# AN AUTOMATED SEGMENTATION SYSTEM FOR OOCYTES WITH SMOOTH ENDOPLASMIC RETICULUM CLUSTERS BASED ON DEEPLAB

**SATOI SHIMOTSU[1], KENTARO MORI[1], KIMITOH KOBAYASHI[2], SHIMPEI MIZUTA[2], TAKUMI TAKEUCHI[2], TOMOMOTO ISHIKAWA[2], YUTAKA HATA[3], NAOMI YAGI[4]**

[1] National Institute of Technology (KOSEN), Maizuru College, Kyoto, Japan

[2] Reproduction Clinic Tokyo, Tokyo, Japan

[3] Graduate School of Information Science, University of Hyogo, Hyogo, Japan

[4] Advanced Medical Engineering Research Institute, University of Hyogo, Hyogo, Japan

E-MAIL: k.mori@maizuru-ct.ac.jp

**Abstract:**

**Intracytoplasmic sperm injection (ICSI) is a widely used infertility treatment, but its success depends on various factors, including smooth endoplasmic reticulum clusters (sERCs) in oocytes. However, evaluation of sERCs is often subjective, leading to inconsistent findings across studies. To address this issue, this study proposes an objective approach using DeepLab, a deep learning-based segmentation model, to analyze regions of sERC, oocyte, and needle in ICSI procedure videos. A DeepLab model was trained to segment these regions, and features were extracted automatically to assess their consistency with manual annotations. Key metrics such as sERC circularity and area ratio showed good agreement, while features involving needle tip position and small-scale values showed larger relative errors due to boundary detection challenges. Despite these limitations, the results demonstrated that this method provides a reliable foundation for consistent, quantitative analysis in ICSI studies and may support more standardized assessments of factors influencing treatment outcomes.**

**Keywords:**

**DeepLab; ICSI; Infertility; Segmentation; sERC;**

## 1. Introduction

In contemporary Japan, declining birthrates have become a serious social issue, and this trend has led to increasing public interest in infertility treatments. Among these, intracytoplasmic sperm injection (ICSI), in which a single sperm is directly injected into an oocyte, has attracted considerable attention. ICSI is considered one of the most promising methods of artificially assisting fertilization. The method can theoretically achieve fertilization with a single sperm by delivering it directly into the cytoplasm of an oocyte. Furthermore, it reduces the risk of polyspermy by enabling complete control over the fertilization process. The method has a high fertilization rate of approximately 40% to 60% [1].

It is known that various factors influence the success of ICSI. Understanding and managing these factors appropriately is crucial for improving treatment outcomes. Previous studies have highlighted the importance of oocyte quality in determining ICSI success [2, 3, 4]. Oocyte morphology [2], maternal age [3], and the presence of ovarian stimulation techniques [4] have been reported as common assessment indicators. The present study focuses on smooth endoplasmic reticulum clusters (sERCs), which are considered a key factor in the success of ICSI. sERCs are vesicular structures observed in oocytes. Figure 1 shows an example of the sERC in an oocyte, where a region of the sERC is enclosed by a red dotted line. A study reported that oocytes with sERCs have lower fertilization and pregnancy rates, as well as delayed embryo development [5]. It also suggested that the size and distribution of sERCs may influence ICSI outcomes. However, another study found no significant correlation between the presence of sERCs and the success of ICSI [6]. In addition, this study reported no direct link between sERCs and chromosomal abnormalities in embryos. The observed inconsistency can be attributed to the inherent subjectivity of sERC evaluation, which is contingent upon the expertise of seasoned embryologists. Moreover, the heterogeneity of patient populations and the constraints of limited sample sizes are likely contributing factors to these disparate findings.

The application of artificial intelligence (AI) technologies is expected to contribute to solving this issue. In our previous work, we conducted evaluations of the oocytes and uterus in the context of infertility treatment using AI-based approaches [7, 8]. In the present study, we apply AI techniques to the analysis of sERCs. Segmentation technologies based on AI, particularly deep learning, are capable of accurately extracting complex morphological structures. These techniques offer a promising approach for

the objective evaluation of sERCs and oocyte regions. In this study, we aim to develop a reliable method to analyze the shape of sERCs in ICSI procedure videos. We use DeepLab [9], a deep learning-based segmentation method, to extract the regions of sERC, oocyte, and needle. We call the results from automatic segmentation "automatic extraction" and the results from manual annotation "manual extraction" to compare the two methods. By comparing automatic extraction with manual extraction, we evaluate the accuracy and reliability of our method.

Although intersection over union (IoU) is useful for measuring the overlap of regions, it does not show differences in size, shape, or angle. These differences are important for understanding how sERCs affect the success of ICSI. Since our future studies will investigate the relationship between sERC, injection methods, and pregnancy outcomes using these feature values, it is important that the automatically extracted features closely match those obtained manually. If the feature values are too different, they may affect the results of the future studies. Therefore, we compare these feature values to make sure the automatic extraction method is reliable in studies with large numbers of data.

Our main goal is to find an objective and repeatable way to measure these feature values. Traditional methods rely on subjective judgment, which can lead to mistakes. By using a numerical and automatic approach, we aim to make the measurement process more standardized. Efficiently processing large amounts of ICSI data should help us find patterns in sERC characteristics and how they relate to pregnancy outcomes. We need to make sure that the feature values we extract are reliable. So, in this study, we compare the feature values from automatic and manual extraction to calculate the error and check their reliability and consistency.

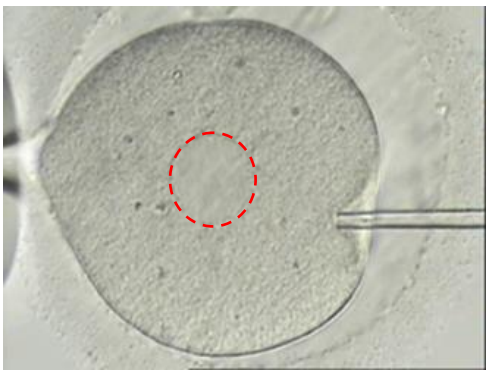


**FIGURE 1.** sERC in oocytes

## 2. Methods

In this study, we compare the feature values calculated from automatic and manual extraction results to evaluate the accuracy of the proposed automatic extraction method. In this section, we describe the dataset, the segmentation model, the feature values, and the evaluation methods.

### 2.1. Dataset

In this study, we used video data recorded at a Reproduction Clinic. The videos were recorded in WMV format (resolution: 480×640 pixels), and data from 11 cases were collected. The video lengths ranged from 38 to 89 seconds, with an average of 52.6 seconds. For dataset construction, one frame was extracted every second (24 frames) from each video and used as a still image. Table 1 shows the number of images per case.

Regions of sERC, oocyte, and needle were annotated manually using LabelMe [10], an open-source annotation tool. Figure 2 shows an example of an annotated image. The green region indicates sERC, the red region indicates the oocyte, and the yellow region indicates the needle. Annotation involved drawing polygons to appropriately identify each region and labeling them as "sERC," "oocyte," and "needle."

**TABLE 1.** Number of images for each case

| Case | Number of images |
|---|---|
| 1 | 49 |
| 2 | 50 |
| 3 | 65 |
| 4 | 45 |
| 5 | 43 |
| 6 | 49 |
| 7 | 89 |
| 8 | 38 |
| 9 | 48 |
| 10 | 50 |
| 11 | 53 |

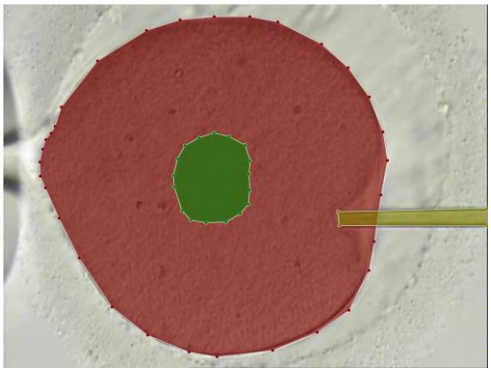


**FIGURE 2.** Annotated image

## 2.2. Segmentation Model

In this study, we used DeepLab, a segmentation method, to automatically extract regions of sERC, oocyte, and needle. For DeepLab training, the annotated dataset was divided using leave-one-out cross-validation, where one case was used as the test dataset, and the remaining cases were used for training. This process was repeated for all 11 cases, ensuring that each case was used once as the test set. Training hyperparameters included a batch size of 4, 25,000 steps, an optimization method of Adam, and a learning rate of 0.00001. We employed a pre-trained DeepLab model composed of ResNet-50 trained using the ImageNet published in reference [11].

To visualize the progress of the training, the loss values were recorded, and a learning curve was created. The learning curve shown in Figure 3 represents the average loss across all training iterations. The vertical axis indicates the loss relative to the training or test data, and the horizontal axis represents the number of steps. The blue line represents the average learning curve for the training data. The loss decreased rapidly in the early stages of training and tended to converge after approximately 10,000 steps. This result indicates that the number of training iterations was sufficient.

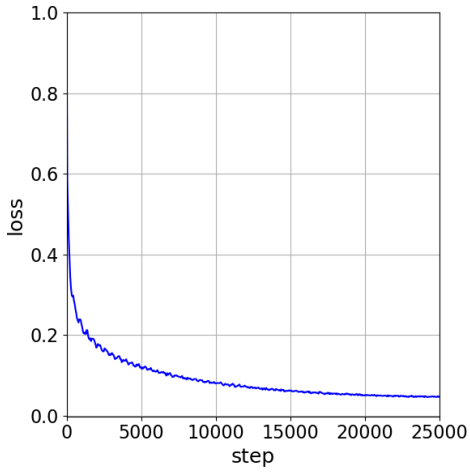



**FIGURE 3.** Learning curve

## 2.3. Feature Value Extraction

In this study, we employed four types of feature values—circularity of sERC, area ratio of sERC and oocyte, x-axis distance from center of oocyte to top of needle, and y-axis distance from center of oocyte to top of needle—to evaluate ICSI. These feature values were selected based on previous studies and materials [6, 12]. In this study, we denoted these feature values as $f_1$ to $f_4$, and they were calculated as follows.

$$f_1 = 4\pi \times S_s / L_s^2 \quad (1)$$

$$f_2 = S_s / S_o \quad (2)$$

$$f_3 = C_{ox} - T_{nx} \quad (3)$$

$$f_4 = C_{oy} - T_{ny} \quad (4)$$

Here, the notations $S_s$ and $S_o$ denote the areas of the sERC and the oocyte, the notation $L_s$ denotes the perimeter of the sERC, the notations $C_{ox}$ and $C_{oy}$ denote the x- and y-coordinates of the center of the oocyte, and $T_{nx}$ and $T_{ny}$ denote x- and y-coordinates of the top needle. These values were calculated from mask images showing the regions of sERC, oocyte, and needle.

## 2.4. Comparison with Automatic and Manual Extractions

Relative error of each feature value was used to compare manual and automatic extraction results. Relative error was calculated using the following equation.

$$\text{Relative Error} = (|A - M| / M) \times 100 \quad (5)$$

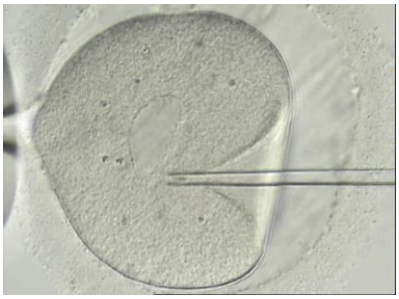
Here, notation $A$ denotes the automatically extracted value of the feature values, and notation $M$ denotes the manually extracted value.

## 3. Experiments

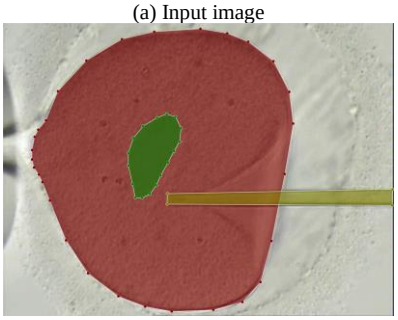
### 3.1. Segmentation Experiments

Figure 4 shows the results of Case 1 as an example of the prediction results. Figures 4(a), 4(b), and 4(c) show the input image, the ground truth labels (manual annotation data), and the output image, respectively. The red region indicates the sERC region, the green region indicates the oocyte region, and the yellow region indicates the needle region.

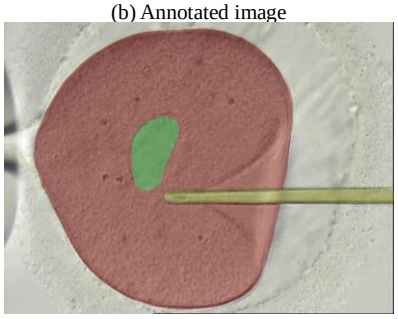
IoU is a measure of the overlap between predicted and ground-truth regions and was used to evaluate the performance of the segmentation model. Table 2 shows the average IoU of each case. The average IoU for the entire validation data was 0.93.



(a) Input image



(b) Annotated image



(c) Prediction image

**FIGURE 4.** Evaluation results

**TABLE 2.** Average IoU

| | Oocyte | sERC | Needle | Average |
|---|---|---|---|---|
| Case1 | 0.99 | 0.94 | 0.95 | 0.96 |
| Case2 | 0.99 | 0.97 | 0.80 | 0.92 |
| Case3 | 0.96 | 0.77 | 0.93 | 0.89 |
| Case4 | 0.99 | 0.93 | 0.94 | 0.95 |
| Case5 | 0.99 | 0.92 | 0.86 | 0.92 |
| Case6 | 0.99 | 0.94 | 0.90 | 0.94 |
| Case7 | 0.99 | 0.86 | 0.92 | 0.92 |
| Case8 | 0.95 | 0.88 | 0.91 | 0.91 |
| Case9 | 0.99 | 0.95 | 0.94 | 0.94 |
| Case10 | 0.99 | 0.78 | 0.96 | 0.91 |
| Case11 | 0.96 | 0.95 | 0.92 | 0.94 |
| Average | 0.98 | 0.90 | 0.91 | 0.93 |

### 3.2. Feature Extraction Experiments

To compare the automatic and manual extraction results, the calculated feature values were evaluated. While IoU is useful for assessing spatial overlap, it does not capture differences in shape, size, or other morphological characteristics. Therefore, comparing feature values allows us to evaluate not only the overlap but also the precision of the extracted regions' morphology, which is crucial for clinical analysis.

Table 3 shows the average and standard deviation of the relative errors for each feature value of automatic and manual extraction. These values are calculated for each case.

**TABLE 3.** Average and standard deviation of relative errors

| | $f_1$ [%] | | $f_2$ [%] | | $f_3$ [%] | | $f_4$ [%] | |
|---|---|---|---|---|---|---|---|---|
| | AVG | SD | AVG | SD | AVG | SD | AVG | SD |
| Case1 | 4.56 | 4.38 | 5.95 | 5.09 | 16.53 | 46.79 | 18.46 | 9.14 |
| Case2 | 4.29 | 5.48 | 7.42 | 5.20 | 3.17 | 4.58 | 10.82 | 3.76 |
| Case3 | 3.00 | 4.44 | 16.81 | 22.31 | 4.30 | 7.46 | 59.86 | 82.96 |
| Case4 | 7.18 | 9.00 | 29.22 | 39.43 | 4.87 | 13.08 | 15.60 | 8.60 |
| Case5 | 2.62 | 4.03 | 4.60 | 4.06 | 2.57 | 2.04 | 12.11 | 8.57 |
| Case6 | 3.63 | 7.04 | 4.13 | 2.88 | 5.29 | 9.59 | 19.17 | 16.43 |
| Case7 | 10.36 | 5.15 | 10.36 | 11.96 | 4.03 | 4.24 | 22.61 | 11.09 |
| Case8 | 8.21 | 8.30 | 10.36 | 6.66 | 12.17 | 19.12 | 71.59 | 49.97 |
| Case9 | 7.63 | 9.51 | 15.99 | 9.80 | 5.33 | 9.02 | 27.18 | 15.78 |
| Case10 | 5.14 | 6.10 | 10.28 | 14.58 | 5.33 | 7.72 | 26.73 | 23.55 |
| Case11 | 4.20 | 4.08 | 10.24 | 8.22 | 1.63 | 1.92 | 19.86 | 13.66 |
| AVG | 5.53 | 6.14 | 11.40 | 11.84 | 5.93 | 11.41 | 27.64 | 22.14 |

AVG: average, SD: standard deviation

## 4. Discussion

In this study, the regions of sERC, oocyte, and needle were extracted from the same image by both automatic and manual methods, and their characteristics were compared for application to future studies. As a result, the reproducibility of $f_1$ by automatic extraction was high and the relative error was small. On the other hand, the relative error for $f_2$ tended to be large in some cases. This is due to the fact that the value of $f_1$ was 0.6-0.8, while the value of $f_2$ was 0.025-0.2, so that even a small difference in the feature value would result in a large relative error.

The relative error for $f_3$ was small, but the error tended to be larger for $f_4$, and in one case the relative error exceeded 70%. Figure 5 shows an image of the region of the oocyte and needle in Case 8. The red dot represents the center of gravity of the oocyte and the black region enclosed by green line represents the area of the needle. From this image, it can be visually confirmed that the tip of the needle is extremely close to the center of the oocyte. In such cases, even a small extraction error can lead to a large deviation in the $f_4$ value.

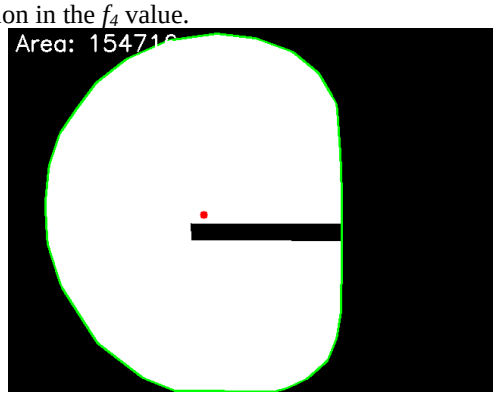
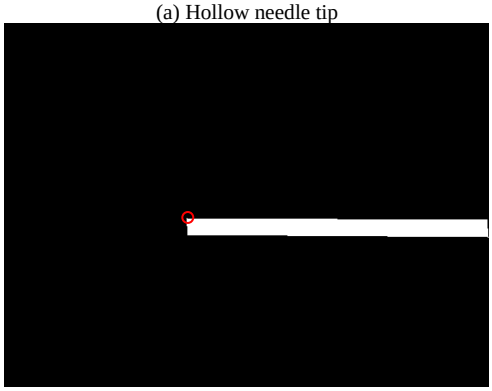



**FIGURE 5.** Manual extraction image of oocyte and needle

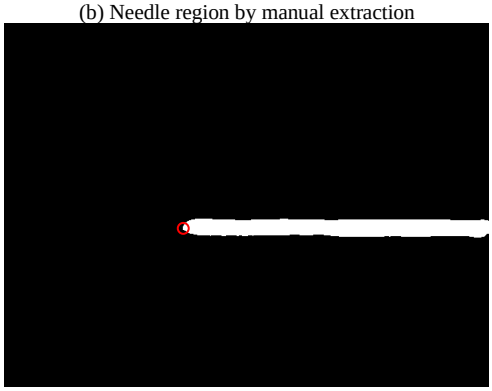
The boundaries between the oocyte and sERC and the background are relatively clear, making it easy to extract their centers and contours. However, the tip of the needle is hollow and, as shown in Figure 6(a), its boundary with the background is very difficult to distinguish. As a result, in automatic extraction, the tip may be recognized as rounded, leading to a deviation from the true tip position and increasing the $f_4$ error. Figures 6(b) and 6(c) show manually and automatically extracted images of the needle in Case 8. Figure 6(b) shows that the tip of the needle is clearly depicted. In contrast, Figure 6(b) shows that the tip of the needle has been smoothly rounded off. The red dots in the two images are the detected tips, and it can be seen that the rounding causes the $f_4$ value to deviate from the actual value. The discrepancy between the actual shape and the extraction result is the cause of the error.

Nevertheless, the IoU was high in all regions. The IoU was also high for the needle region, which is due to the fact that the extraction was very accurate for the areas other than the tip. In order to improve the accuracy in the future, it is especially important to detect needle tips more accurately. For example, it is expected that the tip identification accuracy can be improved by enhancing the contour of the needle in preprocessing.

(a) Hollow needle tip



(b) Needle region by manual extraction



(c) Needle region by automatic extraction

**FIGURE 6.** Needle Image and Corresponding Masks

## 5. Conclusion

This study demonstrated the usefulness of image analysis techniques for ICSI. DeepLab was used to isolate regions in the image and its accuracy was compared to manual extraction. The results showed that the average IoU of manual and automatic extraction was 0.93. Comparing manual and automatic extraction, there was little difference, proving that the DeepLab-based method is as accurate as manual.

Furthermore, by analyzing the feature values of the segmented regions, it was confirmed that the size and circularity of each region could be measured numerically. This enabled us to move from a subjective evaluation based on experience to a more reliable and reproducible method. Specifically, we quantitatively extracted four types of characteristic quantities: circularity of sERC, area ratio of sERC to oocyte, x-axis distance from the center of oocyte to the top of the needle, and y-axis distance from the center of oocyte to the top of the needle. This helps us understand how these shapes affect the success rate of ICSI.

In the future, it will be important to analyze a larger dataset to better understand how the shape and position of sERC and the needle affect the success rate of ICSI. Clarifying these relationships may lead to the development of new, objective evaluation criteria, which could help improve the accuracy and reliability of infertility treatments.

Furthermore, image analysis using deep learning has the potential to process large volumes of medical data efficiently. By applying this technology to other areas of medicine, it may become possible to support more accurate diagnoses and treatment planning. The findings of this study suggest that such methods can contribute to more effective and standardized medical care.

## References


[1] K. Hoshi, “ICSI (Intracytoplasmic Sperm Injection)”, Yamanashi Medical University Journal, Vol. 11, No. 4, pp. 65-71, 1996.

[2] R. Sciorio, and M. Meseguer, “Focus on time-lapse analysis: blastocyst collapse and morphometric assessment as new features of embryo viability”, Reproductive BioMedicine Online, Vol. 43, No. 5, pp. 821-832, Aug. 2021.

[3] A. La Marca, M. Capuzzo, M. Longo, M. G. Imbrogno, G. A. Spedicato, F. Fiorentino, F. Spinella, P. Greco, M. G. Minasi, and E. Greco, “The number and rate of euploid blastocysts in women undergoing IVF/ICSI cycles are strongly dependent on ovarian reserve and female age”, Human Reproduction, Vol. 37, No. 10, pp. 2392-2401, Sep. 2022.

[4] Nishitan ART, “Ovarian Stimulation (Ovulation Induction)”, https://nishitan-art.jp/treatment/art/ovarian-stimulation/, accessed Apr. 1, 2025.

[5] A. S. Setti, R. C. S. Figueira, D. P. A. F. Braga, M. C. Azevedo, A. Iaconelli Jr., and E. Borges Jr., “Oocytes with Smooth Endoplasmic Reticulum Clusters Originate Blastocysts with Impaired Implantation Potential”, Fertility and Sterility, Vol. 106, No. 7, pp. 1718-1724, Oct. 2016.

[6] C. S. Chiu, T. Hung, M. Lin, R. K. Lee, Y. Weng, and Y. Hwu, “Metaphase II (MII) Human Oocytes with Smooth Endoplasmic Reticulum Clusters Do Not Affect Blastocyst Euploid Rate”, Taiwanese Journal of Obstetrics and Gynecology, Vol. 61, No. 4, pp. 585-589, Jul. 2022.

[7] N. Yagi, H. Tsuji, T. Morimoto, T. Maekawa, S. Mizuta, T. Ishikawa, and Y. Hata, “Rupture Prediction for Microscopic Oocyte Images of Piezo Intracytoplasmic Sperm Injection by Principal Component Analysis,” Journal of Clinical Medicine, Vol. 11, No. 21, 2022.

[8] K. Mori, K. Kitaya, T. Ishikawa, and Y. Hata, “A Pregnancy Prediction System based on Uterine Peristalsis from Ultrasonic Images”, Intelligent Automation & Soft Computing, Vol. 29, No.2, pp.335-352, Jun. 2021.

[9] B. Cheng, M. D. Collins, Y. Zhu, T. Liu, T. S. Huang, H. Adam, and Liang-Chieh Chen, “Panoptic-DeepLab”, Proceedings of ICCV 2019 Joint COCO and Mapillary Recognition Challenge Workshop, 2019.

[10] K. Wada, “Labelme - AI Image Annotation & Dataset Creation”, https://labelme.io/, accessed Apr. 1, 2025.

[11] Q. Yu, and Y. Zhu, “Download the pretrained checkpoints”, https://github.com/google-research/deeplab2/blob/main/g3doc/projects/imagenet_pretrained_checkpoints.md, accessed Apr. 1, 2025.

[12] H. Hattori, Y. Nakamura, Y. Nakajo, Y. Araki, and K. Kyono, “Deliveries of babies with normal health derived from oocytes with smooth endoplasmic reticulum clusters”, Assisted Reproduction Technologies, Vol. 31, pp. 1461-1467, Sep. 2014.