 Fertility & Sterility (2019)

 doi: 10.1016/j.fertnstert.2019.07.1322.

 Rienzi L, Cimadomo D, Delgado A, Minasi MG, Fabozzi G, Del Gallego R, Stoppa M, Bellver J, Giancani A, Esbert M, Capalbo A, Remohì J, Greco E, Ubaldi FM, Meseguer M.

4. Assessment of embryo implantation potential with a cloud-based automatic software.

Reproductive BioMedicine Online (2020) Sep.

 doi: 10.1016/j.rbmo.2020.09.032.

 Alegre L, Del Gallego R, Bori L, Loewke K, Maddah M, Aparicio-Ruiz B, Palma-Govea A, Marcos J, Meseguer M.

3.2 National or International Conferences

1. Time-lapse technology combined with a novel automated analysis method for embryo selection; clinical validation. (Poster; P-347)

Authors: L Alegre; E Palma; J J Marcos; C Albert; R Del Gallego; A Pellicer; M Meseguer.

Conference: 74th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: San Antonio (USA)

Date: 28/10 – 1/11/2017

2. Nuevo biomarcador no invasivo para la selección embrionaria basado en el estado oxidativo del medio de cultivo. (Poster; P-008)

Authors: S Arrones; L Alegre; R Del Gallego; M J de los Santos; J Remohí; M Meseguer.

Conference: IX Congreso de la Asociación para el estudio de la Biología de la Reproducción (ASEBIR)

City of event: Madrid (Spain)

Date: 15-17/11/2017

3. Oxidative status of the embryo's spent culture media as a new noninvasive tool for embryo selection. (Poster)

Authors: V García-Láez; L Alegre; R Del Gallego; T Cnaani; S Shnizer; M Meseguer.

Conference: ALPHA 12th Biennial Conference

City of event: Reykjavik (Iceland)

Date: 17-20/05/2018

4. Clinical validation of a non-invasive embryo selection algorithm combining time-lapse morphokinetics and the oxidative status of spent embryo culture media. (Poster; P-182)

Authors: R Del Gallego; L Alegre; T Cnaani; S Shnizer; M Meseguer.

Conference: 68th Annual Meeting of the European Society of Human Reproduction and Embryology

City of event: Barcelona (Spain)

Date: 01-04/07/2018

5. Humid vs. Dry embryo culture conditions on embryo development: a continuous embryo monitoring assessment. (Poster)

Authors: R. Del Gallego, C. Albert, J. Marcos, Z. Larreategui, L. Alegre, M. Meseguer

Conference: 74th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Denver (USA)

Date: 6/10 – 10/10/2018

6. Embryologist team vs. Automated annotation software outcomes. (Poster)

Authors: L. Alegre, R. Del Gallego, T. Peura, L. Bori, B. Aparicio Ruiz, A. Adam, A. Coello, D. Castello, M. Meseguer

Conference: 74th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Denver (USA)

Date: 6/10 – 10/10/2018

7. Time of morulation and trophectoderm quality are associated with live birth after euploid blastocyst transfer: a multicenter study. (Oral Communication; O-006)

Authors: L.F. Rienzi, D. Cimadomo, A. Delgado, M.G. Minasi, G. Fabozzi, R. Del Gallego, M. Stoppa, J. Bellver, A. Giancani, M. Esbert, A. Capalbo, J. Remohì, E.Greco, F.M.Ubaldi, M.Meseguer.

Conference: 69th Annual Meeting of the European Society of Human Reproduction and Embryology

City of event: Vienna (Austria)

Date: 23 - 26/06/2019

8. Embryo development detection by automated software vs. embryologist team. (Oral Communication; O-171)

Authors: R. Del Gallego, L. Alegre, L. Bori, T. Peura, S. Azaña, M. Meseguer.

Conference: 69th Annual Meeting of the European Society of Human Reproduction and Embryology

City of event: Vienna (Austria)

Date: 23 - 26/06/2019

9. Assessment of embryo implantation potential with a cloud-based automatic software. (Poster; P-164)

Authors: L. Alegre, R. Del Gallego, L. Bori, N. Basile, K. Loewke, M. Mahnaz, B. Aparicio-Ruiz, A. Palma-Govea, J. Marcos, M. Meseguer

Conference: 69th Annual Meeting of the European Society of Human Reproduction and Embryology

City of event: Vienna (Austria)

Date: 23 - 26/06/2019

10. High levels of Follicular fluid oxidative stress are present in younger patients and fertile donors. (Poster; P-232)

Authors: D. Castello, R. Del Gallego, L. Bori, I. Hervas, L. Alegre, T.C. Cnaani, S.S. Shnizer, M.Meseguer

Conference: 69th Annual Meeting of the European Society of Human Reproduction and Embryology

City of event: Vienna (Austria)

Date: 23 - 26/06/2019

11. High culture media oxidative profile as a biomarker of good quality embryos: a non-invasive tool to select the embryo to transfer. (Poster; P-246)

Authors: M.D.M. Nohales Corcoles, R. Del Gallego, L. Bori, L. Alegre, S. Shnizer, M.Meseguer

Conference: 69th Annual Meeting of the European Society of Human Reproduction and Embryology **City of event**: Vienna (Austria)

Date: 23 - 26/06/2019

12. Would an automated system detecting embryo developmental events select the same embryo as an embryologist using a morphokinetic algorithm? (Oral Communication, O-9)

Authors: Raquel Del Gallego, Lorena Bori, Lucia Alegre, Teija Peura, Manuel Ugidos, Marcos Meseguer.

Conference: 75th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Philadelphia (USA)

Date: 12/10 – 16/10/2019

13. The effect of high humidity on embryo culture media oxidation. (Oral Communication, O-109)

Authors: Carmela Albert, Raquel Del Gallego, Lucia Alegre, Zaloa ZL. Larraeategui, Julian Marcos, Belen Aparicio-Ruiz, Marcos Meseguer.

Conference: 75th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Philadelphia (USA)

Date: 12/10 – 16/10/2019

14. Is there any room to improve embryo selection? Artificial intelligence technology applied for live birth prediction on blastocysts. (Oral Communication, O-184)

Authors: Marcos Meseguer, Cristina Hickman, Lorena Bori, Lucia Alegre, Marco Toschi, Raquel Del Gallego, Jose Celso Rocha.

Conference: 75th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Philadelphia (USA)

Date: 12/10 – 16/10/2019

15. Group embryo culture strategies affect the oxidative status of the spent culture media and embryo development results. (Poster, P-39)

Authors: Lorena Bori, Raquel Del Gallego, Lucia Alegre, Silvia Azaña, Thamara Viloria, Marcos Meseguer.

Conference: 75th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Philadelphia (USA)

Date: 12/10 – 16/10/2019

16. A massive embryo morphokinetics comparison system is able to select embryos with high implantation potential enhancing single embryo transfer policy. (Poster, P-415)

Authors: Lucia Alegre, Raquel Del Gallego, Lorena Bori Arnal, Manuel Muñoz, Antonio Pellicer, Marcos Meseguer.

Conference: 75th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Philadelphia (USA)

Date: 12/10 – 16/10/2019

17. Continuous monitoring of the embryo development: a leap towards automated systems. (Video Session, V-6)

Authors: Lorena Bori, Raquel Del Gallego, Lucia Alegre, Antonio Pellicer, Marcos Meseguer

Conference: 75th Annual Meeting American Society for Reproductive Medicine (ASRM)

Publication: "Supplement to Fertility and Sterility".

City of event: Philadelphia (USA)

Date: 12/10 – 16/10/2019

18. Aplicación de la inteligencia artificial para la selección embrionaria combinando el análisis proteico del medio del cultivo en contacto con el blastocisto, la morfocinética y la morfología en d5 de desarrollo (The application of Artificial Intelligence for embryo selection combining a proteomic analysis of the spent culture media, morphokinetics and day 5 blastocyst morphology.) (Oral Communication, 008). Merck-ASEBIR Innovation Award.

Authors: M. Meseguer Escriva, L. Bori, M. Toschi, Raquel Del Gallego, L. Alegre, C. Hickman, C. Rocha

Conference: X Congreso ASEBIR Cáceres 2019

Publication: ASEBIR.

City of event: Caceres (Spain)

Date: 23/10 – 25/10/2019

19. Efecto del cultivo en ambiente húmedo en el desarrollo embrionario y el perfil oxidativo del medio de cultivo (The effect of the humid environment on the embryo development and the oxidative profile of the culture media) (Oral Communication, 002).

Authors: C. Albert Rodríguez, Raquel Del Gallego, L. Alegre Ferri, Z. Larreategui Laiseca, J. Marcos Alises, B. Aparicio Ruiz, P. Gamíz Izquierdo, JM. De los Santos Molina, M. Meseguer Escrivá

Conference: X Congreso ASEBIR Cáceres 2019

Publication: ASEBIR.

City of event: Caceres (Spain)

Date: 23/10 – 25/10/2019

20. ¿Es un sistema automatizado capaz de detectar los eventos del desarrollo embrionario como un embriólogo? (Is an automated system able to detect embryo developmental events as an embryologist?) (Poster, 031).

Authors: Raquel Del Gallego, L Bori Arnal, L Alegre Ferri, S Azaña Gutiérrez, I Hervás Herrero, M Meseguer Escrivá

Conference: X Congreso ASEBIR Cáceres 2019

Publication: ASEBIR.

City of event: Caceres (Spain)

Date: 23/10 – 25/10/2019

21. Software automático de análisis morfocinético es capaz de detectar embriones con mayor potencial de implantación. Establecimiento de una política de transferencia de un único embrión. (Poster, 063).

Authors: T Viloria Samochín, L Alegre Ferri, Raquel Del Gallego, L Bori Arnal, A Pellicer Martínez, M Meseguer Escrivá

Conference: X Congreso ASEBIR Cáceres 2019

Publication: ASEBIR.

City of event: Caceres (Spain)

Date: 23/10 – 25/10/2019

22. Efecto del cultivo en grupo de embriones humanos sobre el perfil oxidativo del medio de cultivo (Embryo group culture effect on the oxidative profile of the spent culture media) (Poster, 095).

Authors: L Bori Arnal, Raquel Del Gallego, L Alegre Ferri, I Hervás Herrero, S Azaña Gutiérrez, M Meseguer Escrivá

Conference: X Congreso ASEBIR Cáceres 2019

Publication: ASEBIR.

City of event: Caceres (Spain)

Date: 23/10 – 25/10/2019

3.3 Prices, Mentions and Distinctions

Innovation Award MERCK-ASEBIR.

X Congreso ASEBIR – Cáceres, 10/2019.

Aplicación de la inteligencia artificial para la selección embrionaria combinando el análisis proteico del medio del cultivo en contacto con el blastocisto, la morfocinética y la morfología en d5 de desarrollo (The application of Artificial Intelligence for embryo selection combining a proteomic analysis of the spent culture media, morphokinetics and day 5 blastocyst morphology.) (Oral Communication, 008).

Authors: M. Meseguer Escriva, L. Bori, M. Toschi, Raquel Del Gallego, L. Alegre, C. Hickman, C. Rocha

Conference: X Congreso ASEBIR Cáceres 2019

Publication: ASEBIR.

City of event: Caceres (Spain)

Date: 23/10 – 25/10/2019

REFERENCES

REFERENCES

Adamson GD, Abusief ME, Palao L, Witmer J, Palao LM, Gvakharia M. Improved implantation rates of day 3 embryo transfers with the use of an automated time-lapse–enabled test to aid in embryo selection. *Fertil Steril* [Internet] 2016;**105**:369-375.e6. Elsevier Inc.

Adolfsson E, Andershed AN. Morphology vs morphokinetics: A retrospective comparison of interobserver and intra-observer agreement between embryologists on blastocysts with known implantation outcome. *J Bras Reprod Assist* 2018;**22**:228–237.

Agarwal A, Durairajanayagam D, Plessis SS du. Utility of antioxidants during assisted reproductive techniques: an evidence based review. *Reprod Biol Endocrinol* 2014;**12**:112.

Agarwal A, Sharma R, Roychoudhury S, Plessis S Du, Sabanegh E. MiOXSYS : a novel method of measuring oxidation reduction potential in semen and seminal plasma. *Fertil Steril* [Internet] 2016;**106**:566-573.e10. Elsevier Inc.

Aguilar J, Rubio I, Muñoz E, Pellicer A, Meseguer M. Study of nucleation status in the second cell cycle of human embryo and its impact on implantation rate. *Fertil Steril* 2016;**106**:291-299.e2.

Alegre L, Gallego R Del, Bori L, Loewke K, Maddah M, Aparicio-Ruiz B, Palma-Govea AP, Marcos J, Meseguer M. Assessment of embryo implantation potential with a cloud-based automatic software. *Reprod Biomed Online* [Internet] 2021;**42**:66–74. Elsevier Ltd.

Alegre L, Palma A, Marcos J, Albert C, Gallego R Del, Pellicer A, Meseguer M. Time-lapse technology combined with a novel automated analysis method for embryo selection; clinical validation. *Fertil Steril* 2017;**108**:e239.

ALPHA Scientists in Reproduction medicine and ESHRE SIG of Embryology. The Istanbul consensus workshop on embryo assessment: Proceedings of an expert meeting. *Hum Reprod* 2011;**26**:1270–1283.

Aparicio-Ruiz B, Basile N, Pérez Albalá S, Bronet F, Remohí J, Meseguer M. Automatic time-lapse instrument is superior to single-point morphology observation for selecting viable embryos: retrospective study in oocyte donation. *Fertil Steril* 2016;**106**:1379-1385.e10.

Aparicio-Ruiz B, Romany L, Meseguer M. Selection of preimplantation embryos using time-lapse microscopy in in vitro fertilization: State of the technology and future directions. *Birth Defects Res* [Internet] 2018;**110**:648–653.

Armstrong S, Arroll N, Lm C, Jordan V, Farquhar C. Time-lapse systems for embryo incubation and assessment in assisted reproduction. *Cochrane Database Syst Rev* 2015;CD011320.

ASEBIR. Asociacion para el Estudio de la Biologia de la Reproduccion (ASEBIR). Criterios ASEBIR de Valoracion Morfologica de Ovocitos, Embriones Tempranos y Blastocistos Humanos. Cuadernos de embriologia clinica. 2015;**Madrid**, **Sp**:

Athayde Wirka K, Chen AA, Conaghan J, Ivani K, Gvakharia M, Behr B, Suraj V, Tan L, Shen S. Atypical embryo phenotypes identified by time-lapse microscopy: High prevalence and association with embryo development. *Fertil Steril* [Internet] 2014;**101**:1637-1648.e5. Elsevier Inc.

Azzarello A, Hoest T, Hay-Schmidt A, Mikkelsen AL. Live birth potential of good morphology and vitrified blastocysts presenting abnormal cell divisions. *Reprod Biol* [Internet] 2017;**17**:144–150. Society for Biology of Reproduction & the Institute of Animal Reproduction and Food Research of Polish Academy of Sciences in Olsztyn.

Azzarello A, Hoest T, Mikkelsen AL. The impact of pronuclei morphology and dynamicity on live birth outcome after time-lapse culture. *Hum Reprod* 2012;**27**:2649–2657.

Barrie A, Homburg R, McDowell G, Brown J, Kingsland C, Troup S. Preliminary investigation of the prevalence and implantation potential of abnormal embryonic phenotypes assessed using time-lapse imaging. *Reprod Biomed Online* [Internet] 2017a;**34**:455–462. Elsevier Ltd.

Barrie A, Homburg R, McDowell G, Brown J, Kingsland C, Troup S. Examining the efficacy of six published time-lapse imaging embryo selection algorithms to predict implantation to demonstrate the need for the development of specific, in-house morphokinetic selection algorithms. *Fertil Steril* [Internet] 2017b;**107**:613–621. Elsevier Inc.

Basile N, Nogales MDC, Bronet F, Florensa M, Riqueiros M, Rodrigo L, García-Velasco J, Meseguer M. Increasing the probability of selecting chromosomally normal embryos by time-lapse morphokinetics analysis. *Fertil Steril* [Internet] 2014;**101**:699-704.e1. Elsevier Inc.

Basile N, Vime P, Florensa M, Aparicio Ruiz B, García Velasco JA, Remohí J, Meseguer M. The use of morphokinetics as a predictor of implantation: A multicentric study to define and validate an algorithmfor embryo selection. *Hum Reprod* 2015;**30**:276–283.

Baxter Bendus AE, Mayer JF, Shipley SK, Catherino WH. Interobserver and intraobserver variation in day 3 embryo grading. *Fertil Steril* 2006;**86**:1608–1615.

Bhide P, Maheshwari A, Cutting R, Seenan S, Patel A, Khan K, Homburg R. Time lapse imaging: is it time to incorporate this technology into routine clinical practice? *Hum Fertil* 2017;**20**:74–79.

Boivin J, Bunting L, Collins J. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum Reprod* 2007;**22**:1506–1512.

Bori L, Dominguez F, Fernandez EI, Gallego R Del, Alegre L, Hickman C, Quiñonero A, Nogueira MFG, Rocha JC, Meseguer M. An artificial intelligence model based on the proteomic profile of euploid embryos and blastocyst morphology: a preliminary study. *Reprod Biomed Online* [Internet] 2021;**42**:340–350. Elsevier Ltd.

Botros L, Sakkas D, Seli E. Metabolomics and its application for non-invasive embryo assessment in IVF. *Mol Hum Reprod* 2008;**14**:679–690.

Braude P, Bolton V, Moore S. Human gene expression first occurs between the four- and eight-cell stages of preimplantation development. *Nature* 1988;**332**:459–461.

Bromer JG, Seli E. Assessment of embryo viability in assisted reproductive technology: shortcomings of current approaches and the emerging role of metabolomics. *Curr Opin Obstet Gynecol* [Internet] 2008;**20**:234–241.

Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Hickman CFL. Modelling a risk classification of aneuploidy in human embryos using non-invasive morphokinetics. *Reprod Biomed Online* [Internet] 2013a;**26**:477–485. Reproductive Healthcare Ltd.

Campbell A, Fishel S, Bowman N, Duffy S, Sedler M, Thornton S. Retrospective analysis of outcomes after IVF using an aneuploidy risk model derived from time-lapse imaging without PGS. *Reprod Biomed Online* [Internet] 2013b;**27**:140–146. Reproductive Healthcare Ltd.

Carmen Nogales M Del, Bronet F, Basile N, Martínez EM, Liñán A, Rodrigo L, Meseguer M. Type of chromosome abnormality affects embryo morphology dynamics. *Fertil Steril* 2017;**107**:229-235.e2.

Carrasco B, Arroyo G, Gil Y, Gómez MJ, Rodríguez I, Barri PN, Veiga A, Boada M. Selecting embryos with the highest implantation potential using data mining and decision tree based on classical embryo morphology and morphokinetics. *J Assist Reprod Genet* 2017;**34**:983–990. Journal of Assisted Reproduction and Genetics.

Cetinkaya M, Pirkevi C, Yelke H, Colakoglu YK, Atayurt Z, Kahraman S. Relative kinetic expressions defining cleavage synchronicity are better predictors of blastocyst formation and quality than absolute time points. *J Assist Reprod Genet* 2014;**32**:27–35.

Chamayou S, Patrizio P, Storaci G, Tomaselli V, Alecci C, Ragolia C, Crescenzo C, Guglielmino A. The use of morphokinetic parameters to select all embryos with full capacity to implant. *J Assist Reprod Genet* 2013;**30**:703–710.

Chavez SL, Loewke KE, Han J, Moussavi F, Colls P, Munne S, Behr B, Reijo Pera RA. Dynamic blastomere behaviour reflects human embryo ploidy by the four-cell stage. *Nat Commun* [Internet] 2012;**3**:1–12. Nature Publishing Group.

Ciray HN, Campbell A, Agerholm IE, Aguilar J, Chamayou S, Esbert M, Sayed S. Proposed guidelines on the nomenclature and annotation of dynamic human embryo monitoring by a time-lapse user group. *Hum Reprod* 2014;**29**:2650–2660.

Cobo A, Meseguer M, Remohi J, Pellicer A. Use of cryo-banked oocytes in an ovum donation programme: a prospective, randomized, controlled, clinical trial. *Hum Reprod* 2010;**25**:2239–2246.

Cohen J, Grudzinskas G, Johnson MH. Embryonic DNA sampling without biopsy: The beginnings of non-invasive PGD? *Reprod Biomed Online* [Internet] 2013;**26**:520–521. Reproductive Healthcare Ltd.

Conaghan J, Chen AA, Willman SP, Ivani K, Chenette PE, Boostanfar R, Baker VL, Adamson GD, Abusief ME, Gvakharia M, *et al.* Improving embryo selection using a computer-automated time-lapse image analysis test plus day 3 morphology: Results from a prospective multicenter trial. *Fertil Steril* [Internet] 2013;**100**:412-419.e5. Elsevier Inc.

Coticchio G, Lagalla C, Sturmey R, Pennetta F, Borini A. The enigmatic morula: Mechanisms of development, cell fate determination, self-correction and implications for ART. *Hum Reprod Update* 2019;**25**:422–438.

Coticchio G, Renzini MM, Novara P V., Lain M, Ponti E De, Turchi D, Fadini R, Canto MD. Focused time-lapse analysis reveals novel aspects of human fertilization and suggests new parameters of embryo viability. *Hum Reprod* 2018;**33**:23–31.

Cruz M, Gadea B, Garrido N, Pedersen KS, Martínez M, Pérez-Cano I, Muñoz M, Meseguer M. Embryo quality, blastocyst and ongoing pregnancy rates in oocyte donation patients whose embryos were monitored by time-lapse imaging. *J Assist Reprod Genet* 2011;**28**:569–573.

Cruz M, Garrido N, Herrero J, Pérez-Cano I, Muñoz M, Meseguer M. Timing of cell division in human cleavage-stage embryos is linked with blastocyst formation and quality. *Reprod Biomed Online* 2012;**25**:371–381.

Dal Canto M, Coticchio G, Mignini Renzini M, Ponti E De, Vittoria Novara P, Brambillasca F, Comi R, Fadini R. Cleavage kinetics analysis of human embryos predicts development to blastocyst and implantation. *Reprod Biomed Online* [Internet] 2012;**25**:474–480.

Danuser G. Computer Vision in Cell Biology. *Cell* [Internet] 2011;**147**:973–978. Elsevier Inc.

Desai N, Goldberg JM, Austin C, Falcone T. Are cleavage anomalies, multinucleation, or specific cell cycle kinetics observed with time-lapse imaging predictive of embryo developmental capacity or ploidy? *Fertil Steril* [Internet] 2018;**109**:665–674. Elsevier Inc.

Desai N, Ploskonka S, Goodman L, Attaran M, Goldberg JM, Austin C, Falcone T. Delayed blastulation, multinucleation, and expansion grade are independently associated with live-birth rates in frozen blastocyst transfer cycles. *Fertil Steril* [Internet] 2016;**106**:1370–1378. Elsevier Inc.

Desai N, Ploskonka S, Goodman LR, Austin C, Goldberg J, Falcone T. Analysis of embryo morphokinetics, multinucleation and cleavage anomalies using continuous time-lapse monitoring in blastocyst transfer cycles. *Reprod Biol Endocrinol* 2014;**12**:1–10.

Desch L, Bruno C, Luu M, Barberet J, Choux C, Lamotte M, Schmutz E, Sagot P, Fauque P. Embryo multinucleation at the two-cell stage is an independent predictor of intracytoplasmic sperm injection outcomes. *Fertil Steril* 2017;**107**:.

Diamond MP, Suraj V, Behnke EJ, Yang X, Angle MJ, Lambe-Steinmiller JC, Watterson R, Athayde Wirka K, Chen AA, Shen S. Using the Eeva Test™ adjunctively to traditional day 3 morphology is informative for consistent embryo assessment within a panel of embryologists with diverse experience. *J Assist Reprod Genet* 2015;**32**:61–68.

Dominguez F, Meseguer M, Aparicio-Ruiz B, Piqueras P, Quiñonero A, Simón C. New strategy for diagnosing embryo implantation potential by combining proteomics and time-lapse technologies. *Fertil Steril* 2015;**104**:908–914.

Ebner T, Höggerl A, Oppelt P, Radler E, Enzelsberger SH, Mayer RB, Petek E, Shebl O. Time-lapse imaging provides further evidence that planar arrangement of blastomeres is highly abnormal. *Arch Gynecol Obstet* 2017;**296**:1199–1205. Springer Berlin Heidelberg.

Elder K, Dale B. In-vitro fertilization. *3rd ed Cambridge Cambridge Univ Press* 2011;**112**:.

ESHRE Working group on Time-lapse technology, Apter S, Ebner T, Freour T, Guns Y, Kovacic B, Clef N Le, Marques M, Meseguer M, Montjean D, *et al.* Good practice recommendations for the use of time-lapse technology. *Hum Reprod Open* 2020;**2020**:1–26.

Feyeux M, Reignier A, Mocaer M, Lammers J, Meistermann D, Barrière P, Paul-Gilloteaux P, David L, Fréour T. Development of automated annotation software for human embryo morphokinetics. *Hum Reprod* 2020;**35**:557–564.

Fragouli E, Wells D. Aneuploidy in the human blastocyst. *Cytogenet Genome Res* 2011;**133**:149–159.

Franasiak JM, Forman EJ, Hong KH, Werner MD, Upham KM, Treff NR, Scott RT. The nature of aneuploidy with increasing age of the female partner: A review of 15,169 consecutive trophectoderm biopsies evaluated with comprehensive chromosomal screening. *Fertil Steril* [Internet] 2014;**101**:656-663.e1. Elsevier Inc.

Fréour T, Fleuter N Le, Lammers J, Splingart C, Reignier A, Barrière P. External validation of a time-lapse prediction model. *Fertil Steril* 2015;**103**:917–922.

Gardner DK, Balaban B. Assessment of human embryo development using morphological criteria in an era of time- lapse, algorithms and ‘OMICS’: Is looking good still important? *Mol Hum Reprod* 2016;

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**:1175–1180.

Gardner DK, Schoolcraft WB. In Vitro Culture of Human Blastocyst. *Towar Reprod Certain Infertil Genet Beyond* 1999;**Parthenon**:377–388.

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**:1981–1986.

Garrido-Gómez T, Ruiz-Alonso M, Blesa D, Diaz-Gimeno P, Vilella F, Simón C. Profiling the gene signature of endometrial receptivity: Clinical results. *Fertil Steril* 2013;**99**:, p.

1078–1085.

Glujovsky D, Farquhar C, Quinteiro Retamar AM, Alvarez Sedo CR, Blake D. Cleavage stage versus blastocyst stage embryo transfer in assisted reproductive technology. *Cochrane Database Syst Rev* 2016;**2016**:.

Goodman LR, Goldberg J, Falcone T, Austin C, Desai N. Does the addition of time-lapse morphokinetics in the selection of embryos for transfer improve pregnancy rates? A randomized controlled trial. *Fertil Steril* [Internet] 2016;**105**:275-285.e10. Elsevier Inc.

Gore A, Chappell V, Fenton S. EDC-2: The Endocrine Society's Second Scientific Statement on Endocrine-Disrupting Chemicals. *Endocr Rev* 2015;**36**:E1–E150.

Hanson BM, Tao X, Hong KH, Comito CE, Pangasnan R, Seli E, Jalas C, Scott RT. Noninvasive preimplantation genetic testing for aneuploidy exhibits high rates of deoxyribonucleic acid amplification failure and poor correlation with results obtained using trophectoderm biopsy. *Fertil Steril* [Internet] 2021;**115**:1461–1470. Elsevier Inc.

Insua MF, Cobo AC, Larreategui Z, Ferrando M, Serra V, Meseguer M. Obstetric and perinatal outcomes of pregnancies conceived with embryos cultured in a time-lapse monitoring system. *Fertil Steril* [Internet] 2017;**108**:498–504. Elsevier Inc.

Iwata K, Yumoto K, Sugishima M, Mizoguchi C, Kai Y, Iba Y, Mio Y. Analysis of compaction initiation in human embryos by using time-lapse cinematography. *J Assist Reprod Genet* 2014;**31**:421–426.

Jedrusik A. Making the first decision: lessons from the mouse. *Reprod Med Biol* 2015;**14**:135–150.

Juneau C, Franasiak J, Treff N. Challenges facing contemporary preimplantation genetic screening. *Curr Opin Obstet Gynecol* [Internet] 2016;**28**:151–157.

Kahraman S, Cetinkaya M, Pirkevi C, Yelke H, Kumtepe Y. Comparison of blastocyst development and cycle outcome in patients with eSET using either conventional or time lapse incubators. A prospective study of good prognosis patients. *J Reprod Biotechnol Fertil* [Internet] 2013;**3**:55–61.

Kaser DJ, Bormann CL, Missmer SA, Farland L V., Ginsburg ES, Racowsky C. A pilot randomized controlled trial of Day 3 single embryo transfer with adjunctive time-lapse selection versus Day 5 single embryo transfer with or without adjunctive time-lapse selection. *Hum Reprod* 2017;**32**:1598–1603.

Kaser DJ, Racowsky C. Clinical outcomes following selection of human preimplantation

embryos with time-lapse monitoring: A systematic review. *Hum Reprod Update* 2014;**20**:617–631.

Kato K, Ueno S, Berntsen J, Ito M, Shimazaki K, Uchiyama K, Okimura T. Comparing prediction of ongoing pregnancy and live birth outcomes in patients with advanced and younger maternal age patients using KIDScore^{TM} day 5: a large-cohort retrospective study with single vitrified-warmed blastocyst transfer. *Reprod Biol Endocrinol* 2021;**19**:1–11. Reproductive Biology and Endocrinology.

Kieslinger DC, Gheselle S De, Lambalk CB, Sutter P De, Kostelijk EH, Twisk JWR, Rijswijk J Van, Abbeel E Van Den, Vergouw CG. Embryo selection using time-lapse analysis (Early Embryo Viability Assessment) in conjunction with standard morphology: A prospective two-center pilot study. *Hum Reprod* 2016;**31**:2450–2457.

Kirkegaard K, Hindkjaer JJ, Grøndahl ML, Kesmodel US, Ingerslev HJ. A randomized clinical trial comparing embryo culture in a conventional incubator with a time-lapse incubator. *J Assist Reprod Genet* 2012a;**29**:565–572.

Kirkegaard K, Juhl Hindkjaer J, Ingerslev HJ. Human embryonic development after blastomere removal: A time-lapse analysis. *Hum Reprod* 2012b;**27**:97–105.

Kirkegaard K, Kesmodel US, Hindkjær JJ, Ingerslev HJ. Time-lapse parameters as predictors of blastocyst development and pregnancy outcome in embryos from good prognosis patients: A prospective cohort study. *Hum Reprod* 2013;**28**:2643–2651.

Kovacs P, Matyas S, Forgacs V, Sajgo A, Rarosi F, Pribenszky C. Time-lapse embryo selection for single blastocyst transfer – results of a multicenter, prospective, randomized clinical trial. *Fertil Steril* [Internet] 2013;**100**:S90. Elsevier Ltd.

Kramer YG, Kofinas JD, Melzer K, Noyes N, McCaffrey C, Buldo-Licciardi J, McCulloh DH, Grifo JA. Assessing morphokinetic parameters via time lapse microscopy (TLM) to predict euploidy: are aneuploidy risk classification models universal? *J Assist Reprod Genet* 2014;**31**:1231–1242.

Leaver M, Wells D. Non-invasive preimplantation genetic testing (niPGT): The next revolution in reproductive genetics? *Hum Reprod Update* 2020;**26**:16–42.

Leese H, Guerif F, Allgar V, Brison D, Lundin K, Sturmey R. Biological optimization, the Goldilocks principle, and how much is lagom in the preimplantation embryo. *Mol Reprod Dev* 2016;**83**:748–754.

Lipari CW, Garcia JE, Zhao Y, Thrift K, Vaidya D, Rodriguez A. Nitric oxide metabolite

production in the human preimplantation embryo and successful blastocyst formation. *Fertil Steril* 2009;**91**:1316–1318.

Lissak A, Wiener-Megnazi Z, Reznick AZ, Shnizer S, Ishai D, Grach B, Lahav-Baratz S, Shiloh H, Koifman M, Dirnfeld M. Oxidative stress indices in seminal plasma, as measured by the thermochemiluminescence assay, correlate with sperm parameters. *Fertil Steril* 2004;**81**:792–797.

Liu Y, Chapple V, Feenan K, Roberts P, Matson P. Time-lapse videography of human embryos: Using pronuclear fading rather than insemination in IVF and ICSI cycles removes inconsistencies in time to reach early cleavage milestones. *Reprod Biol* 2015a;.

Liu Y, Chapple V, Feenan K, Roberts P, Matson P. Time-lapse deselection model for human day 3 in vitro fertilization embryos: The combination of qualitative and quantitative measures of embryo growth. *Fertil Steril* [Internet] 2016;**105**:656-662e1. Elsevier Inc.

Liu Y, Copeland C, Stevens A, Feenan K, Chapple V, Myssonski K, Roberts P, Matson P. Assessment of human embryos by time-lapse videography: A comparison of quantitative and qualitative measures between two independent laboratories. *Reprod Biol* [Internet] 2015b;**15**:210–216. Society for Biology of Reproduction & the Institute of Animal Reproduction and Food Research of Polish Academy of Sciences in Olsztyn.

Liu Y, Feenan K, Chapple V, Matson P. Assessing efficacy of day 3 embryo time-lapse algorithms retrospectively: impacts of dataset type and confounding factors. *Hum Fertil* [Internet] 2018;**0**:1–9. Informa UK Limited, trading as Taylor & Francis Group.

Mahfouz R, Sharma R, Sharma D, Sabanegh E, Agarwal A. Diagnostic value of the total antioxidant capacity (TAC) in human seminal plasma. *Fertil Steril* [Internet] 2009;**91**:805–811. Elsevier Ltd.

Martínez-Granados L, Serrano M, González-Utor A, Ortíz N, Badajoz V, Olaya E, Prados N, Castilla JA. Inter-laboratory agreement on embryo classification and clinical decision : Conventional morphological assessment vs . time lapse. *PLoS One* 2017;**12**:1–13.

Mascarenhas M, Flaxman S, Boerma T. National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. *PLoS Med* 2012;**9**:e1001356.

Meng Z, Marta N S, Angel M, Chuanxin Z, Mate B, Rachel S M, Richard J P, Matteo A M, Marga E, Richard T S, *et al.* Mechanism of cell polarisation and first lineage segregation

in the human embryo. *bioRxiv* 2020;

Meseguer M. Time-lapse: the remaining questions to be answered. *Fertil Steril* [Internet] 2016;**105**:295–296. American Society for Reproductive Medicine.

Meseguer M, Herrero J, Tejera A, Hilligsøe KM, Ramsing NB, Remohi J. The use of morphokinetics as a predictor of embryo implantation. *Hum Reprod* 2011;**26**:2658–2671.

Meseguer M, Rubio I, Cruz M, Basile N, Marcos J, Requena A. Embryo incubation and selection in a time-lapse monitoring system improves pregnancy outcome compared with a standard incubator: A retrospective cohort study. *Fertil Steril* [Internet] 2012;**98**:1481-1489.e10. Elsevier Inc.

Milewski R, Czerniecki J, Kuczyńska A, Stankiewicz B, Kuczyński W. Morphokinetic parameters as a source of information concerning embryo developmental and implantation potential. *Ginekol Pol* 2016;**87**:677–684.

Milewski R, Kuc P, Kuczynska A, Stankiewicz B, Lukaszuk K, Kuczynski W. A predictive model for blastocyst formation based on morphokinetic parameters in time-lapse monitoring of embryo development. *J Assist Reprod Genet* 2015;**32**:571–579.

Milewski R, Kuczyńska A, Stankiewicz B, Kuczyński W. How much information about embryo implantation potential is included in morphokinetic data? A prediction model based on artificial neural networks and principal component analysis. *Adv Med Sci* 2017;**62**:202–206.

Minasi MG, Colasante A, Riccio T, Ruberti A, Casciani V, Scarselli F, Spinella F, Fiorentino F, Varricchio MT, Greco E. Correlation between aneuploidy, standard morphology evaluation and morphokinetic development in 1730 biopsied blastocysts: A consecutive case series study. *Hum Reprod* 2016;**31**:2245–2254.

Mizobe Y, Ezono Y, Tokunaga M, Oya N, Iwakiri R, Yoshida N, Sato Y, Onoue N, Miyoshi K. Selection of human blastocysts with a high implantation potential based on timely compaction. *J Assist Reprod Genet* 2017;**34**:991–997. Journal of Assisted Reproduction and Genetics.

Motato Y, los Santos MJ de, Escriba MJ, Ruiz BA, Remohí J, Meseguer M. Morphokinetic analysis and embryonic prediction for blastocyst formation through an integrated time-lapse system. *Fertil Steril* 2016;**105**:376-384.e9.

Muñoz M, Cruz M, Humaidan P, Garrido N, Pérez-Cano I, Meseguer M. Dose of recombinant

FSH and oestradiol concentration on day of HCG affect embryo development kinetics. *Reprod Biomed Online* 2012;**25**:382–389.

Nakahara T, Iwase A, Goto M, Harata T, Suzuki M, Ienaga M, Kobayashi H, Takikawa S, Manabe S, Kikkawa F, *et al.* Evaluation of the safety of time-lapse observations for human embryos. *J Assist Reprod Genet* 2010;**27**:93–96.

Park H, Bergh C, Selleskog U, Thurin-Kjellberg A, Lundin K. No benefit of culturing embryos in a closed system compared with a conventional incubator in terms of number of good quality embryos: Results from an RCT. *Hum Reprod* 2015;**30**:268–275.

Paszkowski T, Clarke RN. Antioxidative capacity of preimplantation embryo culture medium declines following the incubation of poor quality embryos. *Hum Reprod* 1996;**11**:2493–2495.

Paternot G, Wetsels AM, Thonon F, Vansteenbrugge A, Willemen D, Devroe J, Debrock S, Hooghe TMD, Spiessens C. Intra- and interobserver analysis in the morphological assessment of early stage embryos during an IVF procedure : a multicentre study. *Reprod Biol Endocrinol* [Internet] 2011;**9**:127. BioMed Central Ltd.

Paulson RJ, Reichman DE, Zaninovic N, Goodman LR, Racowsky C. Time-lapse imaging: clearly useful to both laboratory personnel and patient outcomes versus just because we can doesn't mean we should. *Fertil Steril* [Internet] 2018;**109**:584–591. Elsevier Inc.

Petersen BM, Boel M, Montag M, Gardner DK. Development of a generally applicable morphokinetic algorithm capable of predicting the implantation potential of embryos transferred on Day 3. *Hum Reprod* 2016;**31**:2231–2244.

Pribenszky C, Losonczi E, Molnar M, Lang Z, Matyas S, Rajczy K, Molnar K, Kovacs P, Nagy P, Conceicao J, *et al.* Prediction of in-vitro developmental competence of early cleavage-stage mouse embryos with compact time-lapse equipment. *Reprod Biomed Online* 2010;**20**:371–379.

Pribenszky C, Nilselid AM, Montag M. Time-lapse culture with morphokinetic embryo selection improves pregnancy and live birth chances and reduces early pregnancy loss: a meta-analysis. *Reprod Biomed Online* [Internet] 2017;**35**:511–520. Elsevier Ltd.

Racowsky C, Martins WP. Effectiveness and safety of time-lapse imaging for embryo culture and selection : it is still too early for any conclusions? *Fertil Steril* 2017;**108**:3–5.

Rienzi L, Capalbo A, Stoppa M, Romano S, Maggiulli R, Albricci L, Scarica C, Farcomeni A, Vajta G, Ubaldi FM. No evidence of association between blastocyst aneuploidy and

morphokinetic assessment in a selected population of poor-prognosis patients: A longitudinal cohort study. *Reprod Biomed Online* [Internet] 2015;**30**:57–66. Reproductive Healthcare Ltd.

Rienzi L, Cimadomo D, Delgado A, Minasi MG, Fabozzi G, Gallego R del, Stoppa M, Bellver J, Giancani A, Esbert M, *et al.* Time of morulation and trophectoderm quality are predictors of a live birth after euploid blastocyst transfer: a multicenter study. *Fertil Steril* [Internet] 2019;**112**:1080-1093.e1. Elsevier Inc.

Rubio C, Navarro-Sánchez L, García-Pascual CM, Ocali O, Cimadomo D, Venier W, Barroso G, Kopcow L, Bahçeci M, Kulmann MIR, *et al.* Multicenter prospective study of concordance between embryonic cell-free DNA and trophectoderm biopsies from 1301 human blastocysts. *Am J Obstet Gynecol* [Internet] 2020;**223**:751.e1-751.e13. Elsevier Inc.

Rubio I, Galán A, Larreategui Z, Ayerdi F, Bellver J, Herrero J, Meseguer M. Clinical validation of embryo culture and selection by morphokinetic analysis: A randomized, controlled trial of the EmbryoScope. *Fertil Steril* 2014;**102**:1287–1294.

Rubio I, Kuhlmann R, Agerholm I, Kirk J, Herrero J, Escribá MJ, Bellver J, Meseguer M. Limited implantation success of direct-cleaved human zygotes: A time-lapse study. *Fertil Steril* 2012;**98**:1458–1463.

Ruiz de Assin R, Clavero A, Gonzalvo MC, Ramirez JP, Zamora S, Fernandez A, Martinez L, Castilla JA. Comparison of methods to determine the assigned value in an external quality control programme for embryo evaluation. *Reprod Biomed Online* 2009;**19**:824–829.

Rutstein S, Shah I. Infecundity infertility and childlessness in developing countries. *Geneva World Heal Organ* 2004;

Sallam HN, Sallam NH, Sallam SH. Non-invasive methods for embryo selection. *Facts, views Vis ObGyn* [Internet] 2016;**8**:87–100.

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.

Segal T, Giudice L. Before the beginning: environmental exposures and reproductive and obstetrical outcomes. *Fertil Steril* 2019;**112**:613–621.

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* [Internet] 2008;**90**:2183–2189. Elsevier Ltd.

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**:1350–1357.

Sharma R, Pasqualotto F, Nelson D, Thomas AJ, Agarwal A. The reactive oxygen species—total antioxidant capacity score is a new measure of oxidative stress to predict male infertility. *Hum Reprod* 1999;**14**:2801–2807.

Shnizer S, Kagan T, Lanir A, Maor I, Reznick AZ. Modifications and oxidation of lipids and proteins in human serum detected by thermochemiluminescence. *Luminescence* 2003;**18**:90–96.

Siristatidis C, Komitopoulou MA, Makris A, Sialakouma A, Botzaki M, Mastorakos G, Salamalekis G, Bettocchi S, Palmer GA. Morphokinetic parameters of early embryo development via time lapse monitoring and their effect on embryo selection and ICSI outcomes: a prospective cohort study. *J Assist Reprod Genet* 2015;**32**:563–570.

Storr A, Venetis C, Cooke S, Kilani S, Ledger W. Time-lapse algorithms and morphological selection of day-5 embryos for transfer: a preclinical validation study. *Fertil Steril* [Internet] 2018;**109**:276-283.e3. Elsevier Inc.

Storr A, Venetis CA, Cooke S, Susetio D, Kilani S, Ledger W. Morphokinetic parameters using time-lapse technology and day 5 embryo quality: a prospective cohort study. *J Assist Reprod Genet* 2015;**32**:1151–1160.

Sundvall L, Ingerslev HJ, Breth Knudsen U, Kirkegaard K. Inter- and intra-observer variability of time-lapse annotations. *Hum Reprod* 2013;**28**:3215–3221.

Swain JE. Decisions for the IVF laboratory: Comparative analysis of embryo culture incubators. *Reprod Biomed Online* [Internet] 2014;**28**:535–547. Reproductive Healthcare Ltd.

Tejera A, Aparicio-Ruiz B, Meseguer M. The use of morphokinetic as a predictor of implantation. *Minerva Ginecol* 2017;**69**:.

Tejera A, Castelló D, los Santos JM de, Pellicer A, Remohí J, Meseguer M. Combination of metabolism measurement and a time-lapse system provides an embryo selection

method based on oxygen uptake and chronology of cytokinesis timing. *Fertil Steril* 2016;**106**:119-126.e2.

Tejera A, Herrero J, Los Santos MJ De, Garrido N, Ramsing N, Meseguer M. Oxygen consumption is a quality marker for human oocyte competence conditioned by ovarian stimulation regimens. *Fertil Steril* 2011;**96**:.

The Practice Committee of the Society for Assisted Reproductive Technology and Practice Committee of the American Society for Reproductive Medicine. Preimplantation genetic testing : a Practice Committee opinion. *Fertil Steril* [Internet] 2008;**90**:136–143.

Ueno S, Berntsen J, Ito M, Okimura T, Kato K. Correlation between an annotation-free embryo scoring system based on deep learning and live birth/neonatal outcomes after single vitrified-warmed blastocyst transfer: a single-centre, large-cohort retrospective study. *J Assist Reprod Genet* [Internet] 2022;**39**:2089–2099. Springer US.

Vera-Rodriguez M, Chavez SL, Rubio C, Reijo Pera RA, Simon C. Prediction model for aneuploidy in early human embryo development revealed by single-cell analysis. *Nat Commun* [Internet] 2015;**6**:1–14. Nature Publishing Group.

Vergouw C, Heymans M, Hardarson T, Sfontouris I, Economou K, Ahlstrom A. No evidence that embryo selection by near-infrared spectroscopy in addition to morphology is able to improve live birth rates: results from an individual patient data meta-analysis. *Hum Reprod* 2014;**29**:455–461.

Vergouw CG, Botros LL, Roos P, Lens JW, Schats R, Hompes PGA, 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**:1499–1504.

Vermilyea MD, Tan L, Anthony JT, Conaghan J, Ivani K, Gvakharia M, Boostanfar R, Baker VL, Suraj V, Chen AA, *et al.* Computer-automated time-lapse analysis results correlate with embryo implantation and clinical pregnancy: A blinded, multi-centre study. *Reprod Biomed Online* [Internet] 2014;**29**:729–736. Reproductive Healthcare Ltd.

Weimar CH, Uiterweer EDP, Teklenburg G, Heijnen CJ, Macklon NS. In-vitro model systems for the study of human embryo – endometrium interactions. *Reprod Biomed Online* [Internet] 2013;**27**:461–476. Reproductive Healthcare Ltd.

Wiener-Megnazi Z, Schnizer S, Lahav-Baratz S, Shiloh H, Koifman M, Dirnfeld M. Electronically excited species (EES) indices correlate with women's age and the chances

of conception in ART. *Fertil Steril* 2002;2002.

Wiener-Megnazi Z, Shiloh H, Sc M, Avraham L, Lahav-baratz S, Ph D. Oxidative parameters of embryo culture media may predict treatment outcome in in vitro fertilization : a novel applicable tool for improving embryo selection. *Fertil Steril* [Internet] 2011;**95**:979–984. Elsevier Ltd.

Wiener-megnazi Z, Vardi L, Lissak A, Shnizer S, Reznick AZ, Ishai D, Lahav-baratz S, Shiloh H, Koifman M, Dirnfeld M. Oxidative stress indices in follicular fluid as measured by the thermochemiluminescence assay correlate with outcome parameters in in vitro fertilization. *Fertil Steril* 2004;**82**:1171–1176.

Wong CC, Loewke KE, Bossert NL, Behr B, Jonge CJ De, Baer TM, Pera RAR. Non-invasive imaging of human embryos before embryonic genome activation predicts development to the blastocyst stage. *Nat Biotechnol* [Internet] 2010;**28**:1115–1121. Nature Publishing Group.

World Health Organization W. International Classification of Diseases, 11th Revision (ICD-11) Geneva: WHO 2018. 2018;

Wu Y, Lazzaroni-tealdi E, Wang Q, Zhang L, Barad DH, Kushnir VA, Darmon SK, Albertini DF, Gleicher N. Different effectiveness of closed embryo culture system with time-lapse imaging (EmbryoScope TM) in comparison to standard manual embryology in good and poor prognosis patients : a prospectively randomized pilot study. *Reprod Biol Endocrinol* [Internet] 2016;**14**:1–11. Reproductive Biology and Endocrinology.

Yang L, Cai S, Zhang S, Kong X, Gu Y, Lu C, Dai J, Gong F, Lu G, Lin G. Single embryo transfer by Day 3 time-lapse selection versus Day 5 conventional morphological selection: A randomized, open-label, non-inferiority trial. *Hum Reprod* 2018;**33**:869–876.

Yang L, Kong X, Zhang S, Dai J, Gong F, Lu G, Lin G. Single embryo transfer on cleavage-stage(D3) using timelapse selection VS on blastocyst(D5) using traditional morphological selection in patients with good prognosis: a prospective randomized controlled trial. *Hum Reprod* 2017;**32**:i102-103.

Yang ST, Shi JX, Gong F, Zhang SP, Lu CF, Tan K, Leng LZ, Hao M, He H, Gu YF, *et al.* Cleavage pattern predicts developmental potential of day 3 human embryos produced by IVF. *Reprod Biomed Online* 2015;**30**:625–634.

Yeung S, Downing NL, Fei-Fei L, Milstein A. Bedside Computer Vision — Moving Artificial Intelligence from Driver Assistance to Patient Safety. *N Engl J Med* [Internet]

2018;**378**:1271–1273.

Zaninovic N, Irani M, Meseguer M. Assessment of embryo morphology and developmental dynamics by time-lapse microscopy: is there a relation to implantation and ploidy? *Fertil Steril* [Internet] 2017;**108**:722–729. Elsevier Inc.

Zaninovic N, Zhan Q, Clarke R, Ye Z, Pereira N, Rosenwaks Z. Impact of time-lapse and reduced oxygen culture on live birth rate and its correlation with infertility diagnosis. *Fertil Steril* [Internet] 2015;**104**:e95.

Zegers-Hochschild F, Dickens B, Dughman-Manzur S. Human rights to in vitro fertilization. *Int J Gynecol Obstet* 2013;**123**:86–89.

Zhan Q, Ye Z, Clarke R, Rosenwaks Z, Zaninovic N. Direct unequal cleavages: Embryo developmental competence, genetic constitution and clinical outcome. *PLoS One* 2016;**11**:1–19.