

# Agreement between maternal recall of distant first birth events with hospital birth records: A cohort study

Mindy Ebrahimoff (✉ [mindyebb@gmail.com](mailto:mindyebb@gmail.com))

---

## Research article

**Keywords:** Maternal recall, birth events, hospital records, agreement, nulliparous daughters

**Posted Date:** October 5th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-951334/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**BACKGROUND** Inter and intra-generational birth cohorts could be particularly useful for predicting the likelihood of labour and birth events for nulliparous women. However, maternal recall of their first childbirth may be imprecise, and hospital records can be inaccurate. Establishing the extent of agreement between mothers' recall and hospital reports of historical first birth events could be the basis of a prediction tool that could contribute to better health care practices during daughter's perinatal period.

**METHODS** In 2015, women who had their first baby between 1967 and 1997 were asked to recall gravidity, method of labour onset, type of pain relief, length of labour, birth outcome, and infant's gender, birthweight and gestational age  $\geq 17$  years postpartum. Responses were compared to hospital birth records. Agreement was evaluated using Bland-Altman's plots and Kappa statistics ( $k$ ). Logistic regression modeling was used to determine factors influencing discrepant recall.

**RESULTS** Of 150 questionnaires distributed, 101 records were complete. Up to 49 years after birth there was strong agreement for birthweight measured at interval (mean discrepancy  $-28.69g$ ,  $SD = 170.91g$ , Bland-Altman 95% limits of agreement  $(-363.66g, 306.28g)$ ) and category level birthweight  $k=0.83$ , good agreement for gestational age (GA) in weeks, at interval level (mean difference  $=0$ ,  $SD = 0.90$ , Bland-Altman 95% limits of agreement  $(-1.766, 1.766)$ ) and at category level GA  $k=0.56$ . There was moderate agreement for labour length ( $\leq 10hrs / >10hrs$ )  $k=0.54$ ; 43% of records did not record this information. For gravidity  $k=0.43$ , labour onset  $k=0.79$ ; any pain relief  $k=0.61$ ; and birth outcome  $k=0.91$ . Univariate logistic regression showed better agreement on infant birthweight in women with higher levels of education, lower agreement for onset of labour method with increasing maternal age at birth, and higher agreement for use of pethidine, but lower agreement for use of epidural in women who had their first babies more recently.

**CONCLUSIONS** Mothers accounts of first birth events generally agree with hospital records. Familial birth data may contribute to more individualised care for nulliparous women, and may limit rising interventions based on population level guidelines. Future research in other settings is warranted before diagnostic criteria may be used in clinical settings.

## Background

Research on familial factors influencing reproductive health depends on good quality data about birth events. Women who are currently pregnant may act upon familial birth information for more individualized care. Reproductive histories show babies of the same mother tend to have similar weight for gestation<sup>[1]</sup>, small for gestational age (GA) babies show intergenerational recurrence<sup>[2]</sup>, and within-family labour and birth characteristics may predict labour length for first time mothers<sup>[3]</sup>. Intrapartum events are often obtained by maternal recall. Agreement analysis between recall and recorded birth data is essential so that birth histories can be based on robust evidence.

Previously published data on maternal recall and recorded birth data can be divided into three time

intervals. Ten studies cover the period up to one year after birth (mean period since birth 20.66 weeks)<sup>[4-13]</sup>. Eight studies span more than one year to less than 10 years since birth (mean period since birth 6 years)<sup>[14-21]</sup>. Eleven studies include women who gave birth more than ten years ago (mean 22.59 years)<sup>[22-32]</sup>. The longest maternally recalled time-span for infant birthweight was 35-70 years after delivery<sup>[30]</sup>. Despite the number of studies suggesting that agreement between hospital records and maternal recall of birthweight and/or GA, and/or mode of birth is sufficiently reliable (summary range for birthweight  $k=0.71-1$ <sup>[17,18]</sup>; for GA  $k=0.6-0.85$ <sup>[11,16]</sup>; for mode of birth  $k=0.80-1.00$ <sup>[15,24,29]</sup>), there is evidence that maternal characteristics, such as education<sup>[5,8,9-13,15,17,18,20,23,29,30]</sup>, socioeconomic status<sup>[5,9,10,12,13,16,18,20,22,25,27,28,30]</sup>, ethnicity<sup>[5,8,9,10,16,17,20,22,25,28,30]</sup>, and gravidity/parity<sup>[19,22,29]</sup> may impact on strength of agreement, with disparities reported across studies. There has been one systematic review and meta-analysis of the validity of recalled and recorded birthweight<sup>[33]</sup>; high agreement was found (range 86-129g,  $n=29,293$ ). Maternal recall of gravidity and parity were only investigated in the  $\geq 10$  years from delivery time-frame (range  $k=0.74-0.98$ )<sup>[22,29]</sup>. Use of analgesia and labour length have been investigated less frequently across all time-frames with kappa coefficients for analgesia agreement ranging from 0.58-0.85<sup>[4,13,15,29,32]</sup>; and for labour length agreement from  $k=0.21$  to 71% agreement<sup>[8,11,18,26,29,32]</sup>. As far as we are aware, there have been no studies conducted comparing maternal recall agreement with birth records for method of labour onset. Some studies assume that hospital records are accurate<sup>[18]</sup>, however, there is evidence that data are not always correctly recorded in hospital records<sup>[11]</sup>. This study does not assume that the gold standard is the hospital account. Most studies have focused on one or a few variables. Only Buka et al. (2004)<sup>[29]</sup> analysed six recalled and recorded birth outcomes in 96 women, 22 years later and found that age at birth and parity were accurately reported, more educated women generally recalled events more accurately than less educated women, and women with major medical events such as cesarean section ( $k=1$ ), were more likely to recall these events in line with what was recorded in hospital records. This study examines agreement between maternal recall and hospital birth records for 8 intrapartum variables, including labour onset which has not been addressed previously. Mothers were asked to complete a questionnaire on average 33 years after delivery reporting on: gravidity, length of labour, pain relief, birthweight, GA, mode of delivery, infant gender and labour onset. Reasons for discordance in agreement were identified.

## Methods

Women who were accompanying family members or friends attending antenatal clinics in either of two Israeli maternity hospitals between 2015-2016 were recruited by the attending midwife. Eligible women were those who had given birth to their first child in that hospital more than 17 years ago (to exclude very young women who may have health consequences such as pre-term birth and lower birthweight). Questionnaires, participant information sheets and consent forms were provided. Completed questionnaires were returned at ante-natal visits or via Freepost. Project approval was granted by the ethics committees of the Sourasky Medical Centre, ref:0039-14,12.06.2014; Ma'ayney Hayeshua Hospital,

ref:72.14,30.07.2014; and University of Central Lancashire Research Ethics Committee, UK; ref:STEMH 255,09.09.2014.

Questionnaires included name, telephone number, national identity number (for tracking records), demographic information, general and obstetric health histories, and prenatal and perinatal information. Maternal history items included mother's first birth age, delivery date, gravidity, weight gain during pregnancy, duration of pregnancy, signs of labour, length of labour, use of pain relief medication, birth outcome, GA at birth, infant gender, birthweight, and Apgar of newborn. Responses were either categorical or continuous. Length of labour was measured on a categorical scale with four time intervals ( $\leq 2$  hours,  $>2-6$  hours,  $>6-10$  hours,  $>10$  hours) based on the findings of a systematic review which identified the mean active labour duration for 7,009 nulliparous women as 6 hrs  $\pm$  SD 3.5 hrs<sup>[34]</sup>.

Data were anonymized, numbered and assessed manually for errors. Missing data or ambiguous values were queried by telephone conversation. Recalled first birth events were compared for agreement with data in hospital records. A pilot study (n=10) confirmed that recruitment procedures, questionnaire use and return of data were satisfactory. Pilot study results (n=9) were included in the main results.

## Statistical analysis

Sample characteristics were described using descriptive statistics,

means and standard deviations (*SD*), or medians and interquartile range (IQR) for continuous variables, and frequencies and percentages for categorical variables. Multiple births were included in the analysis with each child treated as a single unit. Bland-Altman<sup>[35]</sup> plots were used to determine agreement between maternal recall and medical records for infant birthweight and GA as continuous variables. To measure agreement for birthweight and GA as categorical variables optimal cut-offs for clinical significance were applied: low ( $\leq 2,499$ g), normal (2,500-3,999g) and high ( $\geq 4,000$ g); and preterm ( $\leq 37$  weeks), term (37-40 weeks) and post-term ( $\geq 41$  weeks). Mothers length of labour was dichotomized ( $\leq 10$ hrs/ $>10$ hrs). Participants were asked to report on labour length within a time category because poor specificity over time is common for women reporting time-bound peripartum events<sup>[36]</sup>. Kappa statistics together with 95% confidence intervals (CI) was calculated for categorical and dichotomised variables. Strength of agreement was classified according to Landis and Koch (1977)<sup>[37]</sup> with kappa values of  $\geq 0.8$  indicating 'excellent', 0.61–0.8 'substantial' and 0.41–0.6 'moderate' agreement.

Logistic regression analysis was used to predict discrepant recall of birth events. For continuous and categorical variables, the corresponding binary variables were created indicating inaccurate or accurate recall (coded as 0 and 1 respectively). The impact of independent variables on maternal recall is presented as odds ratios (OR) with 95% CI and *p* values. Variables that had no effect were deleted from the regression analysis in a stepwise manner. Recalled and categorical independent variables were also

analysed using chi-squared test. Statistical significance was defined as  $p < 0.05$ . Data were analysed with SPSS version 24.0 (Armonk, NY:IBM Corp.).

## Results

The number of questionnaires distributed was set by the available women in the time period of the study. Of the 150 questionnaires distributed (10 in pilot study, 140 in main study), 121 completed questionnaires were returned (81%). Of these, 20 were excluded from the analysis, due to missing hospital birth records (n=14), and having a birth less than 17 years prior to the study (n=6), making a final study sample size of 101 (Figure 1).

Demographic characteristics of participants are presented in Table 1.

**Table 1: Demographic Characteristics of Study Participants Based on Questionnaire Data (Maternal Recall)**

<b>Variable</b>	<b>-</b>	<b>Participants</b>	
Age	Total sample size (N)	101	
	Mean ( <i>SD</i> )	55.17 (7.37)	
Age at first delivery	Number of usable replies (N)	101	
	Mean ( <i>SD</i> )	24.18 (4.53)	
	Median (IQR)	23.0 (5)	
Years elapsed since first birth	Number of usable replies (N)	101	
	Mean ( <i>SD</i> )	30.99 (7.88)	
Country of origin	Number of usable replies (N)	99	%
	Israel	76	75.2
	Europe	15	14.9
	Asia/Africa	5	5.0
	Americas		
	Missing	3	3.0
		2	2.0
Mother's origin	Number of usable replies (N)	97	%
	Israel	24	23.8
	Asia/Africa	33	32.6
	Europe	35	34.6
	Americas	5	5.0
	Missing	4	4.0
Father's origin	Number of usable replies (N)	97	%
	Israel	21	20.8
	Asia/Africa	29	28.7
	Europe	43	32.6
	Americas	4	4.0
	Missing	4	4.0
Birth hospital	Number of usable replies (N)	101	%
	†Kiriya	90	89.1
	Ma'aynei	7	6.9
	†Lis	4	4.0
	-		
Education	Number of usable replies (N)	97	%
	Full high school	30	29.7
	Vocational school	28	27.7
	University degree	28	27.7
	Partial high school	11	10.9
	Missing	4	4.0
Family status	Number of usable replies (N)	101	%
	Married	84	83.1
	Divorced	10	9.9
	Widowed	5	5.0
	Other	2	2.0
Religious status	Number of usable replies (N))	98	%
	Secular	44	43.6
	Traditional	31	30.7
	Ultra-orthodox	16	15.8
	Religious	7	6.9

Missing

3

3.0

† Kiriya=Lis

Participants gave birth to their first child between 1967 and 1998, were currently aged 37-70 years, were aged 18-41 years at first delivery, and the years elapsed since first birth was 17-49 years. Most participants were educated to high school level (30%). Most women were Israeli (75%), secular (44%) and married (83%).

Recalled and recorded perinatal and new-born outcome information is presented in Table 2. (see end of manuscript for Table 2)

Multiple births included one set of triplets and three sets of twins giving a total of n=106 infants.

We reiterate, there is no single gold standard for agreement status to use as the reference. Bland-Altman plots for agreement between recalled and recorded continuous infant birthweight and GA showed small differences for birthweight records with a mean difference of -28.69g, *SD*=170.91, and 95% limits of agreement (-363.7, 306.3). There was an observed tendency towards lower maternal estimation than recorded data among normal birthweight infants of between 2700–3500g (Figure 2).

Gestational age agreement was within an acceptable range. Whilst the estimated mean difference was equal to 0, the 95% limits of agreement were relatively wide,  $\pm 1.77$  weeks. There was an observed trend towards mothers reporting a lower GA where the records reported 39 and 41 weeks of gestation, and a trend towards maternal reports of higher GA for those recorded as over 41 weeks' gestation (Figure 3).

When comparing data from medical records and mothers recall using the categories low ( $\leq 2,499$ g), normal (2,500-3,999g) and high ( $\geq 4,000$ g) for birthweight, and GA groups of preterm ( $\leq 36$  weeks), term (37-40 weeks) and post-term ( $\geq 41$  weeks) it was found that mothers of low and high birthweight babies tended to recall them as being smaller or larger respectively than recorded in the hospital records. Good agreement between records and maternal recall for babies of normal birthweight was found (Table 3).

**Table 3: Distribution of agreement of Maternal Recall and Hospital Birth Records of Infant Birth Weight**

Infant Birth Weight (g)	IBW error* (g), mean (SD)	n
$\leq 2,499$	136.15 (186.84)	12
2,500-3,999	71.73 (141.32)	84
$\geq 4,000$	184.00 (214.55)	2
Total	85.86 (153.13)	98

\*Compared with hospital birth records

Similarly, mothers of preterm and post term babies reported lower and higher birthweights respectively when compared to the hospital records, whereas length of gestation for term babies had much closer agreement between maternal recall and the records (Table 4).

**Table 4: Distribution of agreement of Maternal Recall and Hospital Birth Records of Gestational Age**

Gestational age (weeks)	GA error* (w)	mean (SD)	n
≤36	0.8	(0.98)	7
37-40	0.49	(1.03)	56
≥41	0.93	(0.51)	11
Total	0.6	(1.82)	74

\*Compared with hospital birth records

Agreement between recall and medical records for five categorical variables and five dichotomous variables are shown in Table 5.

**Table 5. Agreement between Maternal Recall and Hospital Birth Records for Categorical and Dichotomous variables**

Categorical Variable	Kappa	Strength of agreement*	<i>p</i>	95% CI for kappa
Gestational age	0.563	Moderate	0.001	0.342, 0.784
Infant birth weight	0.830	Almost perfect	0.001	0.641, 0.989
Onset of labour	0.790	Substantial	0.001	0.601, 0.979
** Any pain relief	0.618	Substantial	0.001	0.477, 0.759
Mode of delivery	0.919	Almost perfect	0.001	0.842, 0.995

Dichotomous Variable	Kappa	Strength of agreement*	<i>p</i>	95% CI for kappa
Length of labour ≤10hrs/>10 hrs	0.536	Moderate	0.001	0.305, 0.767
Gravidity	0.437	Moderate	0.001	0.176, 0.698
Induction	0.758	Substantial	0.001	0.556 - 0.960
Epidural	0.759	Substantial	0.001	0.632, 0.886
Pethidine	0.224	Fair	0.024	-0.035, 0.448

For categorical variables infant birthweight and mode of delivery almost perfect agreement was found with an exact match in 42% of women. For infant birthweight 30%, 15% and 7% of the women had up to 100g difference, between 101 to 300g difference and >300g difference respectively. Categorised GA had the lowest level of agreement. An exact match between recall and records was found for GA in 35% of women, with 29% and 6% of women reporting a one-week difference and two-week difference from their records respectively. Some 30% of women did not report on this parameter. For the majority of these, data were available in their records. For method of onset of labour and type of pain relief, an exact match was found in 61% and 57% of women respectively.

Agreement results for dichotomous variables show moderate agreement for length of labour ≤10hrs/>10hrs, however, 43% of data for labour onset was missing from birth records. Agreement between women's accounts and records for the number of previous pregnancies was 95%. Fifteen women recalled having had a previous pregnancy while birth records only reported on 9 women with a previous pregnancy. The nature of the pregnancy loss is not specified in the records. Having labour induced and having an epidural (yes/no) were both found to be substantially in agreement between recall and records. Of the 46 women who recalled receiving an epidural, 42 had this recorded in their hospital records. Only fair agreement between recall and records was found for use of pethidine.

Logistic regression analysis showed statistically significant predictors for inaccurately recalled variables (Table 6).

**Table 6: Univariate Logistic Regression for the Prediction of Disagreement between Maternal Recall and Hospital Records**

<b>*Dependant variable</b>	<b>Independent variable</b>	<b>B</b>	<b>OR</b>	<b>95% CI</b>	<b>p</b>
<b>Recall of IBW (± 100g)</b>	Education (ref.: academic)				0.10
	Partial ed.	1.79	6.0	1.18-30.62	0.031
	Full high school	0.88	2.4	0.63-9.16	0.20
	Higher ed.	0.04	1.04	0.23-4.67	0.96
<b>Recall of onset of labour</b>	Age of mother	0.132	1.14	1.01-1.29	0.038
<b>Recall of use of epidural</b>	Years since delivery	-0.13	0.88	0.79-0.97	0.013
<b>Recall of use of pethidine</b>	Years since delivery	0.12	1.13	1.04-1.22	0.002

\* DV = 1 corresponds to inaccurate recall  
*p* < 0.05 is a significant finding.

The pain relief variables (epidural and pethidine) were influenced by time elapsed since delivery. There was less recall-record agreement for use of pethidine for women who had more years elapsed since birth. However, in the case of epidural analgesia, women who had *less* years elapsed since birth showed *less* agreement between recalled and recorded epidural use than women who had more years elapsed since birth. Women who were older at the time of their first birth showed less agreement between recalled and recorded type of onset of labour than those who were younger women.

Finally, high school vs higher education was a statistically significant factor associated with higher differences between recall and records (*p*=0.029). For women with education up to high school level, the odds of disagreement between recalled and recorded infant birthweight by more than 100g were higher (OR=3.1, 95% CI (1.1,8.4)) compared to women with higher education. Analysis of ordinal education variables with four categories, as presented in Table 4, shows statistically significant linear-by-linear association test for trends between education and inaccurate recall (*p*=0.024).

## Discussion

Our data suggest that long-term recall of most perinatal events results in close agreement with hospital records up to 41 years later. Agreement for mode of birth and infant birthweight was generally higher than the other variables studied, consistent with other studies researching agreement ≥10 years from birth for mode of birth<sup>[24,29,32]</sup> and infant birthweight<sup>[23-30,32]</sup>. Furthermore, the agreement level for these two variables was stable even when adjusted for time elapsed since birth, though recall of birthweight had better agreement for women with higher education levels, and recall for mode of birth had better agreement for women who were younger at time of delivery. High agreement between maternal recall of infant birthweight and birth records may reflect repetition of information to others after delivery<sup>[24]</sup>. The association of lower levels of agreement for women with less years of education is consistent with

findings from five other studies<sup>[5,15,18,29,30]</sup>. However, the actual difference was under 300g for 93% of women, which is likely to be of little clinical relevance. Consistent with one other study<sup>[20]</sup>, Bland-Altman plot showed good agreement for GA as continuous data while agreement was found to be moderate as expressed as a kappa statistic after categorizing these data. This difference may be explained by the categorisation of GA for kappa analysis, although the discrepancies between recalled and documented information were less than one week in all groups, which may be of little clinical significance, unless decisions are being made about induction for post-maturity. In terms of the number of previous pregnancies, 95% of women's accounts and records agreed. Due to the delicacy of the data, information about abortions or a termination of pregnancy may have been withheld.

Only three other studies report recall of length of labour more than ten years since first birth<sup>[26,29,32]</sup>. Our findings of moderate agreement for labour length are in accord with these findings. The fact that nearly half of all records did not report on labour onset (and thus labour length) was unexpected, the consequences of which are seen in the large discrepancy between recall and actual events for this variable. This is a critical omission, as length of labour is an important trigger for routine labour interventions<sup>[38]</sup>. Moreover, it limited our ability to assess the level of agreement with precision. There is currently widespread controversy over the nature and limits of physiological labour length. This debate, focusing specifically on the effects of vaginal birth or interventional delivery on rising cesarean rates is based on epidemiological analysis of labour progress patterns in hospital records<sup>[39]</sup>. If these data are inaccurate, or missing, then future policies and guidelines may also be inaccurate.

To the best of our knowledge, this is the first study that has investigated maternal recall of all modes of labour onset (including natural, pharmacological medication and mechanical or physical approaches). Limited obstetric information on labour onset may weaken rationale for induction, augmentation and cesarean birth for slowly progressing labours. Given the current global burden of rising birth interventions, this is a critical gap in the literature.

Two studies looked at induction of labour (rather than mode of onset). Elkadry et al, (2003)<sup>[8]</sup> interviewed 277 women 10 weeks after birth and found that one of 8 mothers did not know whether her labour had been induced, and of 108 women whose labour had been induced, 1/10 women could not recall the indication that was stated in the records. Bat-Erdene et al., (2013)<sup>[13]</sup> interviewed 755 women four months' post-partum and found that induced labour had lower sensitivity and specificity than the other variables studied. In the current study agreement about augmentation with synthetic oxytocin (Pitocin) was low. Only eight women believed they had been given Pitocin in labour, but records noted that this had happened for 17 women. This study is unable to elucidate the reasons for the discrepancy. There is a possibility that the nine women with no recall of the event were never informed that they had been induced. This finding is supported by similar findings in other studies<sup>[26,29]</sup>. It is also possible that older labour records are less accurate for these details.

The amnesic effect of pethidine may explain the underreporting of pethidine usage by mothers across the whole cohort, when compared to hospital records. It is intriguing that women who had their babies more

recently were less likely to agree with their birth records about whether they had an epidural or not. This remains an open research question. It highlights the need for researchers to use caution when calculating rates of epidural usage further complicated by staff routinely under recording epidurals that were administered (in 3% of cases higher levels were recalled by participants than recorded in the notes). Social desirability response bias can be associated with self-report questionnaires, especially in sensitive areas such as childbirth. High pressure work situations may induce staff to make recording mistakes. However, in general, our study showed acceptable agreement between the two types of data. We believe they are a good barometer of long-term maternal recall. Although our sample was small, item non response was rare in the questionnaire unlike other previous studies of long-term maternal recall<sup>[18,24]</sup>, and neither source (recall or records) were used as the gold standard, in recognition of the potential flaws in both.

## **Strengths and limitations**

The setting for this study included populations in which generations of women tended to give birth in the same hospital, meaning that hospital records existed for even those who had their births two generations previously. Potential errors in chart abstraction were minimised by the use of personalised national identity numbers. Limitations include incomplete or missing personal and demographic data within some hospital records, especially for length of labour. Hospital records also tended to round off GA to whole weeks. Some variables were uncommon in the sample.

## **Conclusions**

This study shows that clinicians may use perinatal information reported by mothers in the absence of medical records, even when this information is required many decades after birth. Further work is needed on how records report labour onset and length, and how women report it, to ensure that what is described in clinical records, and what women report in birth histories, is captured and harmonised. Further studies on agreement data of a broad range of birth events from a range of settings is warranted to determine the replicability of our findings.

## **Abbreviations**

*GA*: gestational age

*L&D*: labour and delivery rooms.

*IBW*: infant birth weight

*CS*: Caesarean section

*PROM*: premature rupture of membranes

*PIH*: pregnancy induced hypertension

*AROM*: artificial rupture of membranes

## Table 2

Table 2: Distribution of Paired Data on Recalled and Recorded Intrapartum Data

		<b>Mother's Recall N</b>	<b>% (out of the no. of valid responses)</b>	<b>Hospital Records N</b>	<b>% (out of the no. of valid responses)</b>
<b>IBW in grams (including multiple deliveries)</b>	Low ≤ 2,499	16	15.2	12	11.9
	Normal 2,500-3,999	84	80.0	87	86.1
	High ≥ 4,000	5	4.8	2	2.0
	Missing	1		5	
		(0.9%)		(4.7%)	
<b>GA in weeks (including multiple deliveries)</b>	Pre-term ≤ 36w	11	10.7	7	9.5
	Term 37- 40w	75	72.8	56	75.7
	Postdate ≥ 41w	17	16.5	11	14.8
	Missing	3		32	
	(2.8%)		(30.2%)		
<b>Gravidity</b>	1	86	85.1	92	91.1
	2			6	5.9
	3			2	2.0
	4			1	1.0
	More than 1 (memory only)	15	14.9		
<b>Onset of labour</b>	Contractions	54	53.4	35	48.6
	Rupture of membranes	31	30.7	26	36.1
	Vaginal bleeding	2	2.0	8	11.1
	Induction	11	10.9	3	4.2
	Elective CS	2	2.0		
	Other	1	1.0		
	Missing	-		29	
			(28.7%)		
<b>Reason for induction</b>	Post date	4	4.0	3	3.0
	PROM	2	2.0	5	5.0
	*Failure to progress-	8	8.0	17	17.0
	augmented	1	1.0	1	1.0
	Oligohydramnion	1	1.0	1	1.0
	Meconium	1	1.0		
	PIH/Toxemia	84	83.0	73	72.0
	No reason given/				
	No induction				
<b>Rupture of membranes</b>	Spontaneous	17	70.8	14	25.5
	AROM	-	-	22	40.0
	PROM	5	20.8	16	29.1
	Elective CS	2	8.4	3	5.4
	Missing data	77		46	
	(76.2%)		(45.5%)		
<b>Anaesthesia</b>	Epidural	46	45.9	42	55.3
	Pethidine	12	12.2	13	17.1
	Spinal	2	2.0	1	1.3
	No recall	3	3.1		
	None	33	33.8	17	22.4
	General	2	2.0	3	3.9
	Missing data	3		25	
	(3%)		(24.8%)		

<b>Length of labour in hours</b>	≤ 2	6	6.3	1	1.7
	>2 - 6	34	35.8	15	25.9
	>6 - 10	16	16.8	29	50.0
	>10	39	41.1	13	22.4
	Missing data	6 (5.9%)		43 (42.6%)	
<b>Birth outcome</b>	Spontaneous vaginal	69	68.3	67	67.0
	Vacuum/Forceps	18	17.8	19	19.0
	Elective CS	2	2.0	4	4.0
	Emergency CS	12	11.9	10	10.0
	Missing data			1 (1.0%)	
<b>Presentation (including multiple deliveries)</b>	Cephalic	97	96.0	88	96.7
	Breech	3	3.0	3	3.3
	Footling	1	1.0		
	Missing data	5 (4.7%)		15 (14.2%)	
<b>Gender (including multiple deliveries)</b>	Male	43	41.3	44	41.6
	Female	61	58.7	62	58.4
	Missing data	2 (1.9%)			
<b>APGAR (including multiple deliveries)</b>	Normal	92	96.8	80	100.0
	Abnormal	1	1.1		
	Don't know	2	2.1		
	Missing data	11 (10.4%)		26 (24.5%)	
<b>Placental expulsion</b>	Spontaneous	84	84.0	66	82.5
	Manual removal in L&D	3	3.0		
	(revision/lysis)	13	13.0	14	17.5
	Surgically removed (CS)	1		21	
	Missing data	(1.0%)		(20.8%)	

L&D: labour and delivery rooms.

IBW: infant birth weight

CS: Caesarean section

GA: gestational age

PROM: premature rupture of membranes

PIH: pregnancy induced hypertension

AROM: artificial rupture of membranes

\*Cases of augmented labour with Pitocin (oxytocin) were excluded from the induction analysis (kappa).

## References

<sup>1</sup> Skjærven R, Bakketeig LS. Classification of small-for-gestational age births: Weight by gestation standards of second birth conditional on size of the first. *Pediatr. Perinat. Epidermal.* 1989;3:432-447.

<sup>2</sup> Jaquet D, Swaminathan S, Alexander GR, et al. Significant paternal contribution to the risk of small for gestational age. *BJOG: An International Journal of Obstetrics & Gynaecology.* 2005;112:153-159.

- <sup>3</sup> Ebrahimoff M, Many A, Downe S, et al. Length of labour in mothers and their daughters: A matched cohort study. *Eur J Obstet Gynecol Reprod Biol.* 2020; 245:77-83.
- <sup>4</sup> Hewson D, Bennett A. Childbirth research data - medical records or women's reports? *American Journal of Epidemiology.* 1987;125:484-491.
- <sup>5</sup> Gayle HD, Yip R, Frank MJ, et al. Validation of maternally reported birth weights among 46,637 Tennessee WIC program participants. *Public Health Reports.* 1988;103(2):143-147.
- <sup>6</sup> Casey R, Rieckhoff M, Beebe SA, et al. Obstetric and perinatal events: the accuracy of maternal report. *Clinical Pediatrics (Philadelphia).* 1992;31(4):200-204.
- <sup>7</sup> Lederman SA, Paxton A. Maternal reporting of pre-pregnancy weight and birth outcome: consistency and completeness compared with the clinical record. *Maternal & Child Health.* 1998;2:123-126.
- <sup>8</sup> Elkadry E, Kenton K, White P, et al. Do mothers remember key events during labor? *American Journal of Obstetrics & Gynaecology.* 2003;189(1):195-200.
- <sup>9</sup> Tate AR, Dezateux C, Cole TJ, et al. Factors affecting a mother's recall of her baby's birth weight. *International Journal of Epidemiology.* 2005;34:688-695.
- <sup>10</sup> Quigley MA, Hockley C, Davidson LL. Agreement between hospital records and maternal recall of mode of delivery: evidence from 12 391 deliveries in the UK Millennium Cohort Study. *British Journal of Obstetrics and Gynaecology.* 2006;114(2):195-200.
- <sup>11</sup> Troude P, L'Hélias LF, Raison-Boulley AM, et al. Perinatal factors reported by mothers: do they agree with medical records? *European Journal of Epidemiology.* 2008;23(8):557-564.
- <sup>12</sup> Poulsen G, Kurinczuk JJ, Wolke D, et al. Accurate reporting of expected delivery date by mothers 9 months after birth. *Journal of Clinical Epidemiology.* 2011;64(12):1444-1450.
- <sup>13</sup> Bat-Erdene U, Metcalfe A, McDonald SW, et al. Validation of Canadian mothers' recall of events in labour and delivery with electronic health records. *BMC Pregnancy and Childbirth.* 2013;13(suppl 1):53.
- <sup>14</sup> Oates RK, Forrest D. Reliability of mothers' reports of birth data. *Australian Paediatric Journal.* 1984;20:185-186.
- <sup>15</sup> Githens PB, Glass CA, Sloan FA, et al. Maternal recall and medical records: an examination of events during pregnancy, childbirth, and early infancy. *Birth.* 1993;20(3):136-141.
- <sup>16</sup> Olson JE, Shu XO, Ross JA, et al. Medical record validation of maternally reported birth characteristics and pregnancy-related events: a report from the children's cancer group. *American Journal of Epidemiology.* 1997;145:58-67.

- <sup>17</sup> Gofin R, Neumark, YD, Adler B. Birthweight recall by mothers of Israeli children. *Public Health*. 2000;114(3):161-163.
- <sup>18</sup> Rice F, Lewis A, Harold G, et al. Agreement between maternal report and antenatal records for a range of pre and peri-natal factors: the influence of maternal and child characteristics. *Early Human Development*. 2006;83:497-504.
- <sup>19</sup> Sou SC, Chen WJ, Hsieh WS, et al. Severe obstetric complications and birth characteristics in preterm or term delivery were accurately recalled by mothers. *Journal of Clinical Epidemiology*. 2006;59:429-435.
- <sup>20</sup> Adegboye AR, Heitmann B. Accuracy and correlates of maternal recall of birthweight and gestational age. *British Journal of Obstetrics and Gynaecology*. 2008;115(7):886-893.
- <sup>21</sup> Liu J, Tuvblad C, Li L, et al. Medical Record Validation of Maternal Recall of Pregnancy and Birth Events from a Twin Cohort. *Twin Research and Human Genetics*. 2013;3:1-16.
- <sup>22</sup> Tilley BC, Barnes AB, Bergstralh E, et al. A comparison of pregnancy history recall and medical records. Implications for retrospective studies. *American Journal of Epidemiology*. 1985;121:269-281.
- <sup>23</sup> Seidman DS, Slater PE, Ever-Hadani P, et al. Accuracy of mothers' recall of birthweight and gestational age. *British Journal of Obstetrics and Gynaecology*. 1987;94(8):731-735.
- <sup>24</sup> Yawn BP, Suman VJ, Jacobsen SJ. Maternal recall of distant pregnancy events. *Journal of Clinical Epidemiology*. 1998;51:399-405.
- <sup>25</sup> Lumey LH, Stein AD, Ravelli AC. Maternal recall of birthweights of adult children: Validation by hospital and well baby clinic records. *International Journal of Epidemiology*. 1994;23:1006-1012.
- <sup>26</sup> Tomeo CA, Rich-Edwards JW, Michels G, et al. Reproducibility and validity of maternal recall of pregnancy-related events. *Epidemiology*. 1999;10:774-777.
- <sup>27</sup> O'Sullivan JJ, Pearce MS, Parker L. Parental recall of birth weight: how accurate is it? *Archives of Diseases in Childhood*. 2000;82:202-203.
- <sup>28</sup> Walton KA, Murray LJ, Gallagher AM, et al. Parental recall of birthweight: a good proxy for recorded birthweight? *European Journal of Epidemiology*. 2000; 16:793-796.
- <sup>29</sup> Buka SL, Goldstein JM, Sparto E, et al. The retrospective measurement of prenatal and perinatal events: accuracy of maternal recall. *Schizophrenia Research*. 2004;71:417-426.

- <sup>30</sup> Catov JM, Newman AB, Kelsey SF, et al. Accuracy and reliability of maternal recall of infant birth weight among older women. *Annals of Epidemiology* 2005; 16(6):429-431.
- <sup>31</sup> Li CY, Wei JN, Lu TH, et al. Mothers tended to over report categorical infant birth weight of their children. *Journal of Clinical Epidemiology*. 2006;59:1319-1325.
- <sup>32</sup> Hopkins LM, Caughey AB, Brown JS, et al. Concordance of chart abstraction and patient recall of intrapartum variables up to 53 years later. *American Journal of Obstetrics & Gynecology*. 2007;196:3.
- <sup>33</sup> Shenkin SD, Zhang MG, Der G, et al. Validity of recalled v. recorded birth weight: a systematic review and meta-analysis. *Journal of Developmental Origins and Diseases*. 2017;8(2):137-148.
- <sup>34</sup> Neal JL, Lowe NK, Ahijevych K, et al. Active labor duration and dilation rates among low-risk, nulliparous women with spontaneous labor onset: a systematic review. *Journal of Midwifery & Women's Health*. 2010;55:308-318.
- <sup>35</sup> Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1:307-310.
- <sup>36</sup> Blanc AK, Warren C, McCarthy KJ, et al. Assessing the validity of indicators of the quality of maternal and newborn health care in Kenya. *Journal of Global Health*. 2016;6:1–13.
- <sup>37</sup> Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159-174.
- <sup>38</sup> Fox H, Callander E, Lindsay D, et al. Evidence of overuse? Patterns of obstetric interventions during labour and birth among Australian mothers. *BMC Pregnancy Childbirth*. 2019;19:226.
- <sup>39</sup> Boerma T, Ronsmans C, Melesse DY, et al. Global epidemiology of use of and disparities in caesarean sections. *Lancet*. 2018;392(10155):1341-8.

## Declarations

### Ethics approval and consent to participate

Project approval was granted by the ethics committees of the Sourasky Medical Centre, ref:0039-14,12.06.2014; Ma'ayney Hayeshua Hospital, ref:72.14,30.07.2014; and University of Central Lancashire Research Ethics Committee, UK; ref: STEMH 255,09.09.2014. All participants signed an informed consent form.

### Consent for publication

Not applicable

### **Availability of data and material**

The data file is available upon request to the corresponding author after receiving approval from the University of Central Lancashire STEMH ethics committee subject to data sharing agreement. Additional data related to this paper may be requested from the authors.

### **Competing interests**

The authors declare that they have no competing interests.

### **Funding**

ME is a doctoral graduand (2019) UCLan, UK. The research was self-funded.

### **Author's contributions**

ME, VHM, SD, and AM designed the study. ME collected data. ME and ST conducted analyses, were responsible for the analysis and interpretation of data, and for the preparation of the results tables. ME wrote the first draft of the manuscript. VHM, SD, ST and AM contributed to interpretation of results, drafting and revising the manuscript. All authors contributed to final revisions of the manuscript, tables and figures. All authors have read and approved the final manuscript. The corresponding author (ME) attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

### **Acknowledgements**

None

### **Figures**

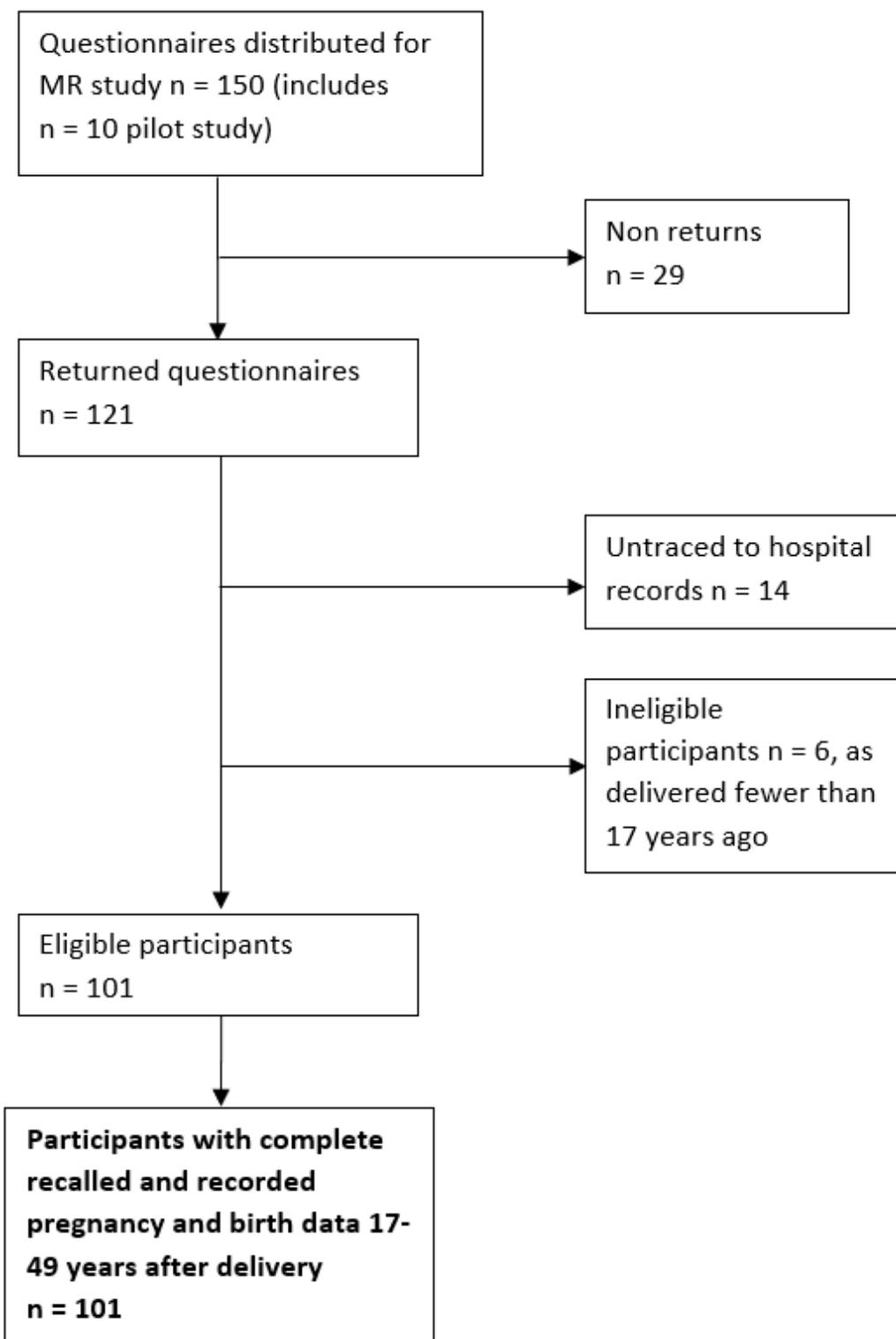


Figure 1

Cohort Profile Maternal Recall Study

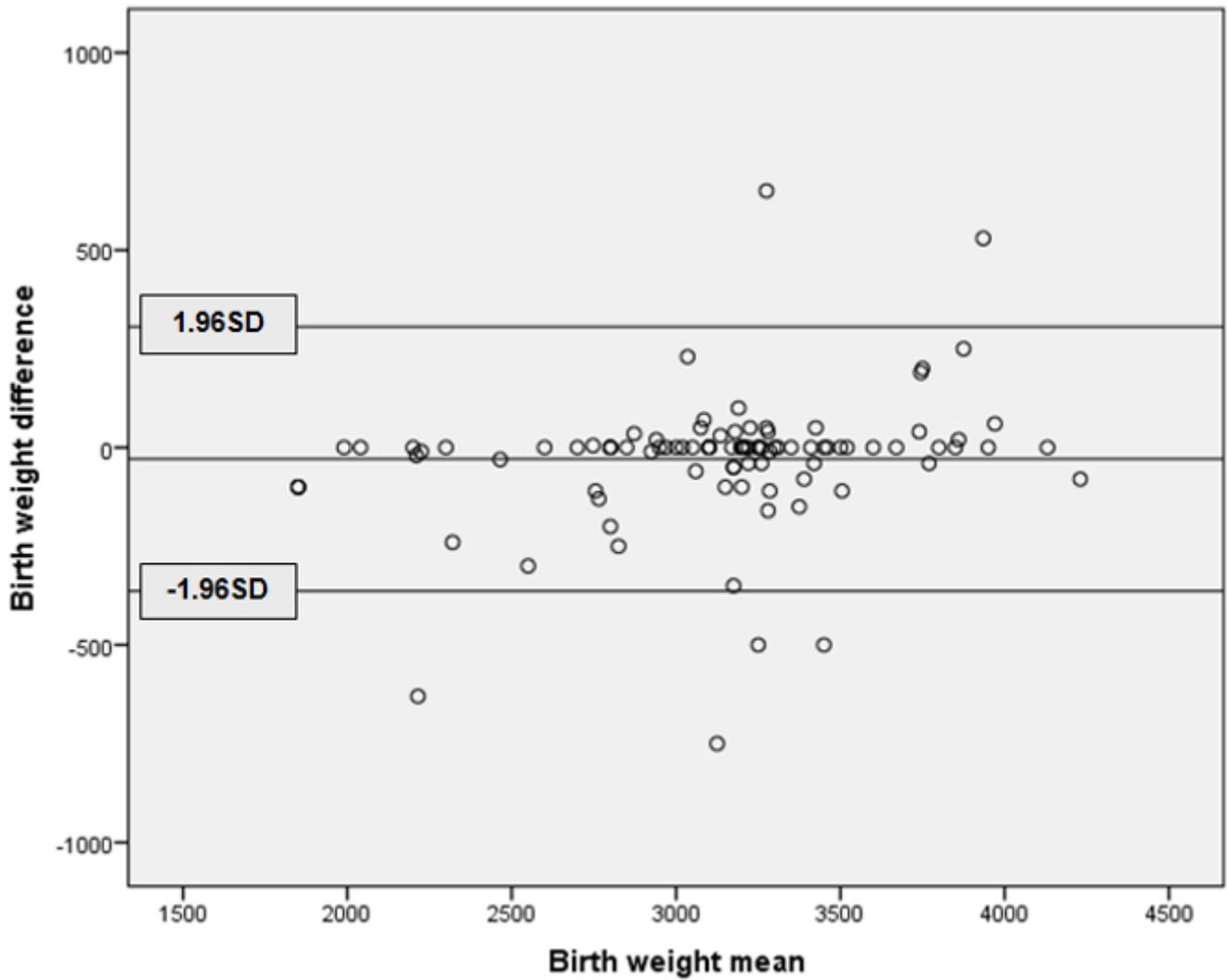


Figure 2

Agreement Between Maternally Recalled and Recorded Infant Birth Weight with 95% Limits of Agreement. Mean difference = -28.69g, and SD = 170.91, 95% limits of agreement (-363.66, 306.28).

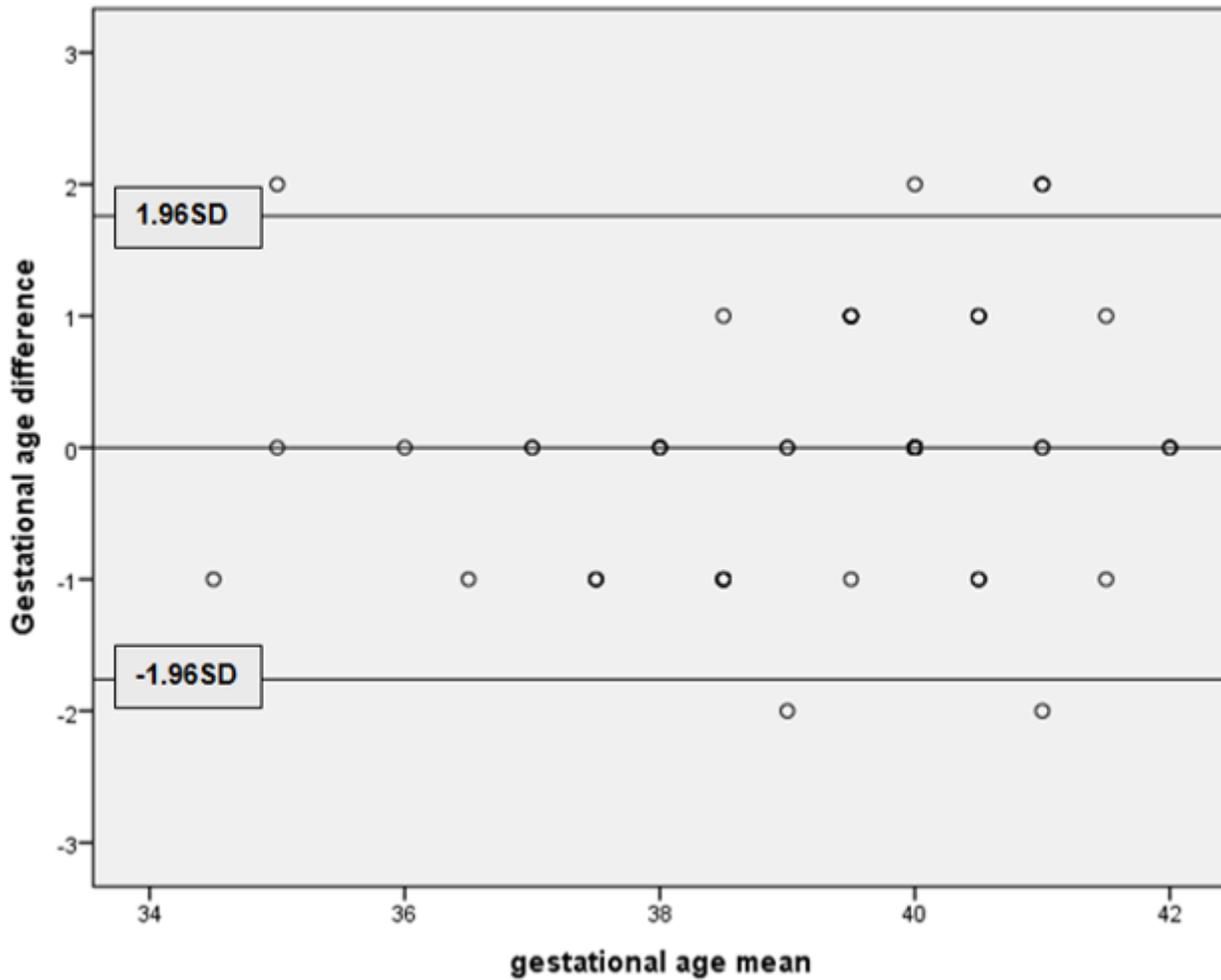


Figure 3

Agreement Between Maternally Recalled and Recorded Gestational Age with 95% Limits of Agreement. Mean difference = 0.0, SD = 0.9, 95% limits of agreement (-1.77, 1.77).