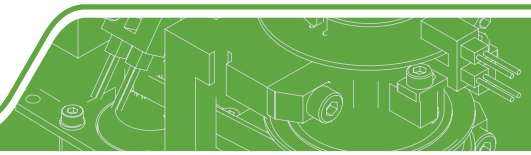

Geri Assess® 2.0

Annotation Accuracy

Geri Assess® 2.0 software has been designed to automatically identify and record embryo key developmental events and observations as they occur (see Table 1 and QRTM242 Geri Assess® 2.0 Automated Software Tool Tech Note for details).

Geri Assess® 2.0 performance testing

To assess software performance in a clinical setting, the accuracy of automated developmental event annotations was tested by comparing them to manual annotations. Manual annotations were performed by three experienced clinical embryologists reviewing the same set of embryo videos as the software. To minimise subjectivity, only events annotated by all three users were defined as a 'true' observed event and included into the analysis. Averages of the event timings were then calculated.

Observation (*e.g.* reverse cleavage and fragmentation) accuracy was tested by only one experienced embryologist, as these observations are less prone to variations between individual annotators and relied solely on the event being detected.

The tests were conducted in two Genea clinics using five Geri® systems. The average incubators' system load was approximately 6 patients or 31 embryos per incubator per week. Test groups comprised of clinically utilised and non-utilised embryos to address both Day 3 and Day 5 embryo transfer scenarios, including biopsied and non-biopsied embryos. Over 3,700 events from >400 embryos from 69 patients were included in the tests.

The tested parameters for key developmental events included:

- **Detection Rate**
- **Accuracy**
- **False Positive Detection**

Detection Rate is whether Geri Assess® 2.0 successfully detected and annotated a 'true' observed event, based on it being manually annotated by 3 annotators. Accuracy was

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
	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.

Table 1 - Key developmental Events and Observations and their definitions automatically annotated by Geri Assess® 2.0. Observations are marked with an asterisk*.

P: +44 1304 897 414
E: info@geneabiomedx.com
W: www.geneabiomedx.com

determined by two independent measures, namely comparing event timings directly between Geri Assess® 2.0 and manual annotations, and by assessing inter-observer error between the two. False Positive Detection *i.e.* the probability of Geri Assess® 2.0 detecting an event when no 'true' Event occurred, was also tested.

The pre-defined criteria for detection and accuracy consisted of pass rates and time ranges, which varied from ±20 minutes for early cleavage events and from ±150 to ±400 minutes for later stage events. These variations reflect increasing biological variation, manual annotations differences, and decreasing risk of impact on clinical outcomes at those stages. Pass rates for most events were 85%, and for False Positives was <20%.

The tested parameters for the observations of Reverse Cleavage/Failed Cytokinesis and Fragmentation, included the rates of:

- **False Positives**
- **False Negatives**

False Positive and False Negative rates are the probability of Geri Assess® 2.0 either recording an observation when it didn't occur, or not recording a 'true' observation.

Fragmentation was assessed by comparing several fragmentation periods during the embryo development between manual and automated annotations and recording disagreements rising from either.

Pass rates for both False Positive and False Negative rates were set at <20%. In addition, Detection Rate was recorded for the Reverse Cleavage/Failed Cytokinesis.

Results

Developmental event Detection and Accuracy results are shown in Table 2, presented as percentage and number of events that passed the testing criteria. It is to be noted that not all embryos underwent all events.

Combined False Positive Detection rate for all developmental events was 16% (42/269).

Detection rate for Reverse Cleavage/Failed Cytokinesis was 72% (18/25).

Fragmentation testing results showed only 5% (6/116) False Positive rates and 17% (10/60) False Negative rates with Geri Assess® 2.0.

Discussion

As anticipated, annotations of normally developing embryos without significant irregular cellular events were more accurate than those with embryos exhibiting irregular morphology or behaviour. This was especially notable in cases where embryos showed abnormal developmental patterns,; such embryos often resulting in disagreements between automated and manual annotations in several developmental events.

The main purpose of these studies was to evaluate the performance of automated annotations software in a clinical setting, rather than to embark on a scientific or clinical validation of the software features. Therefore, it is important to note that automated annotations are not designed to completely replace assessments by embryologists. Hence, users are encouraged to conduct additional manual reviews, especially with poor quality embryos and those missing automated annotations.

Conclusions

Geri Assess® 2.0 software can perform automated annotations of critical embryo developmental events and observations to the level that can support clinics in their embryo assessments. This allows implementation of the Geri system® to assist in improving routine clinical workflows.

For further details about Geri Assess® 2.0, see QRTM242 Geri Assess® 2.0 Automated Software Tool Tech Note.

Feature	PNa	PNd	2-cell	3-cell	4-cell	5-cell	6-cell	M	EB	XB[1]	HgB[1]
Detection	**99%** (70/71)	**96%** (131/136)	**99%** (143/144)	**87%** (122/140)	**88%** (165/188)	**93%** (129/139)	**78%** (145/187)	**89%** (122/137)	**91%** (116/127)	**79%** (46/58)	**83%** (74/89)
Accuracy	**59%** (41/70)	**92%** (121/131)	**97%** (139/143)	**93%** (114/122)	**85%** (141/165)	**87%** (112/129)	**87%** (126/145)	**82%** (100/122)	**92%** (107/116)	**89%** (41/46)	**58%** (43/74)

Table 2 - Automated Annotations Detection and Accuracy outcomes for all embryo development events

1 Includes embryos biopsied for preimplantation genetic testing, representing approx. 77% of embryos.

P: +44 1304 897 414
E: info@geneabiomedx.com
W: www.geneabiomedx.com
Genea BIOMEDX