**St. Luke's International University, Graduate School**

**Doctoral Dissertation, 2016**

# Oxytocin Levels in Low-Risk Primiparas Following Breast Stimulation for Spontaneous Onset of Labor

**Kaori Takahata**

## Table Contents (Main Text)

**Table of Contents (Tables and Figures)**

# Introduction

## Background

Generally, a 40-week gestation is estimated as the date of delivery. The lowest numbers of adverse neonatal outcomes occur between 39 weeks and 40 weeks and 6 days of gestation (Reddy et al., 2011; Tita et al., 2009). Prolonged pregnancy is called late - term at 41 weeks through 41 weeks and 6 days and postterm at 42 weeks and beyond (ACOG committee opinion no 579: Definition of term pregnancy.2013; Spong, 2013). Importantly, the continuation of pregnancy or post-term pregnancy is associated with increased maternal and fetal risks (Olesen, Westergaard, & Olsen, 2003). Synthetic oxytocin infusion for labor induction is used to prevent those increased risks. However, labor induction is one of the factors of a negative birth experience (Henderson & Redshaw, 2013). In addition, there is concern that labor induction might have long-lasting negative effects for the offspring (Freedman, Brown, Shen, & Schaefer, 2015; Gregory, Anthopolos, Osgood, Grotegut, & Miranda, 2013; Kurth & Haussmann, 2011).

While oxytocin infusion for labor induction can have positive effects, there might also be some negative aspects. Therefore, many pregnant women make efforts to induce labor before synthetic oxytocin infusion is required (Chaudhry, Fischer, & Schaffir, 2011; Declercq, Sakala, Corry, Applebaum, & Herrlich, 2014; Kozhimannil, Johnson, Attanasio, Gjerdingen, & McGovern, 2013; Schaffir, 2002). Currently, induction of labor is expected during either membrane sweeping or breast stimulation (Boulvain, Stan, & Irion, 2005; Kavanagh, Kelly, & Thomas, 2005). Membrane sweeping is a medical procedure and it may even be uncomfortable and cause pain. On the other hand, breast stimulation is a natural method that requires no cost or technology, and can be taught by midwives and performed by pregnant women at their own choosing and convenience. Moreover, breast stimulation has been utilized for centuries in obstetrics; medical texts from the 18th century in Europe wrote about breast stimulation as an important method for managing labor (Curtis, Resnick, & Evens, 1986; Curtis, 1999). In Japan, of the 530 puerperants, 93 (17.5%) made an effort to do breast stimulation for induction of labor (Takahata, 2015). Researchers found that breast

stimulation for labor induction reduced the number of women who were not in labor after 72 hours (62.7% versus 93.6%, relative risk 0.67, 95% confidence interval 0.60 to 0.74) (Kavanagh et al. 2005). The minimum breast stimulation time included in the review was one hour per day for three days (Adewole, Franklin, & Matiluko, 1993). However, researchers have yet to establish an effective intervention time for breast stimulation. It is necessary to verify the feasibility and whether or not adverse events might occur for Japanese women.

There are various theories regarding the mechanisms of the onset of labor, which appear to be associated with a variety of different factors (Buckley, 2015; Cunningham et al., 2014). Among them, it is considered to be associated with oxytocin levels between the breast simulation and onset of labor (Kavanagh, Kelly, & Thomas, 2005; Razgaitis & Lyvers, 2010). Breast stimulation promotes the secretion of oxytocin from the posterior pituitary. Plasma oxytocin binds to the oxytocin receptor of the uterine muscle, resulting in an increase in phospholipase C activity and an increase in cytoplasmic calcium levels. Increased calcium levels work to stimulate the contraction of the myometrial cells, and leads to organized uterine contractions (Buckley, 2015; Fuchs, Fuchs, Husslein, Soloff, & Fernstrom, 1982; Kamel, 2010). A few studies have measured the oxytocin level during short breast stimulation for pregnant women at various minute intervals. The oxytocin level as induced by breast stimulation for pregnant women was reportedly non-significant (Leake, Fisher, Ross, & Buster, 1984; Ross, Ervin, & Leake, 1986) and significant increased (Amico & Finley, 1986; Christensson, Nilsson, Stock, Matthiesen, & Uvnas-Moberg, 1989; Finley, Silverman, & Lannin, 1989).

Those results were not consistent and the oxytocin levels with breast stimulation. Moreover there is a gap in knowledge related to repeat and long breast stimulation, oxytocin levels and spontaneous onset of labor. A preliminary study was conducted before this study (Takahata et al., 2016). The analysis using plasma was affected by the pulsing of oxytocin, moreover it was highly invasive to the participants, therefore, measurement with saliva was chosen. It was suggested that the possibility of oxytocin increases with breast stimulation in third intervention day. It was an acceptable

method for pregnant women in Japan, and this previous research confirmed the feasibility of protocols. However, the limitations of this preliminary study were the lack of a control group and half of the collected samples had insufficient amount of saliva for analysis of the oxytocin level. Accordingly, this study focused on the oxytocin levels to investigate the physiological background for long hours of breast stimulation with control group.

**Objectives**

This study aimed to examine the effect on the salivary oxytocin levels in low-risk primiparas following breast stimulation for spontaneous onset of labor and the feasibility of the breast stimulation intervention protocol.

**Study Outcomes**

Primary Outcome: salivary oxytocin level

1. The intervention group would have a higher oxytocin level at 30 minutes on the third day than the control group.
2. The intervention group would have a higher basal oxytocin level before the intervention on the third day compared to the first day.

Secondary outcomes:

1. The intervention group would have a higher rate of spontaneous onset of labor than the control group.
2. The intervention group would have a higher rate of *adequate uterine contraction* than the control group.
3. The oxytocin level would be related to background factors and delivery outcomes.

Feasibility:

1. The majority of participants would confirm the acceptability and practicality of the protocol thus confirming its feasibility.

## Significance of the Study

Practicing breast stimulation for spontaneous onset of labor could result in a significant reduction in the number of women not in labor after 72 hours. Moreover, to investigate the salivary oxytocin levels during breast stimulation may make a contribution to physiological evidence for a method of labor induction that depends on the pregnant woman's own initiative.

In the future, if breast stimulation were effective for inducing labor an evidenced-based breast stimulation program would be a viable addition to complementary and alternative medicine to reduce the medical intervention for onset of labor.

Oxytocin is attracting a great deal of attention from various areas of study. This study might not only benefit pregnant women, as the assay of oxytocin from saliva is also noninvasive but could also be easily used for other situations and populations. This study would contribute to the development of a method to assay salivary oxytocin.

## Definition of Terms

1. Oxytocin level: This is oxytocin concentration in plasma or saliva.
2. Low-risk pregnant women: Pregnant women who will have a singleton birth by spontaneous cephalic delivery, with a normal progression of pregnancy and without medical or pregnancy complications.
3. Breast stimulation: This procedure is gently massaging own nipple.
4. Spontaneous onset of labor: The natural initiation of labor leading to delivery without medical or instrumental intervention.
5. Induction of labor: The initiation of labor using artificial methods with the aim of achieving vaginal delivery.
6. Adequate uterine contractions: At least three contractions lasting for 40 seconds each occurring within 10 minutes are *adequate contractions*.

# Literature Review

## Onset of Labor

Labor indicates uterine contraction that eventually leads to parturition. The onset of labor refers to a cycle of uterine contractions occurring in less than ten minutes, or at a frequency of six or more cycles per hour. Labor that does not lead to parturition is called false labor or Blackstone Hicks contraction.

The exact mechanism of labor onset in humans remains unclear, and diverse hypotheses have been suggested. Species difference is reported to be large, making accurate elucidation more difficult (R. Smith, 2007). In particular, parturition is classified into four stages: *Phase 0*, Quiescence; *Phase 1*, Activation; *Phase 2*, Stimulation; *Phase 3*, Involution (Cunningham et al., 2014). The appearance of labor onset indicates changes that occur from *Phase 1* to the early stage of *Phase 2*.

The mechanisms of labor onset are reported to be related to the corticotrophin releasing hormone hypothesis, prostaglandins, oxytocin, relaxin, uterine stretch, and surfactant protein-A. However, labor is a phenomenon that occurs as a result of temporal and spatial complications leading to the disappearance of factors related to the maintenance of pregnancy and the appearance of factors that induce labor (Cunningham et al., 2014).

## Induction of Labor

### Prolonged pregnancy.

Generally, researchers have found that 10% of pregnant women continue pregnancy past 42 weeks gestation (Olesen 2003). The most common cause of prolonged pregnancies attributed to inaccurate dating (Whitworth, Bricker, & Mullan, 2015). Caughey, Nicholson, & Washington (2008) investigated the gestation period in 7894 women. When women received ultrasound before 12 weeks of their estimated delivery date, it was wrong only 2.7% of the time compared to 3.7% ($p = 0.02$) when evaluated by ultrasound between 12 and 24 weeks. Prolonged pregnancy has been associated with primiparas (Campbell, Ostbye, & Irgens, 1997), their own mothers

having had postterm pregnancy (Mogren, Stenlund, & Hogberg, 1999), obesity (Stotland, Washington, & Caughey, 2007) , male fetus (Divon, Ferber, Nisell, & Westgren, 2002) and duration of pregnancy associated with race (Patel, Steer, Doyle, Little, & Elliott, 2004).

**Medical induction of labor.**

The rate of induction of labor is rising worldwide (ACOG Committee on Practice Bulletins -- Obstetrics, 2009). A national U.S. online survey in 2011 of 2,400 mothers found that 30 percent reported medically induced labor (Declercq et al., 2014). Medical induction also affects the psychological state of women. For low risk pregnant women, because they do not anticipate medical intervention, they exhibited a strong fear against induction of labor (Gammie & Key, 2014). Unexpected medical intervention such as childbirth induction is considered to be one of the factors for expressing a negative birth experience (Shetty, Burt, Rice, & Templeton, 2005; Waldenstrom, Hildingsson, Rubertsson, & Radestad, 2004).

In the use of oxytocin infusion and using epidural anesthesia, it has been found to decrease the amount of postpartum oxytocin levels (Handlin et al., 2009).

**Complementary alternative medicine.**

Some complementary alternative medicine options for spontaneous onset of labor have been studied. For example there are studies of using coitus (Kavanagh, Kelly Anthony, & Thomas, 2001; Omar, 2013), acupuncture (C. A. Smith, Crowther, & Grant, 2013), homeopathy (C. A. Smith, 2003), castor oil (Kelly Anthony, Kavanagh, & Thomas, 2013), but there were no significant effects or insufficient evidence.

**Oxytocin**

Oxytocin not only affects a wide range of human physiological processes including the development of the brain and peripheral organs but also relates to learning, attachment, and social cognition and contributes to maintaining normal conditions in humans (Lee, Macbeth, Pagani, & Young, 2009; Yang, Wang, Han, &

Wang, 2013). In the 1900s, oxytocin was described to cause contractions of the mammalian uterus (Dale, 1906). In the 1950s, it was identified as being composed of nine amino acid peptide hormones (du Vigneaud, Ressler, & Trippett, 1953), followed shortly by its synthesis (du Vigneaud, Ressler, Swan, Roberts, & Katsoyannis, 1954). Synthetic oxytocin is commonly used in the clinical setting for the induction and augmentation of labor (ACOG Committee on Practice Bulletins -- Obstetrics, 2009).

Oxytocin level and the expression of oxytocin receptors were reported to be higher in pregnant women (Carter et al., 2007; Zingg & Laporte, 2003). In late pregnancy, the number of oxytocin receptors in the uterine muscle layer and decidual tissue increases markedly (Fuchs, Fuchs, Husslein, & Soloff, 1984; Kimura et al., 1996; Wathes, Borwick, Timmons, Leung, & Thornton, 1999). Oxytocin, which is secreted into the plasma from the posterior pituitary gland, binds to oxytocin receptors in the uterine muscle, causing a sequential increase in phospholipase C activity, cytoplasmic calcium level, and uterine contraction (Buckley, 2015; Fuchs et al., 1982). Oxytocin promotes the release of prostaglandin from the chorion, decidua, and amniotic membrane, and was reported to be synthesized directly from the deciduous membrane, fetus, and placenta (Chibbar, Miller, & Mitchell, 1993; Zingg et al., 1995). Moreover, excessive oxytocin secretion can cause abnormal uterine contraction and preterm birth (Arthur, Taggart, & Mitchell, 2007; Terzidou, Blanks, Kim, Thornton, & Bennett, 2011). Based on these reports, oxytocin is thus considered to be a hormone involved in labor.

Measurement with saliva is a noninvasive way to measure oxytocin. With regard to the correlation between salivary oxytocin level and plasma oxytocin level, a moderate correlation ($r = 0.41$-$0.59$) has been confirmed (Feldman, Gordon, Schneiderman, Weisman, & Zagoory-Sharon, 2010; Grewen, Davenport, & Light, 2010). Notably, a high correlation ($r = 0.89$) has also been reported by other studies in women with anorexia nervosa without self-induced vomiting (Hoffman, Brownley, Hamer, & Bulik, 2012). Although there have been studies on salivary oxytocin in relation to breast stimulation in postpartum women (Grewen et al., 2010), no studies have apparently investigated the effects of breast stimulation on the changes in oxytocin level during pregnancy.

## Oxytocin Receptor

Research on oxytocin receptor synthesis and regulation of the myometrium has been carried out in rodents. Disruption of the oxytocin receptor genes in mice has been found to have no effect on parturition; also in case of lack of milk ejection (Nishimori et al., 1996). The human oxytocin receptor is a polypeptide consisting of 389 amino acids. The oxytocin receptors gene is located on the short arm of chromosome three (3q25) and has three introns and four exons (Israel et al., 2008). Oxytocin receptors are present in the uterine muscle and nipple (Kimura, 1995). In particular, the number of oxytocin receptors in the uterine muscle increases 50 times or more in late pregnancy, and it is reported to increase up to 200-300 times by delivery (Fuchs et al., 1982; Kimura et al., 1996). The increase in the number of oxytocin receptors is reportedly correlated with an increase in the uterine contraction response to the oxytocin stimulus (Fuchs, Periyasamy, Alexandrova, & Soloff, 1983). Therefore, even if the same oxytocin stimulus is given, the uterine contraction response varies depending on the individual and gestational weeks. It has been pointed out that a prolonged pregnancy is related to a delayed increase in the number of oxytocin receptors (Fuchs & Fuchs, 1984).

Recently, it has been reported that oxytocin sensitivity is decreased by oxytocin receptor gene single nucleotide polymorphism. It has been confirmed that in the GG type of rs53576, the duration of labor at the first stage is prolonged (Terkawi et al., 2012). In addition, it has been found that the GG type of rs2254298 is related to a significantly low plasma oxytocin level (Feldman et al., 2012).

## Breast Stimulation

### Breastfeeding.

The structure of the breast consists of the skin and subcutaneous tissue, fat, breast parenchyma (ducts and lobules), nipple, and areola. The breast is a modified cutaneous exocrine gland. In the nipple and areola, free nerve endings are present, and the nerves distributed in the skin travel radially from the mammary gland body toward the nipple. The mammary gland consists of dendritically divergent acini, and each

acinus is covered with a capillary network and smooth muscle-like epithelial cells (Jesinger, 2014).

The best-known response to oxytocin is milk ejection in puerperium. Stimulation such as suckling of a baby stimulates the hypothalamic neurons through the spinal cord thalamic tract, increases the plasma oxytocin level, and causes milk discharge by contraction of the myoepithelial cells surrounding the acini (Crowley & Armstrong, 1992; McNeilly, Robinson, Houston, & Howie, 1983). After suckling, the plasma oxytocin level increases within one minute (Leake, Waters, Rubin, Buster, & Fisher, 1983) and returns to the basal value within six minutes after suckling (Riordan & Wambach, 2010). This oxytocin response to suckling was reportedly similar to the oxytocin response in late postpartum as long as breastfeeding continues (Leake et al., 1983). A previous study confirming the milk ejection phenomenon noninvasively by ultrasound imaging of the breast in lactating women has shown that the actual milk discharge occurs between 45 seconds and 3.5 minutes (Ramsay, Kent, Owens, & Hartmann, 2004).

Another effect of the increased oxytocin level due to lactation is the promotion of uterine involution via uterine contractions. The uterus is expected to contract rhythmically for 20 minutes not only during lactation but also after lactation (Riordan & Wambach, 2010).

**Uterine construction.**

Breast stimulation during pregnancy has been mainly studied in the 1980s for the contraction stress test (CST) as an alternative to the use of synthetic oxytocin (Chapman, Pincombe, & Harris, 2013; Marshall, 1986). CST is one of the antenatal fetal assessment techniques for evaluating the well-being of the fetus by mimicking labor during childbirth. CST is performed by stimulating the nipple and areola for a short time using clothes or towels until sufficient uterine contraction is confirmed. In these studies breast stimulation was performed with the participant in the left lateral or semi- Fowler's position (Marshall, 1986). In these studies, when *adequate uterine contractions* were not obtained with only one-sided stimulation, bilateral stimulation

was performed or additional long stimulation was added. Thus, the rate of *adequate uterine contractions* was higher than 80%. The rate of hyperstimulation in the study was 5-21% for high-risk pregnancy with prolonged and bilateral nipple stimulation.

However, CST is rarely performed at present. The current mainstream techniques used for assessing fetal reserve performance are the non-stress test, in which the fetal heart rate (FHR) is monitored without artificial stimulation, and image diagnosis.

Results of previous reports have demonstrated that the plasma oxytocin level shows a small change even if the level increases or no change when induced by breast stimulation during pregnancy (**Table 1**). Moreover, there is apparently no research regarding changes in breast stimulation-induced oxytocin levels in reference to a control group. Studies investigating the effects of breast stimulation for several days on changes in oxytocin levels are also lacking.

**Spontaneous onset of labor.**

Thus far, there are six randomized and quasi-randomized controlled studies of breast stimulation for the spontaneous onset of labor (**Table 2**). What was common among these studies was the stimulation of the nipple and areola for long hours by self-massage at home. The studies found that breast stimulation for labor induction reduced the number of women not in labor after 72 hours (62.7% versus 93.6%, relative risk 0.67, 95% confidence interval 0.60 to 0.74) (Kavanagh et al. 2005). In addition, Adewole et al. (1993) reported that postpartum hemorrhage was significantly reduced in multiparas. Kavanagh et al. (2005) reports that the rate of postpartum hemorrhage was also reduced (0.7% versus 6.0%, relative risk 0.16, 95% confidence interval 0.03 to 0.87).

In some studies, the instructions for breast stimulation were shown to the participants in a movie for two minutes (Singh, Tripathi, Mala, & Yedla, 2014), and breast stimulation was performed at the hospital for the first time (Salmon, Kee, Tan, & Jen, 1986). In most studies, the stimulation methods were not strictly controlled. Studies using lubricants were considered as direct stimulation (Elliott & Flaherty, 1984; Singh

*Table 1*

***Literature Review of Reports on Plasma Oxytocin Level in Relation to Short Breast Stimulation in Pregnant Women***

| **Reference** | **Sample** | **Risk of pregnancy** | **Parity** | **Gestational weeks** | **Intervention** | **Time** (minutes) | **Plasma sampling** | **Results** Oxytocin/ adequate UC[a] | **Adverse event** |
|---|---|---|---|---|---|---|---|---|---|
| **Christensson et al. (1989)** | 10 | Low | Primiparas and multiparas | 38-39 | Self/ Only unilateral nipple | 30 | Every 15 seconds at uterine contraction (total 6 contractions per women) | Significant increased in the sixth contraction compared with the base/ 90.0% adequate UC | 1 uterine tachysystole with FHR deceleration |
| **Ross et al. (1986)** | 20 | High | Primiparas and multiparas | Unclear | Self/ alternating stimulation for 10 minutes | Max 40 | Pre (samples 1, 2, 3), during (4, 5), post (6) | *ns* / 45.0% adequate UC and FHR was monitored | Unclear |
| **Amico et al. (1986)** | 19 | Unclear | Unclear | 28-40 | Self or electrical breast pump/ Huddleston methods[b] | Max 12 | Pre and post stimulation | Significant increased/ UC was not reported. | Unclear |
| **Finley et al. (1989)** | 19 | High and low | Primiparas and multiparas | 28-40 | Self or electrical breast pump/ Huddleston methods[b] | Max 12 | Pre and every 45 second after stimulation (Max 6 samples per woman ) | Significant increased/ 52.6% adequate UC and FHR was monitored | Unclear |
| **Leake et al. (1984)** | 10 | Low | Unclear | 18-20 (4 women) or 33-44 (6 women) | Elctrical breast pump/ alternating stimulation | 12 | 7 samples every 3 minutes | *ns* / FHR was monitored, but UC was not reported. | Unclear |

*Note.* All studies involved radioimmunoassay for determining plasma oxytocin concentration. FHR = fetal heart rate; *ns* = not significant; UC = uterine contraction.

[a]At least 3-5 contractions lasting for 30-45 seconds each occurring within 10 minutes. [b]Alternating stimulation for 2 minutes and rest for 5 minutes by Huddleston, J. F., Sutliff, G., & Robinson, D. (1984). Contraction stress test by intermittent nipple stimulation. *Obstetrics and Gynecology, 63* (5), 669-673.

**Table 2**

***Literature Review of Reports on Breast Stimulation to Induce Onset of Labor in Pregnant Women***

| Author | Setting | Sample | Risk of pregnancy | Parity | Gestational weeks[a] | Intervention | Period | Duration time × times per day | Adverse events[b] |
|---|---|---|---|---|---|---|---|---|---|
| **Adewole et al[c] (1993)** | Nigeria | 100 | Low | Primiparas | 39 | Self simulation or avoid stimulation | 3 days | 60 minutes × 1 | Nothing |
| **Adewole et al[c] (1993)** | Nigeria | 50 | Low | Multiparas | 39 | Self simulation or avoid stimulation | 3 days | 60 minutes × 1 | Nothing |
| **Salmon et al (1986)** | Singapore | 50 | Low | Primiparas | 38 | Self simulation or avoid stimulation | 3 days | More than 180 minutes × 1 | Nothing |
| **Damania et al. (1992)** | India | 57 | High | Primiparas | 37 | Self simulation or oxytocin infusion or no stimulation | 3 days | 60 minutes × 3 | 3 fetal deaths |
| **Singh et al. (2014)** | India | 199 | Low | Primiparas | 38 | Self simulation or no stimulation | 7 days | 30 - 40 minutes × 3 | 1 intrapartum hyperstimulation |
| **Elliot et al. (1984)** | U.S. | 200 | Low | Primiparas and multiparas | 39 | Self simulation or avoid stimulation | Until derivery | 60 minutes × encouraged over 3 times | Nothing |
| **Chayen et al. (1993)** | U.S. | 62 | High | Primiparas and multiparas | 37 | Elctrical. breast pump or oxytocin infusion | Until every 3 minutes UC[d] | - | Nothing |

Note. Sample size was not calculated in all the studies. [a]This indicate that has completed the gestation weeks. [b]These include the events during experiment and delivery. [c]These was summarized in one report. [d]Participants in this study were pregnant women who required induction of labor, and when breast stimulation failed, they were induced by oxytocin infusion.

et al., 2014). Whether or not the lubricant was used was determined by individual judgment.

There were two studies involving high-risk pregnancy among the participants. The classification of high-risk pregnancy was based on the following: over the due date, post-term, pregnancy-induced hypertension, pre-eclampsia, and fetal growth restriction among others (Chayen, Tejani, & Verma, 1986; Damania et al., 1992). Damania et al. (1992) compared three groups of 57 high-risk women in India. There were three deaths involving the fetus in the breast stimulation group and one death involving a woman in the oxytocin infusion group. It was not clear when this adverse event occurred. In other studies, there were no fetal deaths reported, and other adverse events were rarely reported (see **Table 2)**. Uterine hyperstimulation as an adverse event during the experiment was not reported in those studies. However, there was only two studies (Salmon et al., 1986; Chayen et al., 1986), which confirmed FHR monitoring.

**Summary of Literature Review**

Although various factors have been shown to be related to the onset of labor, oxytocin and its receptors are considered to play an important role in this process. Despite the studies elucidating the relation between breast stimulation and uterine contraction using oxytocin as an indicator, the results are not consistent. Furthermore, the effect of long hours of breast stimulation, which is reported to be effective for labor onset, remains unclear.

## Methods

### Study Design

A quasi-experimental time series design with a control group was conducted to investigate the salivary oxytocin levels during breast stimulation for spontaneous onset of labor.

The first half of participants were assigned to the control group as usual care and the other half were assigned to the intervention group as breast stimulation for three days for one hour per day. Both groups visited the hospital for saliva collection on the 1st and 3rd intervention day (**Figure 1**).

### Setting

This study was conducted at one hospital with more than 2,500 births per year and located in an urban area of Kanagawa, Japan. The caesarean section rate has been less than 4%, and it is a hospital that supports vaginal deliveries as much as possible. In addition, this hospital sometimes advises using breast stimulation for pregnant women. It was also the site of the preliminary study conducted from June to August 2015.

### Participants

Eligible participants were recruited who met all the following at 34 weeks of gestation. They were: (1) between 20 and 39 years of age, (2) planning to have a singleton birth by spontaneous cephalic delivery, (3) between 38 and 40 weeks of gestation, (4) Asian and could read and write Japanese, and (5) able to provide permission to participate in the present research that was provided by the attending physician or midwife. Women were excluded from the study if they had: (1) ongoing medication related to their gestation, (2) medical or pregnancy complications, (3) mental illness, (4) a medical history of assisted reproductive technology treatment, (5) a BMI above 25 before pregnancy, (6) planned on an induced labor, or (7) already conducted breast stimulation over ten minutes per day.

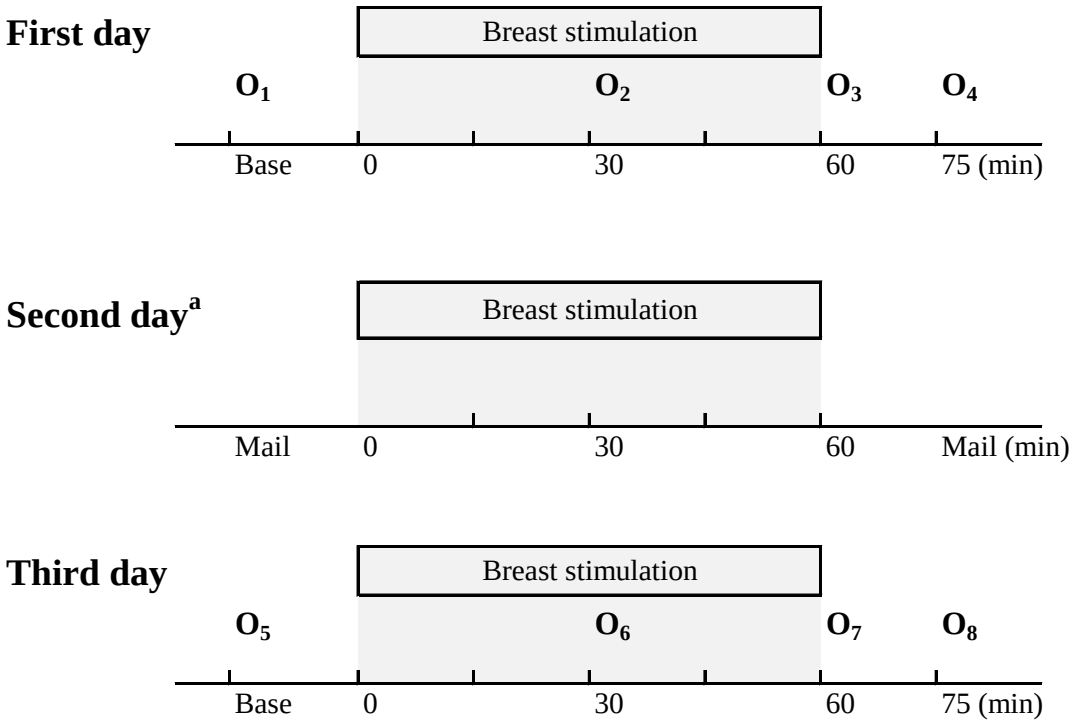



*Figure 1.* Breast stimulation protocol for three days in the intervention group.
[a]Second day protocol was by participant at home.

**Research Period**

Data were collected from February to September 2016 in a single maternity hospital in Kanagawa, Japan.

**Sample Size**

A sample size was determined based on previous research from White-Traut et al. (2009) . In that study the minimal sample size of 11 breastfeeding women indicated significance before, during and after differences in salivary oxytocin during the breastfeeding cycle (White-Traut et al., 2009). Moreover, sample sizes for reports on plasma oxytocin level in relation to short breast stimulation in pregnant women (Table 1) were 10-20 people. The sample size of this research warranted 15 women in each group. In the preliminary research, the assay sample in first intervention day was 40-63% and in third day was 27-37%, because of the insufficient amount of saliva collected for oxytocin level measurement. Considering the dropout rate to be 30% from lack of saliva for assay from the improved saliva collection protocol. The sample size needed was 22 women in each group.

**Procedure for Conducting the Study**

**Recruitment.**

The researcher invited participation using <Appendix 1> and < Appendix 2> with the cooperating institution from the preliminary study. After consent from the institution, the researcher posted the disclosure of research information at the outpatient wards before beginning recruitment <Appendix 3>.

When eligible pregnant women at 34 weeks gestation visited the hospital for a health checkup, the researcher or research assistant provided a verbal explanation about the research objectives and methods using: <Appendix 4 or 5> and < Appendix 6 or 7>. When the researcher or research assistant gained the woman's agreement the *written informed consent to participate in the study form* was signed and dated <Appendix 8> and the original was kept by the researcher, and the consent form (copy) was given to the participant. Finally, the research team and participant together decided her

intervention day.

**Data collection.**

The intervention period was three consecutive days. It occurred in the hospital within the examination time of the outpatient visit, in case of emergencies. Women attended the hospital on first and third intervention day. The intervention was done in private room at the hospital and attendance with the midwife was provided to ensure women's privacy. No one else was permitted in the intervention room during the intervention.

To begin, the participant urinated and then gargled with plain water. Then the researcher explained to the women how to properly collect their saliva. Participants sat in a semi- Fowler's position in a folding chair and they could freely change their posture. Then, a cardiotocograph belt (Fetal Actocardiograph MT-610; TOITU, Tokyo, Japan) was applied to the participants abdomen record both the fetal heart rate and the uterine activity. Next participants started to watch a movie in silence (*See The World by Train* (Sekai no Shasho kara) - Austria Railway, JAN 4988021151528; VAP, Tokyo, Japan) 10 minutes prior to saliva collection for the baseline data. This movie was to control external factors during the intervention in order to minimize environmental aspects that might affect the experimental environment.

After the intervention, the participants answered a post-test < Appendix 9 and 10 or 11 and 12>. If pre-labor contractions happened, the midwife consulted with the participant about whether go home or stay in the hospital. If the researcher was unable to judge it was onset of labor she reported to the head nurse to seek direction.

In a previous research (Takahata, 2015) found that of the 530 puerperants, 296 (55.8%) did breast stimulation before parturition. Moreover, 93 (31.4%) women made an effort to also do it for induction of labor. In the results, stimulation time was 7.5 minutes (*SD* 8.7) per day, and 61% women did it less than 5 minutes and 76 % women reported doing it once a day. From the above, approximately 18% of Japanese women practiced breast stimulation for inducing labor. Moreover, the majority of women in the study practiced breast stimulation for a short period of to prepare for breastfeeding. Due

to the above reasons, researcher decided that breast stimulation for less than 10-minutes per day belong under usual care category.

After all interventions were finished, the researcher requested participants refrain from breast stimulation over 10-minutes per day and when if they did it, the implementation status was written in <Appendix 13>. After giving birth, <Appendix 13> was collected when researcher visited, or participants put it in the prepared envelope and into the collection box in the outpatient wards or posted it. After participating in the study, participants received a fee by bank transfer (¥2,000 or about $20 per day) based on the number of days participating in the experiment.

**Intervention.**

***Breast stimulation procedure.***

Each breast was stimulated for 15 minutes per side, beginning on the right, and alternating for a total of 1 hour per day, because, stimulation at the same time on each side might lead to uterine hyperstimulation (Adewole et al., 1993; Salmon et al., 1986). To assure timing consistency the researcher monitored the time and notified the participant when it was time to change sides or time to stop. The nipple was stimulated directly by the participant using either the left or right hand with pure lanolin nipple cream (PureLan 7g; Medela, Baar, Switzerland). This cream was distributed to each participant and the amount could be freely used. For example using the right hand, stimulation was performed by pinching only the nipple using the thumb in the direction of the “12 o'clock or 9 o'clock” position with nipple and forefinger moving in opposite directions. Women were requested to wear clothes that would allow easy access to their nipples. Women used a prepared nursing cover during the intervention.

***First day.***

After the first saliva collection, women received an explanation for the method of breast stimulation. Attendant researcher demonstrated for about 5 minutes to standardize the breast stimulation. The follow were used to ensure a standardized practice: first, to standardize a stimulation rhythm for 69 beats per minute using an

electric metronome (SQ50V, Seiko Watch Corporation, Tokyo, Japan) with lighting instead of sound, second, explained about positioning of arm or fingers for breast stimulation, using own breast and arm of attendant researcher, third, to indicate the position of finger and nipple showing stimulation method on a breast model made of wool, fourth, to indicate the degree of stimulation pressure by providing visibility using a pressure-measuring instrument (Perineometer, OWOMED, Gyeonggido, South Korea), (recommended stimulation pressure was less than 10 mmHg, which would be gentle and not damage the nipple), and fifth, the attendant researcher stimulated actual women's breast for about five counts according to the rhythm.

The effect of tactile stimulation depends on various factors such as strength and speed (Ellingsen, Leknes, Loseth, Wessberg, & Olausson, 2016). This method was developed from preliminary research (Takahata et al., 2016), and in that study, the stimulus range was set to the nipple and the areola. In this study, to further standardize the stimulation method, the stimulation range was limited to the nipple, and the stimulation finger was also limited. In addition, tactile stimulation was set weakening the intensity of the stimulus to become a gentle touch. At end of the intervention, the women removed any remaining cream from their nipples with a prepared wet towel (Iwatsuki, Tokyo, Japan) depending on their own preferences.

The researcher explained to the women about the methods of the second day of intervention using a <Appendix 14>. The researcher received the women's email addresses. Women reported the estimated start time of breast stimulation. All the necessary goods (metronome, timer, nursing cover and wet towel) were lent to women.

***Second day.***

Breast stimulation in the home was done until noon, in case there was an emergency so they could have access to health care within business hours. Women sent an email to the researcher within 10 min from the start of intervention after preparing for the intervention. The content of the e-mail was only the name and report of start. The researcher sent reminder e-mails to participants that did not report the start of breast stimulation one hour late from estimated time.

At end of the breast stimulation, women sent an e-mail mail to the researcher within 10 min after they went to the lavatory to check any abnormally bleeding. Participants were provided with a consultation from the researcher at any time.

***Third day.***

The researcher at the clinic confirmed participant's status of the second day. The similar process with first day was applied also to the third day.

**Control group.**

On the first day, after first saliva collection, women continued to watch the movie in silence.

On the second day, they resumed their usual life at home.

On the third day, they repeated the process of the first day.

**Outcome Measurements**

**Primary Outcome: salivary oxytocin level.**

***Collection of salivary oxytocin procedure.***

In the preliminary study (Takahata et al., 2016), the measurement point of oxytocin was before the intervention, 15 minutes after intervention, 30 minutes, 60 minutes, 75 minutes, and 90 minutes after the intervention. As a result, the median value suggested that saliva oxytocin increased 30 minutes after intervention. Furthermore, women with the highest oxytocin level increased after 60 minutes were also confirmed. Furthermore, in order to confirm at what point the intervention effect disappeared, 75 minutes after the intervention was set.

From the above, measurement points of this research were set following: before intervention (baseline), 30 minutes after intervention, 60 minutes, 75 minutes. Saliva collection was done from 12:30 - 16:00 to control for diurnal effects. The participants were provided with the following instructions <Appendix 15>: (1) exercise/ alcohol intake is prohibited on the day before the study participation; (2) on the day of study participation, refraining from taking caffeine and take lunch; (3) smoking, brushing,

eating and drinking were prohibited (only water can be ingested), from one hour before collection; (4) no lipstick; (5) gargle before intervention and (6) after starting the intervention refrain from speaking.

Four saliva samples were collected before the intervention and at 30, 60, and 75 minutes after the breast stimulation intervention (See **Figure 1**). The amount of saliva samples was eight per participant. Participants were instructed to wait three minutes for saliva to accumulate naturally in the mouth (passive drool). Then, the participants collected their saliva for each sample in a 2.0 mL polypropylene tube (Eppendorf, NY, USA) using a cut straw. This process was repeated three times. The minimum saliva amount should be 1.5 ml and the goal of the collection target was 2.0 ml. If the amount of saliva is small in the passive drool method, participants were instructed to self-massage the submandibular gland line.

***Analysis of salivary oxytocin.***

After collection the samples were immediately stored in a freezer (Cryo Porter CS-80C, Scinics Corp., Tokyo, Japan) at -80℃. Oxytocin level assayed using Carter et al's (2007) method. We added 500 KIU/μL (Kallikrein Inhibitory Unit) of aprotinin to prevent proteolytic degradation after thawing saliva.

Oxytocin level was run by enzyme-linked immunosorbent assay (ELISA; ENZO Life Sciences, NY, USA) at the Nagasaki University Graduate School of Medicine and Dentistry in Professor Kazuyuki Shinohara's Laboratory. The ELISA manual reports that intra-assay and inter-assay coefficients of variability are 12.6 – 13.3% and 11.9 – 20.9%. The technical report by Salimetrics reported that acceptable values intra-assay and inter-assay coefficients of variability are <10% and <15% respectively. This study reported inter- and intra-assay coefficients of variability from the dilutions of test samples, because the researcher did not measure the same sample in each plate and so the exact variation was unknown. The intra-assay and inter-assay coefficients of variability were <2.68% and <5.49% respectively.

In ELISA analysis, in order to exclude the influence of well uniformity and scratches, the absorbance at 570 nm is also determined as a sub wavelength, and the

difference from the absorbance at 405 nm is treated as true absorbance. In this study, the mean value of the value converted from oxytocin value using software was calculated from the obtained true absorbance. This resulted in oxytocin levels ranging from 14.02 to 492.87pg/mL. The range of duplicate assay values in this study was 0.59 to 8.43 pg/mL. In addition, of the 210 specimens with duplicate assays, the difference of binary values was small: they were less than 2 pg / mL ($n$ = 188 samples; 89.5%). Based on the above results, in this study, samples of single assay were also included in the final analysis.

**Secondary outcomes.**

***Background factors and delivery outcomes.***

The questionnaire contained: age, height, weight, non-pregnancy weight, scheduled delivery date, estimated date of delivery confirmed by echo, whether participant's mother was post-term with participant, whether you have had the opportunity to interact with children, marital status, living with partner, and educational level.

The researcher gathered these data from the medical record: medical history, onset of labor within 72 hours, labor induction, augmentation of labor, the gestational weeks at delivery, the mode of delivery, duration of labor, postpartum hemorrhage, hypotonic contractions, the amount of bleeding delivery, birth weight, fetal gender, Apgar score (1 and 5 min), meconium stained liquor, NICU admission immediately after birth.

***Evaluate the uterine contraction and fetal hart rate.***

During the intervention period, external cardiotocography readings were obtained to record uterine activity and fetal heart rate. Basic record time was from 30 min before the intervention to 90 min after the start of the intervention. The recording speed of the cardiotocography was 3cm/min. The evaluators were composed of two midwives (researcher and another midwife) who had five years of experience. The evaluation of the baseline was defined as the line just before the contraction and

measured by the height: from one to five from baseline to peak of contraction. Tachysystole of uterine activity was defined more than five contractions in 10 minutes, averaged over a 30-minute window (Macones, Hankins, Spong, Hauth, & Moore, 2008).

If participants had not delivered by the prenatal checkup around the estimated date of delivery, they performed a non-stress test at that time even with a healthy pregnancy. This is the routine test that the hospital chose. The researcher collected and evaluated the data from a non-stress test after their childbirth.

***Collection and analyses of polymorphisms of oxytocin receptor.***

Even if the amount of OT secretion increases, when the sensitivity of the receptor is low, the effect may not be confirmed as a clinical outcome. In order to examine that tendency, SNP was measured. Preliminary studies measured three sites that were reported to be associated with oxytocin levels or labor. In this study, only two sites that appear in Japanese in balanced were analyzed.

At the very beginning of experiment after women gargled, buccal mucosa samples were obtained Sample Collection Swab (Epicentre, Wisconsin, U.S.). These swabs collect buccal cells for a safe and gentle DNA extraction and polymerase chain reaction analysis. All samples are analyzed for the two oxytocin receptor gene polymorphisms (i.e., rs53576, rs2254298) by genotyping using the TaqMan single nucleotide polymorphism (SNP) assay (Applied Biosystems, Thermo Fisher, MA, USA) at the Nagasaki University Graduate School of Medicine and Dentistry in Professor Kazuyuki Shinohara's Laboratory.

**Feasibility.**

***Acceptability of intervention method.***

Women reported on three variables using analog scales: fatigue by using a visual analog scale for fatigue from 0 (*no fatigue*) to 100 (*greatest fatigue*) and breast stimulation pain from 0 (*no pain*) to 100 (*worst pain*). It was also used for saliva collection from 0 (*no fatigue*) to 100 (*greatest fatigue).*

Discomfort was indirectly assessed by asking participants: would you want to do it even at the next pregnancy, and would you want to provide information about this technique to friends, using the 4-level Likert scale: 1 (*Very much*) to 4 (*I do not think so*). In addition, participants responded to questions about the environment where intervention was carried out.

***Practicality of intervention methods.***

Determine the impact on the data of adverse events that occurred throughout the intervention. Calculate the proportion of total oxytocin saliva specimens from those that were able to be analyzed, and determine the dropout rate.

**Analyses**

The means, standard deviations, and medians of all variables were calculated. Independent t-test and paired t-tests were carry out for oxytocin analyses. Welch's t test was conducted when the data had unequal variances. When the assumption of normality was not maintained, the Mann-Whitney's U test was conducted. The Bonferroni correction was used for correcting issues related to multiple comparisons. Using the baseline data on the first day as the co-variable, ANCOVA was carried out and presented as an adjusted mean oxytocin value. Repeated measures ANOVA were performed for uterine contraction analysis. Chi-square test was carried out, when the expected value was five or less, Fisher's exact probability test was conduct. Pearson’s correlation coefficient was used to test the correlation between salivary oxytocin level and participants’ characteristics. Statistical analyses were performed using SPSS Statistics version 24. (Static Base and Advanced Statistics, IBM Japan, Tokyo, Japan).

**Ethical consideration**

The study protocol was approved by the Institutional Review Board of St. Luke’s International University, Tokyo, Japan (No. 15-085). This study was registered in the Clinical Trials Registry of the University Hospital Medical Information Network in Japan (UMIN 000020797).

**Ethical consideration of individuals.**

This study was an intervention study involving minimal invasion for humans. In accordance with "Ethical Guidelines on Medical Research for Human beings" (Trial on April 27, 2015), it was implemented in consideration of human rights protection and in compliance with the research plan.

**Profit to be brought to the participants.**

Although this research could not be expected to directly benefit the participants, it will be a social contribution to research.

**Possible disadvantages arising from participating in the research study and prevention of harm.**

Possible adverse events in this study could be premature rupture or uterine hyperstimulation and tachysystole and its accompanying non-reassuring fetal status. In addition, skin symptoms or itching and allergic reactions caused by breast stimulation using a lubricant were possible. All risks and disadvantages arising from research participation were described in the participants' instruction manual and explained orally at recruitment.

Upon intervention, fetal heart rate monitoring was performed. If adverse events occur during this study, they were reported directly and immediately to the obstetrician and appropriate care and recording was taken.

**Treatment and supplementation in case of accidents.**

If an adverse event occurred due to the implementation of this study and health damage to the participant occurred, the researcher and the hospital responded so that appropriate treatment and other necessary measures could be taken. Health insurance was applied to the treatment to be provided, and other monetary compensation was not provided. However, skin symptoms such as allergic symptoms and itching due to breast stimulating lubricant was eligible for compensation by the researcher's insurance.

Before starting the research, the insurance (Will and e-kango of the Association of Japanese Nursing Schools Council of the Japan Institute of Nursing) was prepared to cover treatment and compensation in case of an accident. (Limits up to March 2016: 50 million yen, after April 2016: 100 million yen, since insurance compensation amount will be changed from the next fiscal year).

**Ensure freedom to decline to participate.**

Participation in this study was determined by voluntary intention. Even after consent participants can withdraw it at any time. Participants were told that there was no disadvantage to them due to non-participation in the research. The research confirmed consent to participate in the research by asking the women to sign the consent form, and when giving the consent form, the women also received the refusal form <Appendix 16>.

**Delineation of responsibility for participant's research participation expenses and participant's usual medical expenses.**

There was a possibility that participants could incur additional expenses by participating in the study. For example, transportation expenses would be required to visit the hospital. This transportation fee was included in the remuneration to the participant. In addition, the examination expenses required for this research were covered through the researcher's account. Therefore, participation in this study did not increase the participant's burden of medical expenses. However, the participant payed the inspection expenses, examination fee, and medication expenses carried out in ordinary medical treatment .

**Response after the end of the research.**

After the end of the research, the results will be published in doctoral dissertations and academic journals. Also, report results will be sent to the research cooperation facility. When a participant or medical staff requests disclosure of research results, it can be disclosed by contacting the researcher's contact information described

in the research cooperation instruction.

**Anonymization of participant's personal information.**

After acquiring research consent, a correspondence table between ID and name was created and the information gathering of this research was done under consolidation anonymization. Samples of saliva specimens are discarded immediately after analysis. Data (data collection paper, flash memory etc.) gathered in the research is saved for at least five years, and then all data is erased in a state that cannot be restored, or finely shredded and destroyed.

## Results

### Sample, Demographic Characteristics and Retention Rates

The flow diagram of the data collection is shown in **Figure 2**. Recruitment of the subjects was decided at the season. The first half participants were assigned to the control group. Last half were assigned to the intervention group. Of the 131 recruited women meeting the eligibility criteria 58 refused to participate. Women refused for the following reasons; not meeting inclusion criteria (*n* = 12), their house was far away from the hospital (*n* = 17), declined to participate as schedule was not met (*n* = 20), unknown (*n* = 9). Therefore 73 women consented to participate in the research. Of the 37 women recruited for the intervention group, 15 women were excluded, leaving 22 women to participate in the intervention. The reasons for dropout were: already hospitalized for delivery (*n* = 5), declined to participate (*n* = 3), research schedule conflict (*n* = 7). In the control group, 36 women were recruited, 16 women were excluded, and 20 women finally participated in the research. The reasons for dropout were: already hospitalized for delivery (*n* = 5), due to illness (*n* = 2), declined to participate (*n* = 5), conflict due to research schedule (*n* = 4).

**Table 3** describes baseline characteristics of the participants. The mean ages were 28.8 years (*SD* = 3.8) in the intervention group and 29.3 years (*SD* = 3.7) in the control group. In both groups, mean gestational weeks had started intervention at 38.8 weeks. All women were married and lived with their partner (or had return to their parent's home after birth). Educational levels were largely similar. Most women received ultrasonography at less than 12 weeks gestation to determine the expected delivery date. In addition, there were no significant differences in baseline characteristics of the participants between two groups.

### Primary Outcome: salivary oxytocin level

#### Selection of Samples to be Analyzed.

Four samples were taken per day for each subject, and up to 8 samples were collected in two days (a projected 336 samples), and final collection total was 289

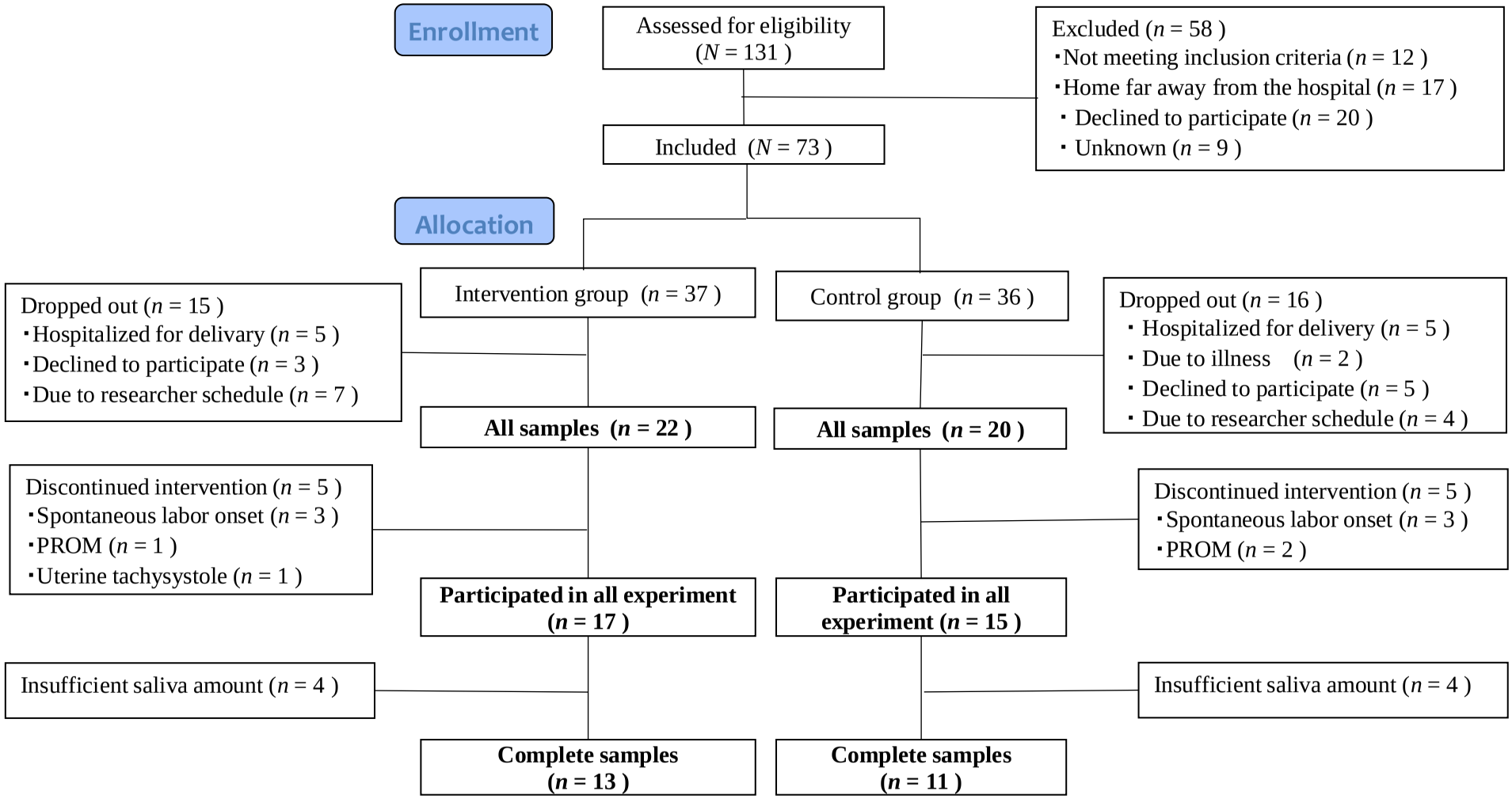



*Figure 2.* Flow diagram of participants from eligibility to data analyses.

**Table 3**

***Comparison of Participants Baseline Characteristics by Intervention and Control Groups (N=42)***

| Characteristics | Intervention ( *n* = 22 ) | | Control ( *n* = 20 ) | | *p* |
|---|---|---|---|---|---|
| **Age (in years): *M* [*SD*]** | 28.8 | [3.8] | 29.3 | [3.7] | .714 |
| **Gestational weeks at start intervention: *M* [*SD*]** | 38.8 | [0.4] | 38.8 | [0.5] | .955 |
| **Living with partner** | 22 | | 20 | | - |
| **Married** | 22 | | 20 | | - |
| **Educational level** | | | | | |
| Junior high school or high school | 4 | (18.2) | 3 | (15.0) | |
| Diploma | 9 | (40.9) | 10 | (50.0) | .839 |
| Undergraduate | 9 | (40.9) | 7 | (35.0) | |
| **Due date confirmation by ultrasonography** | | | | | |
| <12 wks | 21 | | 19 | | 1.000 |
| 13-24 wks | 1 | (4.5) | 1 | (5.0) | |
| **Own mother had postterm pregnancy** | 0 | | 1 | (5.0) | .476 |
| **Had contact with children** | 5 | (22.7) | 4 | (20.0) | 1.000 |
| **Taking medicine** | 6 | | 5 | | .867 |
| **Has sores or wound in the mouth** | 1 | | 0 | | 1.000 |
| **Fetal gender (boy)** | 12 | | 13 | | .491 |
| **Breast stimulation before intervention** | 13 | | 15 | | .275 |
| **Prenatal checkup** | | | | | |
| On 1st day of intervention | 10 | | 8 | | .491 |
| On 3rd day of intervention | 1 | | 2 | | .319 |

*Note.* Values are expressed as mean ± *SD* [range] or numbers (%).

samples from 42 women (151 samples from the intervention groups, and 138 samples from the control group) (Collection rate 86.0%). Of these samples, 32 samples (loss rate 11.1%) were not analyzed because the amount of saliva was insufficient for analysis of the oxytocin level. Insufficient amount of saliva samples were: 13 samples in the intervention group and 19 samples in the control group. Finally, 78 samples were obtained on the first day for the intervention group and 67 samples in the control group, and on the third day there were 60 samples in the intervention group and 52 samples in the control group (total number was 257 samples; analysis rate 88.9 %). *Complete samples* (those with sufficient amount of saliva) that could be analyzed by all eight measurement points were 13 women in the intervention group ($n$ = 104 samples) and 11 women in the control group ($n$ = 88 samples) and total samples were 192.

**Salivary Oxytocin.**

In this section, oxytocin was analyzed in *all samples* (257 samples: intervention group $n$ = 22, control group $n$ = 20) and *complete samples* (192 samples: intervention group $n$ = 13, control group $n$ = 11), which could be analyzed with all eight points measured. (See **Figure 2**).

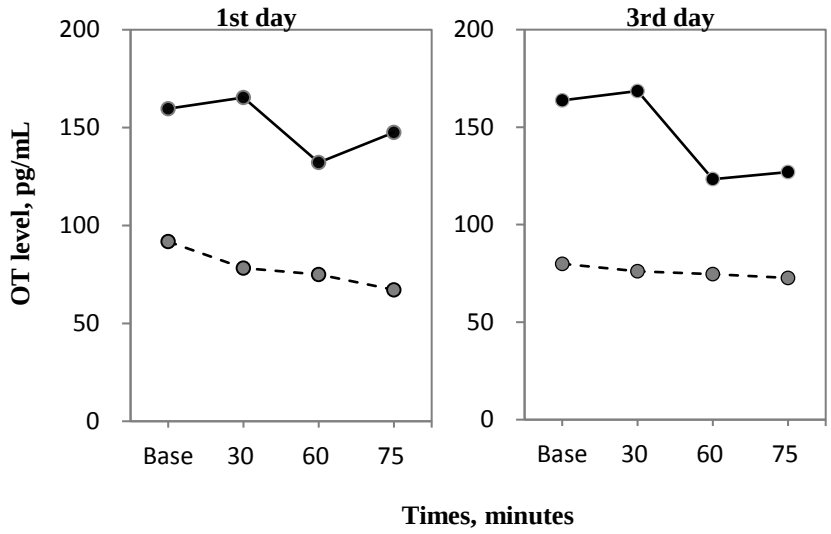
Actual measurement value, amount of change, and percentage of change were analyzed (**Figure 3** and **Figure 4**). For each individual, amount of change of oxytocin at each measurement time point was calculated based on the pre-intervention value for each individual (e.g., $O_2$-$O_1$). For each percentage of change, oxytocin at each measurement time point was converted to a rate of change of 100% for individuals based on pre-intervention value (e.g., $O_2/O_1 \times 100$). For the change amount and the change rate, the mean value of the whole was calculated for each intervention day.

**Table 4** describes mean salivary oxytocin level of both groups for *all samples* and *complete samples*. **Table 5** describes adjusted means salivary oxytocin level of both groups for *all samples* and *complete samples*.

***All samples.***

Oxytocin level at 30 minutes on the third day was significantly higher in the

## Actual measurement value




## Amount of change

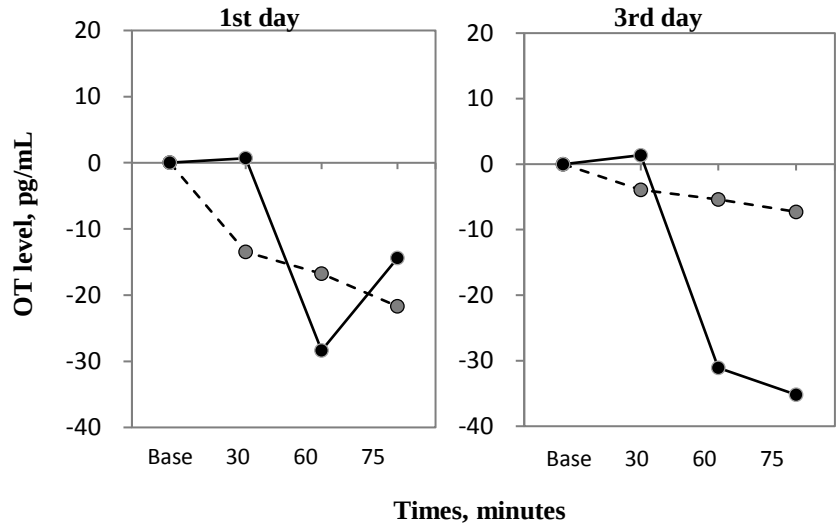
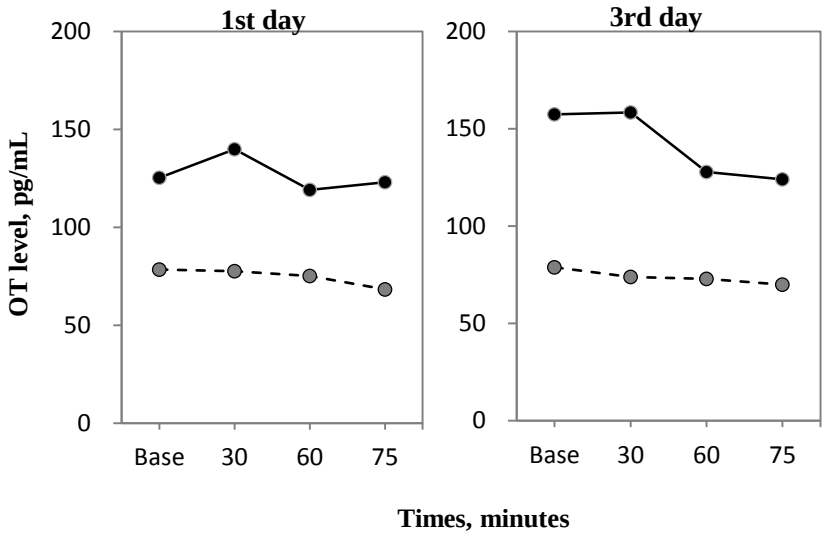



*Figure 3.* The comparison of mean oxytocin levels on days 1 and 3 (*all samples* ). Intervention group (solid line) versus control group (dashed line). (257 samples: intervention group $n$ = 22, control group $n$ = 20)

## Actual measurement value




## Amount of change

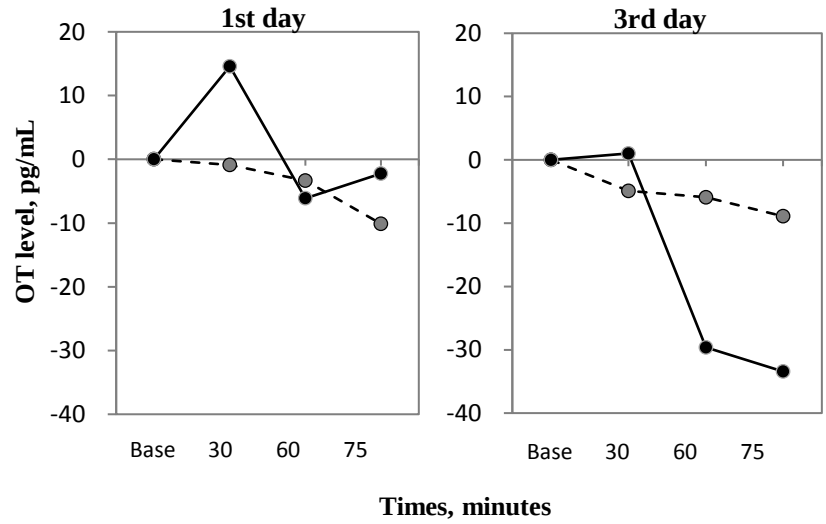



*Figure 4.* The comparison of mean oxytocin levels on days 1 and 3 (*complete samples* ). Intervention group (solid line) versus control group (dashed line). (192 samples: intervention group $n$ = 13, control group $n$ = 11)

**Table 4**

***Comparison of Mean Salivary Oxytocin Levels (pg/mL) in all Samples and Complete Samples by Intervention and Control Groups***

| ***All samples*** ($N$ = 42) | | | **Intervention** | | | **Control** | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | *n* | *M* | *SD* | *n* | *M* | *SD* | *p* [a] |
| **First** | $O_1$ | Base | 21 | 159.6 | 101.7 | 17 | 91.8 | 39.5 | .009 |
| | $O_2$ | 30 min | 19 | 165.4 | 130.0 | 17 | 78.3 | 33.9 | .010 |
| | $O_3$ | 60 min | 19 | 132.1 | 85.9 | 17 | 75.0 | 35.1 | .013 |
| | $O_4$ | 75 min | 20 | 147.6 | 89.2 | 16 | 67.1 | 33.2 | .001 |
| **Third** | $O_5$ | Base | 15 | 163.8 | 122.2 | 13 | 79.9 | 42.5 | .023 |
| | $O_6$ | 30 min | 14 | 168.5 | 110.2 | 13 | 76.0 | 26.5 | .008 |
| | $O_7$ | 60 min | 14 | 123.3 | 100.7 | 13 | 74.5 | 33.1 | .106 |
| | $O_8$ | 75 min | 16 | 127.0 | 99.4 | 13 | 72.7 | 25.7 | .050 |

*Note.* In all samples group, the assumption of normality was adopted in the 9 out of 16 samples. The assumption of normality was possible at $O_1$ and $O_4$ and $O_6$ in both groups. [a]The independent sample t-test was conducted.

| ***Complete samples*** ($n$ = 24) | | | **Intervention** | | | **Control** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | *n* | *M* | *SD* | *n* | *M* | *SD* | *p* [a] | *p* [b] |
| **First** | $O_1$ | Base | 13 | 125.3 | 92.1 | 11 | 78.5 | 41.3 | .134 | .207 |
| | $O_2$ | 30 min | 13 | 139.9 | 134.9 | 11 | 77.6 | 39.6 | .154 | .228 |
| | $O_3$ | 60 min | 13 | 119.2 | 88.6 | 11 | 75.1 | 40.7 | .126 | .277 |
| | $O_4$ | 75 min | 13 | 123.1 | 90.3 | 11 | 68.4 | 36.7 | .063 | .134 |
| **Third** | $O_5$ | Base | 13 | 157.4 | 125.5 | 11 | 78.8 | 46.3 | .053 | .041 |
| | $O_6$ | 30 min | 13 | 158.4 | 107.6 | 11 | 73.9 | 28.1 | .017 | .030 |
| | $O_7$ | 60 min | 13 | 127.7 | 103.4 | 11 | 72.9 | 35.7 | .093 | .331 |
| | $O_8$ | 75 min | 13 | 123.9 | 105.4 | 11 | 69.9 | 26.3 | .096 | .531 |

*Note.* In all samples group, the assumption of normality was adopted in the 8 out of 16 samples. The assumption of normality was possible at $O_4$ and $O_6$ in both groups. [a]The independent sample t-test was conducted. [b]The Mann-Whitney's U test was conducted.

**Table 5**

***Adjusted Means Oxytocin Level (pg/mL) of ANCOVA using Baseline of First Day as Covariable in All Samples and Complete Samples by Intervention and Control Groups***

| *All samples* | | | Intervention | | | Control | | | |
|---|---|---|---|---|---|---|---|---|---|
| ($N$ = 42) | | | *n* | *M* | *SD* | *n* | *M* | *SD* | *p* |
| **First** | $O_1$ | Base | | | | | | | |
| | $O_2$ | 30 min | 19 | 131.1 | 13.5 | 17 | 116.6 | 14.4 | .486 |
| | $O_3$ | 60 min | 18 | 112.8 | 9.5 | 17 | 100.2 | 9.8 | .382 |
| | $O_4$ | 75 min | 19 | 126.2 | 10.2 | 16 | 97.0 | 11.2 | .068 |
| **Third** | $O_5$ | Base | 14 | 141.7 | 15.7 | 12 | 108.5 | 17.0 | .176 |
| | $O_6$ | 30 min | 14 | 146.3 | 12.7 | 12 | 100.1 | 13.7 | .025 |
| | $O_7$ | 60 min | 13 | 108.9 | 11.0 | 12 | 93.4 | 11.5 | .354[a] |
| | $O_8$ | 75 min | 15 | 108.0 | 9.9 | 12 | 97.8 | 11.2 | .516 |

*Note.* In all samples group, the assumption of normality was adopted in the 9 out of 16 samples.The assumption of normality was possible at $O_1$ and $O_4$ and $O_6$ in both groups.

[a]Assumptions violating test of parallelism are indicated in a $O_7$.

| *Complete samples* | | | Intervention | | | Control | | | |
|---|---|---|---|---|---|---|---|---|---|
| ($n$ = 24) | | | *n* | *M* | *SD* | *n* | *M* | *SD* | *p* |
| **First** | $O_1$ | Base | | | | | | | |
| | $O_2$ | 30 min | 13 | 111.3 | 9.5 | 11 | 111.4 | 10.4 | .997 |
| | $O_3$ | 60 min | 13 | 99.6 | 6.8 | 11 | 98.2 | 7.4 | .893 |
| | $O_4$ | 75 min | 13 | 105.3 | 10.7 | 11 | 89.4 | 11.7 | .339 |
| **Third** | $O_5$ | Base | 13 | 133.7 | 16.0 | 11 | 106.7 | 17.5 | .279 |
| | $O_6$ | 30 min | 13 | 138.5 | 13.2 | 11 | 97.4 | 14.4 | .053 |
| | $O_7$ | 60 min | 13 | 107.0 | 10.8 | 11 | 97.4 | 11.8 | .565[b] |
| | $O_8$ | 75 min | 13 | 102.7 | 9.6 | 11 | 95.0 | 10.5 | .602[b] |

*Note.* In all samples group, the assumption of normality was adopted in the 8 out of 16 samples.The assumption of normality was possible at $O_4$ and $O_6$ in both groups.

[b]Assumptions violating test of parallelism are indicated in a $O_7$ and $O_8$.

intervention group than in the control group ($t$ (14.612) = 3.048, $p$ = .008). (See **Table 4**).

ANCOVA results show that the oxytocin level at 30 minutes on the third day was also significantly higher in the intervention group than in the control group ($F$ (1,23) = 5.783, $p$ = .025). (See **Table 5**).

In the intervention group, the basal oxytocin levels before the intervention between the first and third day indicated no significant difference ($t$ (13) = 1.673, $p$ = .118).

In the intervention group, the oxytocin level increased at the 30-minute point after intervention on both days and at the 60-minute point the value was lower than the baseline. In the control group, the oxytocin levels decreased with the passage of time on both days.

***Complete samples.***

The changes in oxytocin level from breast stimulation at the 30 minute on the third day in the actual measurement value, was 158.4 pg/mL (*SD* 107.6) in the intervention group, and 73.9 pg/mL (*SD* 28.1) in the control group. Oxytocin level was significantly higher in the intervention group than in the control group ($t$ (13.974) = 2.723, $p$ = .017) (**Table 4**).

ANCOVA results show that the oxytocin level at 30 minutes on the third day was not significantly different in the intervention group than in the control group ($F$ (1,21) = 4.214, $p$ = .053). (See **Table 5**).

The assumption of normality was adapted in eight out of 16 samples. The analysis of *complete samples* was by Mann-Whitney's U test. The results of these tests are also shown in **Table 4**. As a result, oxytocin level was a significant difference in the both groups at the baseline and at 30 minutes on the third day. There were no significant differences between the two groups in the amount of change and the percentage of change.

The results show that values at baseline, at 30 minutes and 60 minutes were higher in the third day than the first day, however there was no significant difference.

On the third day of the intervention group the baseline was 32.0 pg/mL higher than that of the first day baseline, but as a result of carrying out the paired *t*-test, there was no significant difference ($t$ (12) = 1.874, $p$ = .086). In the control group, the baseline on first and third day were same at 78 pg/mL, oxytocin level was the highest at the baseline on both days and then decreased gradually.

**Secondary outcomes**

**Delivery outcomes.**

There was no significant difference in rate of spontaneous onset of labor from breast stimulation between 18 women (81.8%) in the intervention group and 16 women (80.0%) in the control group (**Table 6**). Participants who had labor within 72 hours and had carried out all three days of protocol were one (5.9%) of 17 women in the intervention group and two (13.3%) out of the control group. There was no significant difference in the number of days from the start of intervention to spontaneous onset of labor in the women who had carried out all three days of protocol. There were no significant differences in other delivery outcomes between two groups.

**Uterine Contractions.**

The rates of *adequate uterine contraction* on the first day of intervention were 16 women (72.7%) out of 22 women in the intervention group, three women (15.0%) out of 20 women in the control group. There was significant difference between the groups ($\chi^2$ (1) = 14.092, $p$ < .001). The rates of *adequate uterine contraction* on the third day of intervention were 13 women (76.5%) out of 17 women in the intervention group, no women (0.0%) out of 15 women in the control group. There was significant difference between the groups ($\chi^2$ (1) = 19.319, $p$ < .001). (**Table 7**)

The frequency of uterine contractions were compared in the following four categories: *baseline* - "30 minutes before intervention"; *0 - 30*: "up to 30 minutes from the start of intervention"; *31-60*: "from 31 minutes after the start of intervention to 60 minutes"; *61-90*: "61 minutes after intervention started to 90 minutes ". In the comparison between groups, the mean frequency of uterine contractions were

**Table 6**

***Comparison of Delivery Outcomes by Intervention and Control Groups (N=42)***

| | Intervention ( *n* = 22 ) | | Control ( *n* = 20 ) | | *p* |
|---|---|---|---|---|---|
| **Gestational weeks at delivery: *M* [*SD*]** | 40.0 | [0.7] | 40.0 | [0.7] | .771 |
| **Onset of labor** | | | | | |
| Spontaneous | 18 | (81.8) | 16 | (80.0) | 1.000 |
| Induced | 4 | (18.2) | 4 | (20.0) | |
| **Onset of labor during intervention period and within 72h after intervention** | 5 | (22.7) | 6 | (30.0) | .734 |
| **Onset of labor only within 72h after intervention** | 1 | (5.9) | 2 | (13.3) | |
| **Duration from the intervention to delivery (hours)[a]** | 7.9 | [3.7] | 6.8 | [3.7] | .403 |
| **Mode of delivery** | | | | | |
| Vaginal delivery | 20 | (90.9) | 16 | (80.0) | |
| Instrumental delivery | 2 | (9.1) | 3 | (15.0) | .489 |
| Cesarean section | 0 | (0.0) | 1 | (5.0) | |
| **Weak labor** | 5 | (22.7) | 8 | (40.0) | .320 |
| **Labor augmentation** | 3 | (13.6) | 7 | (35.0) | .214 |
| **No synthetic oxytocin used during labor** | 7 | (31.8) | 11 | (55.0) | .212 |
| **Duration from the intervention to delivery[a]** | | | | | |
| First stage (min): *M* [*SD*] | 734 | [497.1] | 816 | [663.9] | .650 |
| Second stage (min): *M* [*SD*] | 118 | [179.0] | 153 | [241.4] | .599 |
| Third stage (min): *M* [*SD*] | 4 | [2.9] | 3 | [1.4] | .052 |
| Duration of labor (hours): *M* [*SD*] | 14.2 | [8.6] | 16.9 | [12.4] | .441 |
| **Bleeding** | | | | | |
| Within the 2 hours after birth (g): *M* [*SD*] | 47 | [22.2] | 60 | [39.4] | .208 |
| Total bleeding volume (g): *M* [*SD*] | 317 | [184.1] | 450 | [270.2] | .074 |
| Postpartum hemorrhage | 3 | (13.6) | 7 | (35.0) | .152 |
| **Premature rupture of membranes** | 5 | (22.7) | 3 | (15.0) | .700 |
| **Uterine hyperstimulation** | 0 | (0.0) | 0 | (0.0) | |
| **Apgar Score at 1 min (＜7)** | 0 | (0.0) | 1 | (5.0) | .476 |
| **Apgar Score at 5 min (＜7)** | 0 | (0.0) | 0 | (0.0) | - |
| **Birth weight (in grams): *M* [*SD*]** | 3090 | [287] | 3077 | [277] | .881 |
| **NICU admission immediately after birth** | 1 | (4.5) | 1 | (5.0) | 1.000 |
| **Meconium stained Liqour (≧2+)** | 1 | (4.5) | 4 | (20.0) | .174 |
| **Stillbirth** | 0 | (0.0) | 0 | (0.0) | - |

*Note*. Values are expressed as mean ± *SD* [range] or numbers (%). NICU = neonatal intensive care unit.

[a]Women who completed 3-days protocol without induction of labor (Intervention, n = 13; Control, n = 12).

**Table 7**

***Characteristics of Uterine Contractions by Intervention and Control Groups (N=42)***

| | Intervention | | | Control | | | |
|---|---|---|---|---|---|---|---|
| | *n* | *M* | *SD* | *n* | *M* | *SD* | *p* |
| **Adequate uterine contractions**[a] | | | | | | | |
| First day | 22 | 16 | (72.7) | 20 | 3 | (15.0) | <.001 |
| Third day | 17 | 13 | (76.5) | 15 | 0 | (0.0) | <.001 |
| **Uterine contraction frequency: *M* [*SD* ]** | | | | | | | |
| ***First day*** | | | | | | | |
| Before intervention | 22 | 1.7 | * [1.9] | 20 | 1.1 | [1.0] | .189 |
| 0-30 min | 22 | 5.3 | [3.9] | 20 | 1.6 | [1.3] | <.001 |
| 31-60 min | 22 | 6.7 | [3.9] | 20 | 1.5 | [1.3] | <.001 |
| 61-90 min | 22 | 4.0 | [2.8] | 20 | 1.5 | [1.8] | .001 |
| ***Third day*** | | | | | | | |
| Before intervention | 17 | 0.8 | * [0.8] | 15 | 0.5 | [0.6] | .270 |
| 0-30 min | 17 | 4.1 | [3.2] | 15 | 1.0 | [1.1] | .001 |
| 31-60 min | 17 | 4.8 | [4.0] | 15 | 0.8 | [0.9] | .001 |
| 61-90 min | 17 | 3.9 | [2.9] | 15 | 0.7 | [0.7] | <.001 |
| **Uterine contractions of first non-stress test in prenatal check: *M* [*SD* ]** | 9 | 1.1 | [1.4] | 3 | 0.0 | [0.0] | .202 |
| **Long uterines contraction**[b] | 22 | 5 | (22.7) | 20 | 2 | (10.0) | .414 |
| **Uterine tachysystole with FHR deceleration**[c] | 22 | 1 | (5.9) | 20 | 0 | (0.0) | 1.000 |

*Note.* Values are expressed as mean ± *SD* [range] or numbers (%). FHR = fetal heart rate. [a]At least three contractions lasting for 40 seconds each occurring within 10 minutes. [b]Contractions lasting over 120 seconds without FHR deceleration after stimulation, a sum of three days. [c]After 30 minutes, permission to restart was obtained from the obstetrician. *$p$ <.001

significant different in each of the categories after the start of the intervention *(p* ＜ .05) (**Figure 5)**.

In the group comparison, the number of contractions in the intervention group was significantly increased on the first day ($F$ (3) = 18.593, $p$ < .001) and on the third day ($F$ (3) = 11.054, $p$ <. 001). In the control group, there was no significant difference between the first day ($F$ (3) = 0.893, $p$ = .450) and the third day ($F$ (3) = 1.955, $p$ = .135). (**Table 5**)

During the prenatal check around the estimated date of delivery, nine women in the intervention group and three women in the control group received the non-stress test of routine. There was no significant differences between the groups on the first non-stress test.

**Onset of labor with oxytocin.**

As a result of excluding induction of labor due to medical reasons such as pregnancy-induced hypertension, were three women in the intervention groups and one woman in the control group who were induced because they went beyond the 40 weeks gestation. Among them, one woman from the control group could not be measured because of insufficient saliva and no oxytocin result was present.

The hypothesis was the difference of oxytocin level on the 3rd day in the spontaneous onset of labor and induction of labor. However, there were none in the control group who had induction of labor. Due to insufficient data, it could not be analyzed.

**Polymorphisms of Oxytocin Receptor.**

All samples could be analyzed through SNP detection; the baseline of oxytocin able to analysis on the first day was 21 women in the intervention groups and 17 women in the control groups (**Table 8**).

***rs2254298.***

In the intervention group, there were 10 (47.6%) of the GG, nine (42.9%) of the

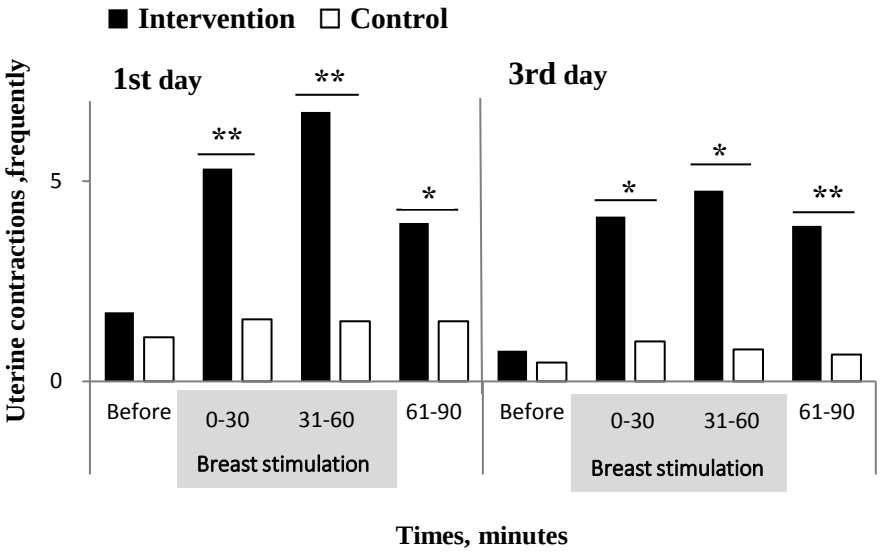



*Figure 5.* Mean number of uterine contractions during 30-min for intervention group ($n$ = 22) versus control group ($n$ = 20)
**$p$ < .001. *$p$ < .05.

**Table 8**

***Comparison of Oxytocin Receptor SNP and Basal Oxytocin Levels (pg/mL) on First Day by Intervention and Control Groups (n = 38)***

| | | Intervention | | | | Control group | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **SNP** | *n* | *M* | *SD* | *p* | *n* | *M* | *SD* | *p* |
| **rs2254298** | GG | 10 | 158 | 98 | .961 | 8 | 78 | 39 | .191 |
| | AG/AA | 11 | 161 | 110 | | 9 | 104 | 38 | |
| **rs53576** | GG | 2 | 206 | 173 | .515 | 2 | 102 | 24 | .707 |
| | AG/AA | 19 | 155 | 98 | | 15 | 90 | 42 | |

*Note.* SNP = single nucleotide polymorphism;GG AND AG/AA.

GA and two (9.5%) of the AA. In the control group, there were eight (47.0%) of the GG, eight (47.0%) of the GA, and one (6.0%) of the AA. The oxytocin level of GG on the first day in both groups was lower than the A allele group, however there was no significant difference (**Table 8**). Furthermore, there was no significant difference between SNP group (3 groups or AA and G allele) and oxytocin level (Data not shown).

Uterine contraction frequency every 30 minutes during the experiment was significantly higher in the GG group (*M* 7.6, *SD* 4.0) than in the A allele group (*M* 3.4, *SD* 2.8) 30 minutes after intervention in the intervention group on the first day ($t$ (20) = 2.898, $p$ = .009). On the third day, the GG group (*M* 5.4, *SD* 3.6) had more contraction times than the A allele group (*M* 3.2, *SD* 2.7) but was not significantly different ($t$ (15) = 1.473, $p$ =.161). In the change of amount from the pre-intervention to 30 minutes after the intervention, but there was no significant difference. Similarly, there was no significant difference in delivery outcomes.

***rs53576.***

In the intervention group, there were two (9.5%) of the GG, five (23.8%) of the GA and 14 (66.7%) of the AA. In the control group, there were two (11.8%) of the GG, seven (41.1%) of the GA, and eight (47.1%) of the AA. The oxytocin level of GG ($n$ = 2) on first day in both groups was higher than A allele, however there was no significant difference (**Table 8**). Furthermore, there was no significant difference between SNP group (3 groups or AA and G allele) and oxytocin level (Data not shown).

SNP and duration of labor were analyzed in 22 women in the intervention group and 20 women in the control group. There was no significant difference between the SNP group (3 groups or GG and A allele or AA and G allele) and duration of labor. In the change of amount from the pre-intervention to 30 minutes after the intervention, there was no significant difference. Similarly, there was no significant difference in delivery outcomes and uterine contraction frequency (Data not shown).

**Other Outcomes and Oxytocin Level.**

There were no correlations between age and educational level and oxytocin

baseline on the first day. Also, there was no significant difference between oxytocin baseline in the first day and whether they were in contact with children on a daily basis or engaged in early childhood education in the past. There was no correlation between oxytocin and the time from mode of labor onset or duration after intervention until labor onset.

**Feasibility**

**Acceptability of Intervention Method.**

No participants dropped out after starting the intervention. A sense of burden for following the protocol was examined for women who participated in the experiment for three days: 17 women in the intervention group and 15 women in the control group (**Table 9**).

For fatigue of saliva collection, both group had a decreasing trend on the third day rather than on the first day. There was no significant difference between the groups. In the intervention group, the fatigue and nipple pain showed no significant change. For the question of discomfort by touching the nipple, women who answered, “I didn't think so” and “I didn't think so at all” accounted for 94.1%. The fatigue of the saliva collection and oxytocin were analyzed by intervention date within each group, but there was no significant correlation. There were also no significant correlations with oxytocin for the fatigue and the pain from breast stimulation. There was no significant association between oxytocin and the presence or absence of positive memory recalled by movie.

In the future, the question: “need to check the methods with the video when doing breast stimulation at home”, there were two women who answered “I didn’t think so at all”. They explained that they understood thoroughly by the first intervention day and could understand it also from the explanation on the handout. Regarding the environment that can be continued (duplicate answers), the number of women who answered, "while watching television” 68.2% was the largest. About those who practiced the intervention who can be continued (duplicate answers), the women who answered to "themselves" 72.7% were the most frequent. Regarding the question of hope to perform this protocol in your next pregnancy, women who answered, “I think so

**Table 9**

***Comparison of Acceptability of Intervention Methods by Intervention and Control Groups (N = 32)***

| | Intervention ( *n* = 17 ) | | Control ( *n* = 15 ) | | *p* |
|---|---|---|---|---|---|
| **Fatigue from saliva collection: *M* [*SD* ]** | | | | | |
| First day | 30.7 | [27.1] | 36.8 | [18.2] | .452 |
| Third day | 25.0 | [21.9] | 34.3 | [17.7] | .194 |
| **Fatigue from nipple stimulation: *M* [*SD* ]** | | | | | |
| First day | 31.3 | [27.8] | | | |
| Third day | 28.0 | [22.2] | | | |
| **Pain from nipple stimulation: *M* [*SD* ]** | | | | | |
| First day | 24.7 | [26.6] | | | |
| Third day | 15.7 | [20.8] | | | |
| **Discomfort from touching nipple** | | | | | |
| I think so very much | 0 | (0.0) | | | |
| I think so | 1 | (11.8) | | | |
| I don't think so | 13 | (76.5) | | | |
| I don't think so at all | 3 | (17.6) | | | |
| **Memories recalled by movie** | | | | | |
| None | 5 | (29.4) | 4.0 | (25.0) | .444 |
| Positive | 12 | (70.6) | 12.0 | (75.0) | |
| Negative | 0 | (0.0) | 0.0 | (0.0) | |
| **Need to check the methods with video** | | | | | |
| I think so very much | 3 | (17.6) | | | |
| I think so | 9 | (52.9) | | | |
| I don't think so | 3 | (17.6) | | | |
| I don't think so at all | 2 | (11.8) | | | |
| **Keep the intervention environment** | | | | | |
| While watching television | 15 | (68.2) | | | |
| While listening music | 7 | (31.8) | | | |
| While talking | 3 | (13.6) | | | |
| None | 1 | (4.5) | | | |
| **Keep the practitioner** | | | | | |
| Self | 16 | (72.7) | | | |
| Electrical breast pump | 3 | (13.6) | | | |
| Partner | 0 | (0.0) | | | |
| **Would do again for next pregnancy** | | | | | |
| I think so very much | 3 | (17.6) | | | |
| I think so | 13 | (76.5) | | | |
| I don't think so | 1 | (4.5) | | | |
| I don't think so at all | 0 | (0.0) | | | |
| **Recommend to your friends** | | | | | |
| I think so very much | 2 | (9.1) | | | |
| I think so | 13 | (79.1) | | | |
| I don't think so | 2 | (11.8) | | | |
| I don't think so at all | 0 | (0.0) | | | |

*Note* .Values are expressed as mean ± *SD* [range] or numbers (%).

very much" and "I think so" accounted for 94.1%. For the question: 'Do you recommend this method to your friends'; women who answered, "I think so very much" and "I think so" accounted for 88.3%.

**Practicality of experimental methods.**

***Adverse events.***

The occurrence of adverse events during delivery is shown in **Table 6**. Premature rupture of membranes, uterine hyperstimulation, Apgar score and other adverse events in the two groups had no significant difference.

The occurrence of adverse events during the experiment is shown in **Table 7**. Lasting contractions over 120 seconds without fetal heart rate deceleration after the start of intervention were five women in the intervention group and two women in the control group. All women had awareness of uterine contractions, and no one complained of pain.

There was one woman in the intervention group who experienced uterine tachysystole with fetal heart rate deceleration. The gestational weeks of the women at the beginning of the intervention was 38 weeks and 6 days, there were no abnormal medical histories and all anticipated normal pregnancies. The researcher confirmed frequent painless uterine contractions two minutes after breast stimulation started. Then stimulation was discontinued 10 minutes after the start. From two minutes after the interruption, severe variable deceleration at the lowest point of 60 bpm was confirmed. After three minutes from the start of the deceleration, the baseline recovered to normal. The researcher did the following: position changed to left lateral position, provided oxygen 5 L/m by a simple face mask for four minutes; oxygen was terminated after it was confirmed that there was no decline in fetal heart rate for 3 minutes. This was immediately reported to the obstetrician and followed up with observation. After that, there were irregular uterine contractions, but there was no pain, and reassuring fetal status was observed. After the lapse of 30 minutes, obstetrician approved the resumption of the intervention, but eventually the intervention was interrupted. Fetal heart rate monitoring continued until 70 minutes after confirming that fetal heart rate deceleration

had declined. This woman was not abnormal afterwards, and labor onset occurred four days after the intervention, and normal delivery occurred in four and one-half hours.

In the intervention group after the end of the third day of intervention, one woman had false labor and she returned home after a 30-minute break in the intervention room. However, this woman did not have spontaneous onset of labor and eventually had induced labor using medical methods.

There was no emergency response, such as telephone, for home intervention at the second day in the intervention group.

In the intervention group, there was one woman who had ruptured membranes at the end of breast stimulation. Reassuring fetal status was observed. The woman commenced spontaneous labor after being hospitalized for premature rupture of membranes, and finally augmentation labor for weak labor.

***Saliva collection and assay.***

There were total of 32 women who could participate in the three days of the experiment. Among them, 24 women provided enough saliva for analysis at the eight collection points (complete samples). The number of specimens that could be collected was 256, but oxytocin level that could be measured was from 192 samples (75.0%). Among of 256 samples, saliva samples exceeding 1 ml were 210 samples (82.0%).

Where possible all samples were run in duplicate, but some samples had insufficient volume to run the analysis. Duplicate assay analysis was 107 samples (77.5%) of the 138 samples in the intervention group and 103 samples (86.6%) of the 119 samples in the control group.

The amount of collected saliva was recorded by eye observation immediately before freezing. The minimum amount of saliva that could be measured in duplicate assay was 0.8 mL in the intervention group and 1.1 mL in the control group. Seven samples (3.6%, n = 4) were unable to be measured in duplicate assay among 193 specimen saliva volumes of 1.5 mL or more. There was one sample that could not be measured in duplicate assay despite the 1.7 mL that was collected. There were no significant differences in the two groups due to saliva collection time and saliva volume.

## Discussion

### Primary Outcome: Salivary Oxytocin Level

At 30 minutes after the intervention on the third day, the oxytocin level was significantly higher in the intervention group than in the control group. The increase in the oxytocin level on the third day showed a similar trend to our preliminary study (Takahata et al., 2016). There is currently no study examining the effects of long hours of breast stimulation on salivary oxytocin level. To our knowledge, this is the first quasi - experimental study to clarify the changes in oxytocin levels by breast stimulation over a three-day period. Notably, oxytocin did not increase sufficiently with only breast stimulation therefore there is a possibility that the previous result was inconsistent.

In the actual measurement of the oxytocin level in the intervention group, the basal 30-minute and 60-minute oxytocin levels were higher on the third day than those on the first day, although the differences were not significant. Upon comparison with the basal oxytocin level before the intervention, there was no significant difference, although the oxytocin level on the third day was higher than the oxytocin level on the first day. On the other hand, in the control group, the oxytocin levels on the first and third days were almost the same. In previous studies, the oxytocin level was reportedly unchanged during pregnancy (Feldman, Weller, Zagoory-Sharon, & Levine, 2007) or increased in late pregnancy (Prevost et al., 2014). The results of the present study showed that the oxytocin level did not extensively fluctuate for several days toward the end of pregnancy in the control group. Therefore, it is possible that the baseline level in the intervention group was increased by the repetitive breast stimulation. In a previous study in which mothers performed massage to their autistic children for 20 minutes during a 3-month period, no significant difference in the oxytocin levels were found before and after the massage; however, the baseline oxytocin level in the mothers and children were significantly higher after the massage (Tsuji et al., 2015). Even in the present study, the oxytocin baseline level of the pregnant women might have been increased by breast stimulation.

On the other hand, the oxytocin level at 60 minutes after the intervention was

lower than the baseline level in the intervention group. As breast stimulation was continued for 60 minutes, it was anticipated that the 60 minutes time point would indicate the peak of the oxytocin level; however, the oxytocin level decreased even when the breast stimulation was continued. As the oxytocin level gradually decreased in the control group, it is possible that the oxytocin level in the intervention group may have naturally decreased after reaching the peak level. In a previous study that investigated the effects of a 15-minute massage on the oxytocin level, it was found that the oxytocin level in the control group also decreased, but not significantly, after sitting in a chair for 15 minutes (Morhenn, Beavin, & Zak, 2012). Although the reason underlying the decrease in the oxytocin level is not clear, the decrease in the oxytocin level at 60 minutes was similar for both days in the intervention group. Several studies have recommended breast stimulation for more than 180 minutes (Damania et al., 1992; Elliott & Flaherty, 1984; Salmon et al., 1986). Elliott & Flaherty (1984) reported that for women who delivered in less than 42 weeks of gestation, stimulation for three hours or longer resulted in significantly shorter days to delivery than stimulation for less than three hours. However, the results of this study suggested that breast stimulation may not increase the oxytocin level even if the breast is stimulated for more than 60 minutes. Earlier studies used longer intervention periods such as 7 days (Singh et al., 2014) or until delivery (Elliott & Flaherty, 1984). Importantly, there is a possibility of increasing the basal oxytocin level using breast stimulation by repeated interventions rather than by a long period of stimulation.

**Secondary Outcomes**

There was no significant difference in the rate of spontaneous onset of labor in the two groups. Moreover, there was no significant difference in the rate of labor within 72 hours after the intervention. In the present study, the oxytocin receptors appeared to have not affected the above results; however, this outcome should be confirmed by re-examination with an appropriate sample size. Another possible reason why this hypothesis is not supported aside from the sample size is the gestational weeks when the intervention was started. For women in their late gestational weeks, oxytocin receptor

mRNA protein levels have been found to increase (Fuchs et al, 1984; Kimura et al., 1996; Liedman, Hansson, Igidbashian, & Akerlund, 2009). In addition, it has been shown that an increase in oxytocin levels and the number of oxytocin receptors are correlated with an increase in response to oxytocin (Fuchs et al., 1983) and have been shown to have a better response to synthetic oxytocin (Satin, Leveno, Sherman, & McIntire, 1992). However, the number of gestational weeks when the breast stimulation intervention was started varies, namely, at the completion of 37 weeks (Damania et al., 1992), 38 weeks (Salmon et al., 1986; Singh et al., 2014), and 39 weeks (Adewole et al., 1993; Elliott & Flaherty, 1984). Thus, the present study cannot be considered as a premature intervention compared with these previous studies.

The rate of *adequate uterine contraction* was significantly increased in the intervention group, and it was confirmed that the breast stimulation had a uterine contraction effect similar to that in a previous study (Christensson et al., 1989). In the intervention group, the mean frequencies of uterine contraction before the intervention on the first and third days were not significantly different, however the frequency of contractions increased significantly up to 30 minutes from the start of intervention. Importantly, this is apparently the first study describing the frequency of uterine contraction during a long period of breast stimulation for labor onset. Regarding mechanical uterine stretch, this is considered to be an important factor in the up-regulation of oxytocin receptor mRNA (Terzidou et al., 2005). Mechanical stimulation to the cervix also promotes prostaglandin secretion and contributes to the cervical ripening mechanism. Previous studies have reported a significant ripening of the cervix as indicated by the Bishop score as a result of breast stimulation for three days (Adewole et al., 1993; Salmon et al., 1986). Future studies involving the measurement of the Bishop score are warranted.

In the present study, the peak frequency of uterine contraction was observed at 30 minutes, which coincided with the peak oxytocin level. (Christensson et al., 1989) showed that the plasma oxytocin level increased with time during the uterine contraction as a result of breast stimulation for 30 minutes. The infusion of synthetic oxytocin is reported to achieve a maximal uterine contraction response at 40 minutes

from the start of the intervention (Crall & Mattison, 1991; Seitchik, Amico, Robinson, & Castillo, 1984). Furthermore, when the infusion dose of oxytocin was not increased, the uterine contraction tended to decrease (Crall et al., 1991).

On the first day of intervention in the present study, after showing a maximal uterine contraction response due to the increase of the amount of oxytocin level by breast stimulation, it is possible that the frequency of the uterine contractions may decrease with decreasing levels of endogenous oxytocin. However, despite the significant increase in the frequency of uterine contractions (the same as on the first day), it was observed that on the third day of the intervention, there was no increase in the oxytocin level during the intervention as compared to the first day. Therefore, the results of uterine contraction response and oxytocin changes were not consistent.

Feldman et al. (2012) reported that the GG type of rs2254298 is related to a significantly low plasma oxytocin level. However, in this study, that was not confirmed. In terms of clinical outcomes, the uterine contraction frequency was higher rather than lower in the A allele group. Also, in this study, rs53576 did not show any significant association with outcomes in either analyses. Both SNPs were different from previous studies.

## Feasibility

### Acceptability of intervention method.

There are studies using breast stimulation for augmentation labor and one study in the U.S. reported that 21 staff (43%) determined it was embarrassing for staff, patients, and their families (Curtis, Resnick, Evens, & Thompson, 1999). However, there have been no studies reporting the acceptability of pregnant women themselves. In this study, about 90% of primiparas women who touched their nipples during pregnancy indicated that this method was not uncomfortable. Moreover, in the 17 women who participated in all the experiments for three days, the burden of fatigue and pain from the breast stimulation had decreased on the third day than on the first day. About 96% of the women stated that they would like to perform breast stimulation in their next pregnancy, and 88% of the women answered that they will recommend the method to

their friends. The study protocol received a favorable evaluation, with a high satisfaction level for the breast stimulation. It is important to know preferences and cultural difference for women; breast stimulation was shown to be a feasible method for Japanese women.

### Practicality of experimental methods.

#### *Adverse events.*

There were no adverse events during delivery reported in this study, which was the same result as several previous studies (Adewole et al., 1993; Elliott & Flaherty, 1984; Salmon et al., 1986).

During the experiment, one woman confirmed painless uterine tachysystole with prolonged FHR deceleration. After 30 minutes of monitoring, the obstetrician allowed the resumption of the experiment. However, the tachysystole of the woman started only two minutes after starting the breast stimulation, thus the experiment was not resumed. FHR tracing shows the acid-base status of the fetus at the time of observation (Macones et al., 2008). Prolonged deceleration is classified as Category II FHR tracings. The tracings in this category are insufficient to judge as predictive of an abnormal acid-base status and are observed daily in routine practice (Macones et al., 2008).

In this study, all participants were low-risk pregnant women who received regular prenatal checkup and had cardiotocography for FHR at least 30 minutes before the intervention. Elliott & Flaherty (1984) recommended daily fetal movement counts at home and the breast stimulation stress test before intervention at home in questionable cases. As just described, it is necessary to pay sufficient attention to safety especially for the first time of intervention.

#### *Saliva collection and assay.*

Of the 256 saliva samples collected in the three-day experiment, oxytocin levels measured at eight points were successfully collected in 192 (75.0%) of the samples. Since there are many sampling points in this study, there is a possibility that some saliva samples were not properly collected. When shortening the intervention time in the

future, by measuring only before and after the intervention, there is a possibility that the number of samples that can provide an adequate saliva amount can be increased. Another possible improvement method, is by inserting the extraction step, whereby the oxytocin analysis rate exceeds 90% even in double measurement with 1 ml saliva volume (Grewen et al., 2010).

In this research, the extraction step was not carried out. Thereby, there is a possibility that individual variations may be largely reflected. In the measurement of salivary oxytocin level, the extraction method is recommended because the non-extraction method shows variations in results compared with extraction method (McCullough, Churchland, & Mendez, 2013; Szeto et al., 2011). However, the extraction method is complicated in terms of the detailed procedure. At present, ELISA with the non-extraction method is widely used. However, if the sample size is small owing to difficulties in recruiting appropriate subjects such as in the present study, oxytocin extraction may reduce the effects of individual variations and may result in a more robust data.

In the present study, 72.7% of the specimens could be measured in duplicates. To ensure statistical validity, having duplicate or triplicate measurements in any ELISA run is the norm. In the future, measures must be considered to make this possible.

**Limitations of the Study and Suggestions for Future Studies**

In all samples, oxytocin levels were significantly higher in the intervention group than in the control group at baseline on the first day. The difference between the two groups before the intervention in this study may be that women had higher OT at baseline at the time of recruitment. There is a possibility that a randomized comparison test can resolve this problem. In addition, a questionnaire could be conducted to check if oxytocin tends to be higher.

The limitations of this study were the limited number of participants. In the previous study (Takahata et al., 2016), only one out of 16 women (6.3%) was unable to continue intervention due to delivery. However, in this study, nine out of 42 women (21.4%) dropped out after the intervention started due to birth. This study was to

examine the effect on oxytocin using an effective long hour breast stimulation. In the previous study, it was considered highly probable that labor would come from carrying out this intervention (Kavanagh et al., 2005). So, considering the fetal and maternal outcome the timing of the intervention cannot be carried out any earlier than was done in the present study. Therefore, when conducting intervention studies at the end of pregnancy, it is required to design a sample size in consideration of about 20% dropout due to parturition. In future randomized control studies, a large sample size and measurement of the Bishop score are warranted to fully clarify the effects of breast stimulation on the oxytocin level in relation to labor induction, onset, and augmentation. In addition, it was suggested that repetitive stimulations might lead to an increase of oxytocin rather than longer stimulation at one time as in this study. Also, it is desirable to set up a comparative study in which the rate of labor onset is set as primary outcome and multiple intervention groups that perform short breast stimulation are set up.

## Conclusion

This study aimed to examine the effect on salivary oxytocin level following breast stimulation in an intervention group (breast stimulation) and a control group (usual care) of low-risk primiparas. Following one hour of breast stimulation for three days, the mean oxytocin level in the intervention group was significantly higher at 30 minutes after the intervention on the third day than in the control group. Repeated breast stimulation for spontaneous onset of labor led to the increase in oxytocin levels. The rate of adequate uterine contraction and the frequency of uterine contraction during the breast stimulation were significantly higher in the intervention group than in the control group on the first and third days. The feasibility of protocol was confirmed in terms of acceptability. In terms of feasibility practicality, improvements of the oxytocin measurement points and assay methods are warranted.