MISSIONE 4
ISTRUZIONE
RICERCA

Department of School of Medicine and Surgery

PhD program: Public Health
Curriculum in Biostatistics

Cycle: XXXVIII

ADDRESSING CONFOUNDING IN MIDWIFERY RESEARCH: FROM OBSERVATIONAL STUDIES TO RANDOMIZED CONTROLLED TRIALS

Candidate: Maria Panzeri

Registration number: 771669

Tutor: Prof. ssa Laura Antolini

Coordinator: Prof. Luigi Badano

Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani
UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

a Pietro e al piccolo uragano Penni

Summary

1. Abstract

This doctoral thesis investigates the use of causal inference methods in midwifery and maternity care research, with the goal of enhancing causal interpretation validity in observational and randomised study designs. In contemporary maternity care research, the estimation of causal effects is often hampered by problems of confounding, selection mechanisms and the complexity of clinical and organisational pathways. Here, causal inference provides a unified conceptual and methodological framework to overcome these limitations by explicitly formalising assumptions about data-generating processes and by distinguishing between association and causation.

The thesis is organised around the central concept of exchangeability, the fundamental condition required for valid causal inference. In observational studies, exposure is not randomly assigned, so exchangeability is usually not guaranteed. This may lead to confounding by baseline prognostic characteristics. In randomised controlled trials, exchangeability is anticipated by design through random allocation, but finite sample variability may still result in residual imbalance in important prognostic factors. In both study designs, causal assumptions are explicitly represented using Directed Acyclic Graphs (DAGs) and counterfactual reasoning to guide covariate selection and identify potential sources of bias.

The methodological approaches for confounding control in observational research including regression adjustment, standardisation, inverse probability of treatment weighting (IPTW) and related g-methods are presented and discussed. These approaches are situated within the target trial emulation paradigm, which aims to reconstruct as closely as possible the conditions of a hypothetical randomised experiment from observational data. The strengths and limitations of each technique are demonstrated through simulated clinical scenarios from midwifery practice.

The empirical part of the thesis is split into two main applications. The first study examines the association between the organisational level of maternity units and maternal birth satisfaction in an observational multicenter cohort of low-risk women. The non-random allocation of women to different hospital settings dictates the use of IPTW based on the propensity scores to address baseline imbalances notably concerning parity and other obstetric characteristics, to improve covariate balance and bolster causal interpretation. The second is a randomised controlled trial of the effectiveness of maternal postural techniques and rebozo manoeuvres to reduce persistent occiput posterior position in labour. In this experiment, stratified randomisation is conducted to improve baseline comparability for parity, which is an important prognostic factor for intrapartum outcomes.

In general, the results of this thesis suggest the importance of explicitly addressing causal assumptions in both observational and experimental research in midwifery. While randomisation remains the gold standard design for causal inference, observational studies require carefully specified analytic strategies to approximate exchangeability. Thoughtful design choices are also helpful for randomised trials when dealing with strong prognostic factors.

In conclusion this work provides a consistent framework to incorporate statistical methodology, study design and clinical reasoning in maternity care research through the perspective of causal inference. In doing so it helps strengthen the validity and interpretability of evidence in midwifery research by bridging

observational and randomised approaches within a common conceptual framework which in turn supports more robust and clinically meaningful decision-making in maternity care.

2. Causal inference in midwifery research: confounding and methods for bias control in observational and randomized studies

In medical and midwifery research, the validity of causal conclusions critically depends on the ability to remove or not include systematic sources of bias that may distort the relationship between exposures and outcomes. Among these, confounding and mediation represents a major challenge in observational studies, where the absence of randomization on exposure makes groups inherently non-exchangeable. At the same time, even in randomized controlled trials (RCTs), residual imbalances and complex clinical pathways may still require careful consideration of causal structures.

Within this context, modern causal inference provides a unified framework to define, represent, and address these issues through tools such as counterfactual reasoning and Directed Acyclic Graphs (DAGs), where causal impact between variables is represented by an arrow and absence of causal impact is represented by the absence of the arrow. In this setting the distinction between confounders and mediators is essential to ensure valid estimation of causal effects and correct interpretation of both total effects and underlying mechanisms (Hernán & Robins, 2020).

The aim of this chapter is twofold: first, to define and describe confounding and mediation within the framework of causal inference; second, to present statistical approaches aimed at addressing the possible sources of bias in both observational and randomized studies using simulated data inspired by real-world clinical scenarios.

2.1. Definition and structure of Confounding

Confounding refers to a systematic source of bias that arises in standard statistical analysis when the exposure variable (or treatment) and the outcome share one or more common causes (confounders), thereby violating the assumption of exchangeability between exposed and unexposed groups. In causal inference, exchangeability means that the distribution of potential outcomes is independent of exposure assignment, so that exposed and unexposed individuals are comparable under the same hypothetical exposure condition. Using the potential outcomes framework, exchangeability can be formalized as the statistical independence between the potential outcomes and the observed exposure assignment:

$$Y^a \perp A$$

where Y^a denotes the potential (*counterfactual*) outcome that would be observed if, possibly contrary to fact, the individual were assigned exposure level a, and A represents the observed exposure. This condition states that the exposure assignment carries no information about the distribution of the counterfactual outcomes. Under exchangeability, the counterfactual risk among exposed and unexposed individuals is the same:

$$P(Y^a = 1|A = 1) = P(Y^a = 1|A = 0)$$

where Y^a represents the counterfactual outcome defined as the hypothetical outcome that would have been observed for the same individual under an alternative exposure condition that, in reality, did not occur (Rubin, 1974).

This assumption implies that the observed exposure assignment does not predict the counterfactual outcome. In randomized experiments, exchangeability is expected to hold by design because treatment allocation is independent of prognostic factors. In observational studies, however, exchangeability is often violated because exposure assignment may depend on individual baseline characteristics associated with the outcome.

More specifically, in observational settings the probability of receiving the exposure may depend on pre-treatment covariates C, such that:

$$P(A = 1|c = C) \neq P(A = 1)$$

and, if these covariates also affect the outcome, then the potential outcomes become associated with the exposure assignment.

Consequently, the observed outcomes among unexposed individuals no longer provide a valid estimate of the counterfactual outcomes that exposed individuals would have experienced under no exposure. In probabilistic terms, this means that:

$$P(Y^0 = 1|A = 1) \neq P(Y^0 = 1|A = 0)$$

where Y^0 denotes the potential outcome that would be observed under no exposure. The term $P(Y^0 = 1|A = 1)$ represents the counterfactual risk among exposed individuals, namely the probability that individuals who were actually exposed would have developed the outcome had they not been exposed. This quantity is unobservable because exposed individuals cannot simultaneously be observed in the unexposed state. Conversely, $P(Y^0 = 1|A = 0)$ represents the observed risk among individuals who were actually unexposed. Under exchangeability, these two probabilities are equal because the unexposed group can serve as a valid substitute for the missing counterfactual outcomes of the exposed group. However, in the presence of confounding, exposed and unexposed individuals differ systematically with respect to determinants of the outcome, and therefore the observed risk among the unexposed no longer provides an unbiased estimate of the counterfactual risk in the exposed group. This inequality thus formalizes the violation of exchangeability.

Thus, confounding arises when the observed association between an exposure and an outcome does not exclusively reflect a causal relationship, since part of the association is generated by external pre-exposure variables that influence both. For this reason, a statistical association between exposure and outcome cannot automatically be interpreted as evidence of causation (Hernán & Robins, 2020).

From a structural perspective, confounding is represented by the causal pathway:

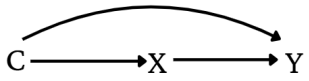


Figure 1

where C denotes the confounder, X the exposure, and Y the outcome. In this structure, the confounder C induces a backdoor path between X and Y ($X \leftarrow C \rightarrow Y$), which creates a non-causal association between exposure and outcome on top of the possible causal association between X and Y.

More formally, if $Y_i(1)$ denotes the outcome for individual i under exposure and $Y_i(0)$ denotes the outcome under no exposure, the individual causal effect can be expressed as:

$$Y_i^{a=1} - Y_i^{a=0}$$

However, only one of these two potential outcomes can be observed for each individual, while the other remains *counterfactual*. This impossibility of simultaneously observing both states for the same individual is commonly referred to as the fundamental problem of causal inference.

The presence of confounding undermines causal identifiability because the observed outcome among unexposed individuals may not accurately represent the counterfactual outcome that exposed individuals would have experienced under no exposure, thus exchangeability does not hold. In the presence of shared causes of exposure and outcome, exposed and unexposed groups can differ systematically even before the exposure occurs. As a consequence, differences in observed outcomes between the two groups (exposed and un-exposed) may reflect these pre-existing differences rather than the true causal effect of the exposure itself (Hernán & Robins, 2020).

Therefore, causal inference requires methods capable of restoring comparability between exposure groups, ideally reproducing the exchangeability conditions that would be achieved under randomization, by adjusting for the confounder.

2.2. Definition and structure of Mediator

A mediator is a variable that lies along the causal pathway between an exposure and an outcome and represents the mechanism (or a part of) through which the exposure exerts its effect. Unlike a confounder, which acts as a common cause of both exposure and outcome, a mediator is causally affected by the exposure and subsequently influences the outcome. In causal inference, mediators are therefore used to investigate not only whether an exposure has a causal effect, but also how that effect operates through intermediate mechanisms (Hernán & Robins, 2020; VanderWeele, 2015).

From a structural perspective, mediation is represented by the causal pathway:

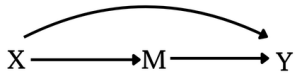


Figure 2

where X denotes the exposure, M the mediator, and Y the outcome. In this structure, part of the causal effect of the exposure on the outcome is transmitted through the mediator and the remaining part not operating through the mediator. Causal mediation analysis aims to decompose the total causal effect, which could be represented with the simple DAG $X \rightarrow Y$, into distinct components: a natural *direct* effect, corresponding to the effect of the exposure (X) on the outcome (Y) not operating through the mediator (M), and a natural *indirect* effect, corresponding to the portion of the effect transmitted through the mediator (M) itself (Hernán & Robins, 2020).

More formally, let $Y^{a=x}$ denote the potential outcome under exposure level x, and let $M^{a=x}$ denote the potential value of the mediator under exposure x. The *direct* effect represents the component of the exposure effect that operates independently of the mediator, while the *indirect* effect captures the extent to which changes in the mediator induced by the exposure are associated with changes in the outcome.

Because mediators are post-exposure variables, adjusting for them in conventional regression analyses may introduce overadjustment bias. By conditioning on the mediator, part of the causal effect (indirect effect) that researchers aim to estimate is artificially blocked and imputed to the mediator itself, leading to biased estimates of the total causal effect. Consequently, mediators should not be treated as standard confounders in adjustment procedures (Hernán & Robins, 2020).

The identification of mediated effects generally requires stronger assumptions than the estimation of total causal effects, including the absence of unmeasured confounding in the relationships between exposure, mediator, and outcome. For this reason, mediation analysis requires careful specification of the underlying causal structure.

2.3. Directed Acyclic Graphs as tools for causal reasoning

In modern causal inference, Directed Acyclic Graphs (DAGs) are increasingly used not merely as graphical representations of variables, but as conceptual tools for formalizing and communicating causal assumptions underlying a scientific research (Hernán & Robins, 2020). Their primary purpose is to support causal reasoning by explicitly representing the hypothesized relationships among exposures, outcomes, and other covariates before conducting the statistical analyses suitable to answer the research question.

One of the main advantages of DAGs is that they force researchers to define, in advance, the assumed causal structure of the phenomenon under investigation. This process promotes a theory-driven approach to data analysis and reduces the risk of relying exclusively on statistical associations when selecting variables for adjustment. In observational research, where exposure assignment is not randomized, this aspect is particularly important because inappropriate adjustment/non adjustment strategies may introduce/not remove substantial bias.

DAGs also provide a practical framework for identifying which variables should be/should be not controlled for in order to estimate causal effects appropriately. In particular, they facilitate the identification of confounders that generate open backdoor paths between exposure and outcome and therefore require adjustment. At the same time, DAGs help researchers avoid controlling for mediators, which may distort the causal effect of interest through overadjustment by opening closed backdoor paths (VanderWeele, 2015).

Another important contribution of DAGs is their ability to improve transparency and reproducibility in epidemiological research. Because the assumed causal relationships are represented explicitly, other researchers can critically evaluate the plausibility of the proposed causal structure and discuss alternative assumptions. Consequently, DAGs make the analytical process more interpretable and scientifically reproducible(Digitale, Martin, & Glymour, 2022).

For these reasons, DAGs are increasingly recommended as a preliminary step in the design and analysis of both observational and experimental studies. Their use encourages researchers to integrate substantive knowledge, clinical expertise, and statistical methodology within a coherent causal framework (Digitale et al., 2022).

2.4. A simulated example mimicking an observational study in midwifery research

To illustrate the concepts of confounding and mediation in a real-world applied setting, consider a multicenter observational cohort study aimed at evaluating whether the level of the hospital where birth occurs (I level Maternity Unit versus II level Maternity Unit) is associated with maternal birth satisfaction, a binary outcome. In this context, the exposure is defined as the type of hospital ($X = 1$ for I level Maternity Unit and $X = 2$ for II level Maternity Unit), which differ in terms of organisational, structural, and technical standards (i.e. high intensity care units, availability of 24-hour services, …).

The simulated study population consists of 7000 women with low to medium obstetric risk, who were eligible to give birth in either type of Maternity Unit, thereby allowing a degree of self-selection in the choice of birth setting. Parity, use of epidural analgesia during labour and maternal satisfaction at birth were collected.

Table 1 describes the characteristics of the simulated sample, including maternal birth satisfaction across Maternity Unit levels.

Table 1: Distribution of baseline characteristics and maternal birth satisfaction in the simulated observational cohort overall and according to Maternity Unit level

		Overall (n=7000)		I level Maternity Unit (n=4500)		II level Maternity Unit (n=2500)	
		n	***%***	***n***	***%***	***n***	***%***
Pluriparous	NO	2000	29	500	11	1500	60
	YES	5000	71	4000	89	1000	40
Epidural analgesia	NO	3313	47	2919	65	394	16
	YES	3687	53	1581	35	2106	84
Satisfaction	NO	2250	32	950	21	1300	52
	YES	4750	68	3550	79	1200	48

As a first attempt to address the study aim, we can hypothesize the association between Maternity Unit level and maternal birth satisfaction through the DAG shown below:

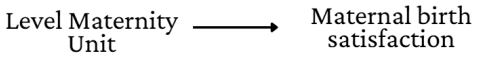


Figure 3

If the research question were addressed using a probabilistic approach, the conditional probabilities of the outcome given the exposure can be computed directly from the observed data.

Specifically, the probability of maternal birth satisfaction among women delivering in I level Maternity Units is:

$$P(Satisfaction = 1 | MaternityUnit = 1) = 3550/4500 = 0.79$$

whereas the corresponding probability among women delivering in II level Maternity Units is:

$$P(Satisfaction = 1 | MaternityUnit = 2) = 1200/2500 = 0.48$$

We could reinforce our results, evaluating the association between Maternity Unit level and maternal birth satisfaction through a regression model estimating $Risk\ Difference$ (RD). Considering maternal birth satisfaction as a binary outcome ($Y = Satisfaction$), a univariate regression model for binary outcomes with identity link estimating RD including the exposure variable ($X = MaternityUnit$) could evaluate the association between the type of Maternity Unit and the probability of reporting satisfaction at birth.

Formally, the model can be expressed as:

$$P(Y = 1) = \beta_0 + \beta_1 X$$

where Y denotes maternal birth satisfaction (1 = satisfied, 0 = not satisfied) and X represents the type of Maternity Unit (1 = I level Maternity Unit, 2 = II level Maternity Unit). In this specification, β_1 captures the crude RD in maternal satisfaction between the two Maternity Unit levels.

Consistently with the probabilistic results, the RD of maternal birth satisfaction in II level Maternity Units compared with I-level Maternity Units is equal to -0.31 ($95\% CI\ -0.33,\ -0.28$; $p < 0.001$), indicating lower risk of satisfaction among women delivering in II level Maternity Units.

From a purely associational perspective, these findings appear to support the hypothesis that giving birth in I-level Maternity Units positively influences maternal birth satisfaction.
However, since this is an observational study, particular attention is required in interpreting these associations as causal effects. In non-randomized settings, the assignment to exposure (i.e. choice of Maternity Unit) may depend on individual characteristics, and therefore exposed and unexposed groups may differ systematically prior to childbirth. Beyond the crude distribution of maternal birth satisfaction across hospital levels, Table 1 also suggests that several baseline characteristics are not equally distributed between women giving birth in I and II level Maternity Units. In particular, the distribution of parity and

epidural analgesia differs across exposure groups compared with the overall study population. Nulliparous women were disproportionately represented in Level II Maternity Units, accounting for 89% of women giving birth in II Level units compared with 11% in I Level units. Similarly, the use of epidural analgesia was more frequent among women giving birth in Level II Maternity Units (84%) than among those giving birth in Level I Maternity Units (35%).

These differences are particularly relevant from a causal inference perspective, as both parity and epidural analgesia may plausibly be associated not only with the choice of Maternity Unit level, but also with maternal birth satisfaction (Figure 4).

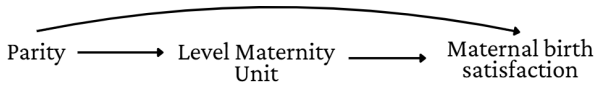


Figure 4

Consequently, the observed association between Maternity Unit level and maternal satisfaction at birth ($Satisfaction = 1$) may partially reflect differences in the underlying characteristics of women attending the two types of Maternity Units rather than the causal effect of Maternity Unit level itself. The unequal distribution of these variables across exposure groups suggests a potential lack of exchangeability between women giving birth in I and II level Maternity Units, leading to inadequate statistical approach without correct considerations.

2.5. Strategies to assess confounding

Evidence of non-exchangeability of women giving birth in I and II level Maternity Units suggests the next stage of the analysis should be to identify and appropriately address potential confounders. The imbalance in baseline characteristics across exposure groups (shown previously) suggests that the observed association between Maternity Unit level and maternal birth satisfaction may be partially explained by pre-existing differences, rather than a causal effect of the exposure itself.

In the motivating simulated example, parity is not evenly distributed across levels of Maternity Units, with nulliparous women ($Pluriparous = 0$) being more frequently represented among those giving birth in II level Maternity Units ($60\% \; versus \; 11\%$).

This relationship can be expressed through the following causal structure where also a causal impact of parity on maternal satisfaction, irrespective of the level of maternity unit, is hypothesized (Figure 4). This DAG implies that parity could be a confounder,as it may influence both the selection of Maternity Unit and maternal birth satisfaction concurrently, establishing a backdoor pathway between exposure and outcome (Figure 4).

This assumption is also supported by existing literature that identifies nulliparity as a relevant negative determinant of maternal birth experience and satisfaction (Fumagalli et al., 2021; Hochman et al., 2023). Therefore, failing to account for parity may lead to biased estimates of the effect of Maternity Unit level on maternal satisfaction.

To provide empirical evidence for the role of parity as a potential confounder in the association between Maternity Unit level and maternal birth satisfaction, we have to first establish whether this variable is associated with both the exposure and the outcome.

First, the association between parity and the choice of Maternity Unit can be tested by using conditional probabilities (orange arrow in the DAG represented in Figure 5).

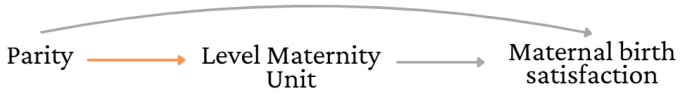


Figure 5

Specifically, the probability of admission to a I level compared to II level Maternity Unit varies by parity status.

$$P(MaternityUnit = 1|Pluriparous = 0) = 500/2000 = 0.25$$

$$P(MaternityUnit = 1|Pluriparous = 1) = 4000/5000 = 0.80$$

The higher prevalence of nulliparous in II level Maternity Units suggest a potential association between parity and place of birth, indicating that nulliparous women have a higher probability of choosing II level Maternity Units. Consistently with the percentages described in Table 1 where different percentages of parity are observed between levels of Maternity Unit. Such imbalance is inherent to non-randomized designs, where baseline characteristics are often evenly distributed across exposure groups. If exchangeability held, one would expect the distribution of parity within each Maternity Unit level to be similar to that of the overall study population (nulliparous$\sim$ 29% while pluriparous$\sim$ 71%).

Second, the association between parity and maternal birth satisfaction can be assessed in a similar way (orange arrow in the DAG represented in Figure 6).

Figure 6

Of note, if parity is considered as a new exposure variable and maternal birth satisfaction remains the outcome, then the above-mentioned DAG, Maternity Unit becomes a mediator. To assess the direct effect of parity on maternal birth satisfaction, the direct path should be blocked by an analysis stratified by Maternity Unit.

Within each Maternity level, the probability of reporting birth satisfaction differs for nulliparous and pluriparous:

$$P(Satisfaction = 1|Pluriparous = 0 \,\&\, MatrnityUnit = 1) = 350/500 = 0.70$$

$$P(Satisfaction = 1|Pluriparous = 1 \ \& MatrnityUnit = 1) = 3200/4000 = 0.80$$

$$P(Satisfaction = 1|Pluriparous = 0 \ \& MatrnityUnit = 2) = 600/1500 = 0.40$$

$$P(Satisfaction = 1|Pluriparous = 1 \ \& MatrnityUnit = 2) = 600/1000 = 0.60$$

These descriptive results suggest that parity may also affect the outcome of interest and to corroborate this intuition we performed a linear regression model for binary outcome with identity link.

Table 2: Distribution of parity in the simulated observational cohort overall and according to Maternity Unit level

		Overall (n=7000)		I level Maternity Unit (n=4500)		II level Maternity Unit (n=2500)	
		n	***%***	***n***	***%***	***n***	***%***
Pluriparous	NO	2000	28.6	500	25.0	1500	75.0
	YES	5000	71.4	4000	80.0	1000	20.0

First, the association between parity and the choice of Maternity Unit, represented in Table 2, was assessed, where the outcome was defined as giving birth in a II level Maternity Unit. The model was specified as:

$$P(MaternityUnit = 2) = \beta_0 + \beta_1 Parity$$

The findings confirm the negative association of being pluriparous with the probability to choose to give birth in a II level Maternity Unit ($RD = \ -0.55; \ 95\%CI \ -0.57, -0.53; \ p < .001$). This suggests a strong association of parity with the exposure, and that nulliparous women are over-represented in II level Maternity Units.

Second, the association between parity and maternal birth satisfaction is evaluated in a similar way, by specifying a generalized linear models with identity link in which maternal birth satisfaction is the binary outcome:

$$P(Satisfaction = 1) = \beta_0 + \beta_1 Parity + \beta_2 MaternityUnit$$

where β_1 estimates the adjusted Risk Difference (aRD) in maternal birth satisfaction associated with parity while holding the level of Maternity Unit constant, whereas β_2 estimates the association between Maternity Unit level and maternal satisfaction after adjustment for parity. The Maternity Unit variable was included in the model to estimate the direct effect of parity on the outcome independently of the level of care (see Figure 6). The result confirmed a positive association between pluriparous women and maternal satisfaction even after adjustment for Maternity Unit level ($aRD = \ 0.15; \ p < .001$). More specifically, pluriparous women reported an absolute 15% greater likelihood of reporting maternal satisfaction at birth than nulliparous women regardless of the Maternity Unit where birth occurred.

Overall, these results consistently demonstrate that parity satisfies the structural definition of a confounding variable, as it is associated with both the exposure (Maternity Unit level) and the outcome (Maternal birth satisfaction).
For this reason, parity must be explicitly taken into account in the analytical strategy to avoid biased estimation of the causal effect of Maternity Unit level on maternal satisfaction. Thus, estimates obtained without appropriate adjustment for parity cannot be considered causal estimates because they are confounded by baseline differences in parity between exposure groups.

2.6. Strategies to control confounding in observational studies

Control of confounding aims to reduce or remove non-causal associations between exposure and outcome in order to obtain unbiased estimates of causal effects. The main goal of these methods in the causal inference framework is to restore exchangeability between exposure groups, ideally reproducing the comparability that would be achieved under randomisation (Hernán & Robins, 2020). Graphically, this is equivalent to blocking all open backdoor paths between the exposure and the outcome through common causes (Figure 7).

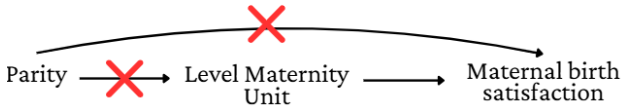


Figure 7

In observational studies exposure assignment is generally driven by prognostic characteristics of individuals which makes confounding an important methodological challenge. Consequently, in observational analysis, the best approach to mimicking the conditions of a hypothetical target trial (Hernán & Robins, 2020), is to use statistical methods that can control for measured confounders. Several analytical approaches have been developed to handle confounding under different assumptions and study settings. Common methods include regression adjustment, inverse probability of treatment weighting (IPTW), and G-estimation. These approaches differ in their statistical implementation, but they have a common purpose to improve comparability between exposure groups and to reduce the estimation of causal effect in observational study.

From a causal inference perspective, it is important to distinguish between marginal and conditional causal effects. Marginal effects refer to comparisons of potential outcomes averaged over the distribution of covariates in the study population (e.g. the Average Treatment Effect, ATE), whereas conditional effects refer to comparisons within levels of baseline covariates (i.e. conditional on confounders such as parity). Regression adjustment typically targets conditional effects, while methods such as inverse probability weighting and G-estimation are used to estimate marginal effects (Hernán & Robins, 2020).

2.6.1. Regression adjustment

One of the most common ways to control for confounding in observational studies is through regression adjustment. In causal inference, regression models are used to estimate the association between exposure and outcome conditional on measured confounders. Regression adjustment seeks to reduce non-causal

associations and improve comparability between exposure groups by adjusting for variables that affect both exposure and outcome at the same time (Hernán & Robins, 2020).

More generally, regression adjustment is interpreted in modern causal inference as a way to emulate a target trial. From this perspective, observational analyses seek to approximate the conditions of an idealised randomised trial as closely as possible by explicitly defining eligibility criteria, treatment assignment, follow-up, and causal contrasts (Hernán & Robins, 2016). Hence, covariates included in regression models should be measured prior to exposure assignment (i.e. at baseline), similarly to the way prognostic variables would be assessed before randomization in a RCT.

From the graphical perspective, conditioning on baseline confounders blocks the open backdoor paths between exposure and outcome, thereby reducing confounding bias. Under the assumption of conditional exchangeability — meaning that all relevant confounders have been correctly measured and correctly included in the model — adjusted regression estimates may approximate the causal effect of the exposure on the outcome.

In our motivating example, parity was identified as a potential confounder because it predicts the choice of Maternity Unit and is associated with the exposure and maternal birth satisfaction. Therefore, parity can be adequately incorporated into the adjustment strategy. Nevertheless, variables that occur after exposure assignment should not be automatically treated as confounders, since conditioning on post-exposure variables may introduce bias or block part of the causal effect of interest (i.e. epidural use in our example).

The adjusted model can be written as a generalised linear model with identity link:

$$P(Satisfaction = 1) = \beta_0 + \beta_1 MaternityUnit + \beta_2 Parity$$

In this model, β_1 shows the adjusted Risk Difference (aRD) associated with Maternity Unit level while adjusting for parity.

The estimated association between Maternity Unit level and maternal birth satisfaction was attenuated after adjustment for parity compared with the crude analysis ($aRD\ for\ parity = -0.23$, p<0.001 versus $crude\ RD = -0.31,\ p < 0.001$), indicating that the association initially observed was partly explained by differences in parity distribution between exposure groups. This finding is indicative of confounding and illustrates how adjustment by regression can partially restore comparability between women giving birth in different Maternity Units.

Because the effect of Maternity Unit level may differ according to parity, an interaction term between Maternity Unit level and parity was additionally considered. This allowed the effect of Maternity Unit level to be estimated separately among nulliparous and pluriparous women:

$$P(Satisfaction = 1) = \beta_0 + \beta_1 MaternityUnit + \beta_2 Parity + \beta_3 (MaternityUnit \times Parity)$$

The interaction model showed that the estimated probability of maternal birth satisfaction among nulliparous women giving birth in I level Maternity Units, corresponding to the reference category, was

0.70 ($\beta_0 = 0.70$). The coefficient for Maternity Unit level ($\beta_1 = -0.30$) therefore represents the RD associated with giving birth in a II-level rather than a I-level Maternity Unit among nulliparous women. Thus, among nulliparous women, giving birth in a II-level Maternity Unit was associated with a 30 percentage-point lower probability of maternal birth satisfaction ($RD = -0.30;\ 95\%\ CI\ -0.35,\ -0.25;\ p < 0.001$).

The coefficient for parity ($\beta_2 = 0.10$) represents the difference in maternal birth satisfaction between pluriparous and nulliparous women among those giving birth in I level Maternity Units. The interaction coefficient ($\beta_3 = 0.10$) indicates that the effect of Maternity Unit level differs by parity. Specifically, the effect of II versus I level Maternity Unit among pluriparous women is obtained by summing the main effect of Maternity Unit and the interaction term ($\beta_1 + \beta_3 = -0.30 + 0.10 = -0.20$). Thus, among pluriparous women, giving birth in a II-level Maternity Unit was associated with a 20 percentage-point lower probability of maternal birth satisfaction. The statistically significant interaction ($p = 0.001$) therefore suggests that the effect of Maternity Unit level on maternal birth satisfaction is modified by parity.

To further illustrate the importance of correctly identifying the causal role of variables before adjustment (*confounders* versus *mediators*), epidural analgesia was additionally included in the regression model together with parity:

$$P(Satisfaction = 1) = \beta_0 + \beta_1 MaternityUnit + \beta_2 Parity + \beta_3 Epidural$$

In this model, β_1 shows the adjusted Risk Difference (aRD) associated with Maternity Unit level while adjusting for parity and epidural use. The association between Maternity Unit level and maternal birth satisfaction was further attenuated ($aRD = -0.18;\ 95\%CI\ -0.20,\ -0.15;\ p < 0.001$). However, this reduction should not necessarily be interpreted as improved control of confounding. Unlike parity, which represents a baseline confounder, epidural analgesia occur after exposure assignment and therefore act as a mediator of the effect of Maternity Unit level on maternal satisfaction. In this context, conditioning on epidural use may artificially block part of the causal pathway through which the Maternity Unit exerts its effect on the outcome. Consequently, adjustment for post-exposure mediators may lead to overadjustment bias and underestimation of the total causal effect of interest.

Rather than treating epidural analgesia simply as a variable to be excluded from the adjustment set, its position on the hypothesised causal pathway provides an opportunity to investigate whether it mediates part of the effect of Maternity Unit level on maternal birth satisfaction. In this context, causal mediation analysis can be used to decompose the total causal effect of Maternity Unit level into a direct effect and an indirect effect operating through epidural analgesia. The indirect effect represents the component of the effect of Maternity Unit level on maternal birth satisfaction that operates through changes in the probability of receiving epidural analgesia. Conversely, the direct effect represents the component of the effect that operates through pathways other than epidural analgesia. In the present example, parity is considered a baseline confounder and is therefore accounted for in the mediation analysis (Figure 8).

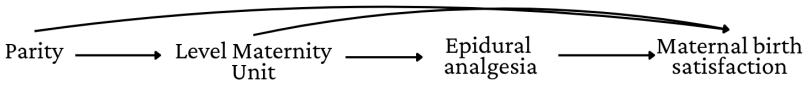


Figure 8

Thus, while adjustment for epidural analgesia in the previous regression model may lead to overadjustment when the total effect is the estimand of interest, treating epidural analgesia as a mediator allows the clinical pathway through which Maternity Unit level may influence maternal birth satisfaction to be investigated explicitly.

To further explore the role of epidural analgesia as a potential mediator, a causal mediation analysis was performed using a regression-based approach. First, the probability of receiving epidural analgesia was modelled as a function of Maternity Unit level and parity.

$$P(Epidural = 1) = \beta_0 + \beta_1 MaternityUnit + \beta_2 Parity$$

The results showed that women giving birth in II level Maternity Units had a higher probability of receiving epidural analgesia compared with women in I level Maternity Units ($aRD = 0.42;\ 95\%\ CI\ 0.39,\ 0.44;\ p < 0.001$), while pluriparous women had a lower probability of receiving epidural analgesia compared with nulliparous women, after adjustment for Maternity Unit level ($aRD = -0.15;\ 95\%\ CI\ -0.18,\ -0.12;\ p < 0.001$). These findings support the hypothesised role of epidural analgesia as a potential mediator of the relationship between Maternity Unit level and maternal birth satisfaction.

Second, maternal birth satisfaction was modelled as a function of Maternity Unit level, epidural analgesia and parity.

$$P(Satisfaction = 1) = \beta_0 + \beta_1 MaternityUnit + \beta_2 Epidural + \beta_3 Parity$$

The results showed that epidural analgesia was associated with a lower probability of maternal birth satisfaction ($aRD = -0.13;\ 95\%\ CI\ -0.15,\ -0.10;\ p < 0.001$), and the association between Maternity Unit level and satisfaction remained negative after adjustment for epidural analgesia and parity ($aRD = -0.18;\ 95\%\ CI\ -0.21,\ -0.15;\ p < 0.001$). Pluriparous women had a higher probability of reporting maternal birth satisfaction compared with nulliparous women ($RD = 0.13;\ 95\%\ CI\ 0.10,\ 0.16;\ p < 0.001$).

Overall, these findings suggest that the associations of Maternity Unit level and parity with maternal birth satisfaction may be partially related to their association with epidural analgesia.

These examples further highlight the importance of considering the causal role of variables in the analytical strategy, distinguishing baseline confounders that require adjustment from post-exposure mediators that may explain part of the causal effect.

However, conventional regression adjustment has important limitations. First, valid estimation requires the correct specification of the functional relationship between exposure, confounders, and outcome.

Second, conventional regression approaches can be invalid in the presence of time-varying confounding by prior exposure, where post-baseline factors prognostic are simultaneously confounders and consequences of prior treatment. In these cases, regression adjustment may add bias and other causal methods such as inverse probability weighting and other g-methods are needed (Hernán & Robins, 2020).

2.6.2. Inverse Probability of Treatment Weighting (IPTW)

Inverse Probability of Treatment Weighting (IPTW) is based on the same probabilistic notion introduced in the previous sections, namely that the assignment is not random but depends on baseline characteristics. However, instead of adjusting for such variables within a regression model, IPTW uses the estimated distribution of these variables to reweight the sample, creating a pseudo-population where exposure is independent of measured confounders.

The basic idea is that each individual is allocated a weight that is the inverse of the probability of receiving the treatment that they actually got. These probabilities are estimated using a propensity score model.

In the motivating example, the probability of giving birth in each level of Maternity Unit, conditional on baseline characteristics such as parity (see Figure 9), can be expressed as:

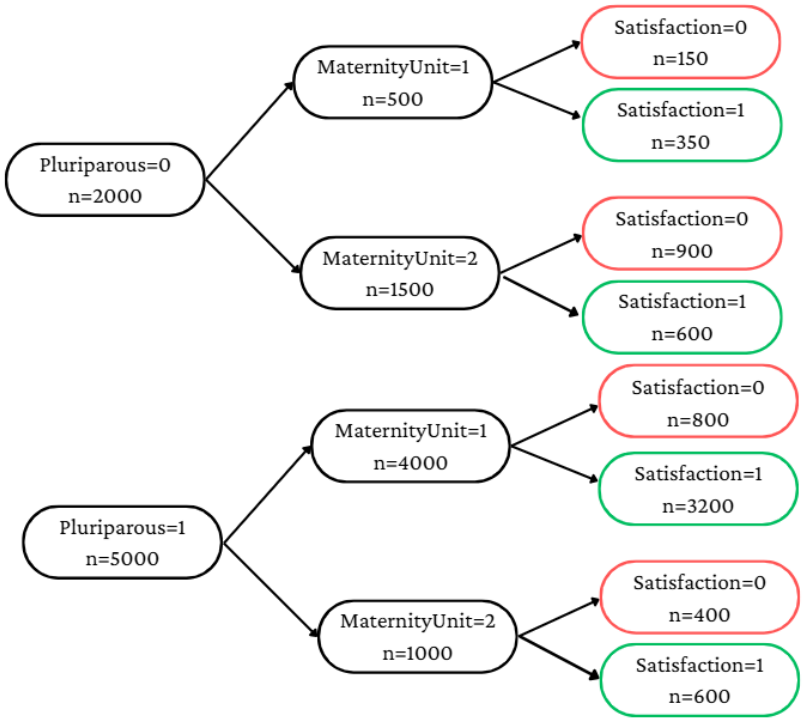


Figure 9

$$P(MaternityUnit = 1|Pluriparous = 0) = 500/2000 = 0.25$$

$$P(MaternityUnit = 2|Pluriparous = 0) = 1500/2000 = 0.75$$

$$P(MaternityUnit = 1|Pluriparous = 1) = 4000/5000 = 0.80$$

$$P(MaternityUnit = 2|Pluriparous = 1) = 1000/5000 = 0.20$$

These probabilities reflect the fact that treatment assignment is systematically related to baseline characteristics and therefore exchangeability fails.

Then, the inverse probability weights are defined as ω_{P0M1}(for nulliparous in I level Maternity Unit), ω_{P0M2}(for nulliparous in II level Maternity Unit), ω_{P1M1}(for pluriparous in I level Maternity Unit) and ω_{P1M2}(for pluriparous in II level Maternity Unit) (Figure 10).

$$\omega_{P0M1} = 1/P(MaternityUnit = 1|Parity = 0) = 1/(500/2000) = 1/0.25 = 4$$

$$\omega_{P0M2} = 1/P(MaternityUnit = 2|Parity = 0) = 1/(1500/2000) = 1/0.75 = 1.33$$

$$\omega_{P1M1} = 1/P(MaternityUnit = 1|Parity = 1) = 1/(4000/5000) = 1/0.80 = 1.25$$

$$\omega_{P1M2} = 1/P(MaternityUnit = 1|Parity = 1) = 1/(100/5000) = 1/0.20 = 5$$

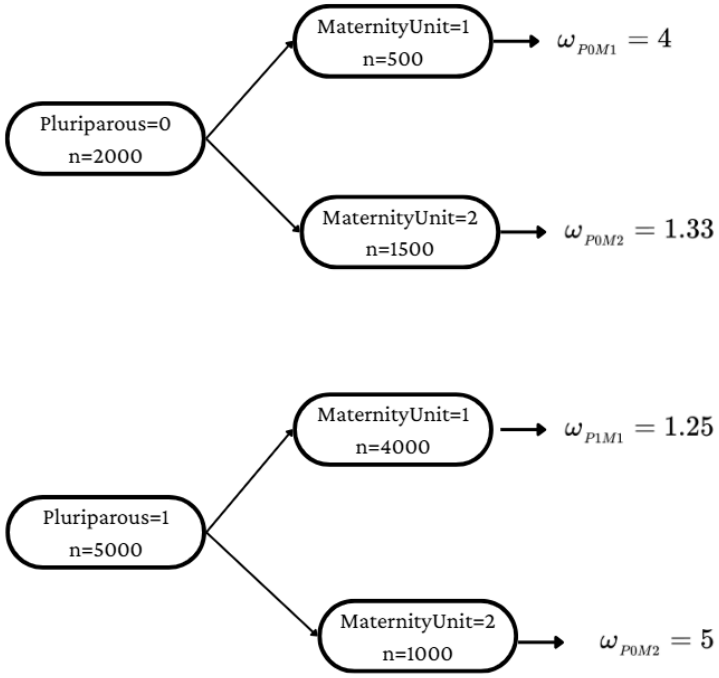


Figure 10

The idea is to implement a weighting procedure that creates a pseudo-population where the distribution of measured baseline confounders is balanced across exposure groups and thus mimics the conditions of a randomized experiment.

Figure 11 shows the pseudo-population generated through IPTW. After estimating the probability of receiving the observed exposure conditional on baseline covariates, each individual is weighted by the inverse of this probability.

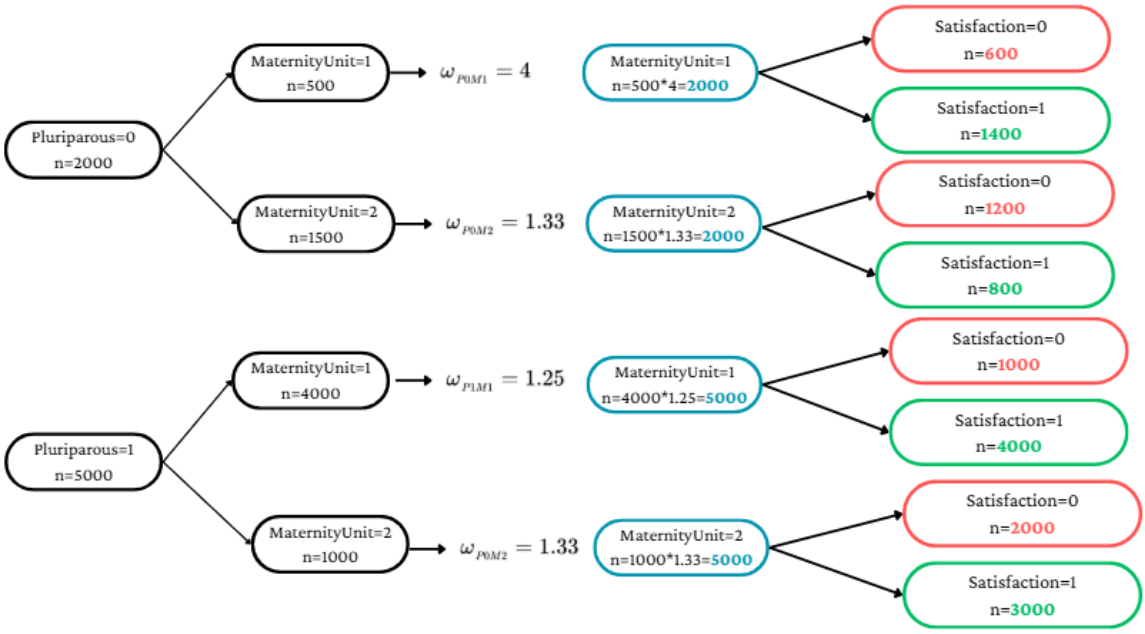


Figure 11

The blue number is the weighted sample after applying these inverse probability weights. In this re-weighted population, the contribution of each individual is modified according to their probability of treatment assignment.

The weighted pseudo-population is obtained by multiplying the observed number of individuals in each exposure-confounder stratum by the corresponding inverse probability weight previously calculated (ω). In practice, each woman contributes to the pseudo-population proportionally to the inverse of her probability of receiving the Maternity Unit level that she actually attended. Women who were underrepresented in a given exposure group receive a higher weight, whereas women who were overrepresented receive a lower weight.

In the present example, the original number of nulliparous women giving birth in I level Maternity Units was 500. Since their inverse probability weight was equal to 4 (ω_{P0M1}), their weighted contribution to the pseudo-population becomes:

$$500 \times 4 = 2000$$

Similarly pseudo-population for each stratum was calculated (see Figure 11)

After weighting, each parity category contributes equally to both Maternity Unit levels according to its distribution in the overall study population. The resulting weighted pseudo-population therefore contains 2000 nulliparous women and 5000 pluriparous women within each exposure group, reproducing the

parity distribution observed in the entire cohort ($P(Pluriparous = 0) = 2000/7000 = 0.29$ and $P(Pluriparous = 1) = 5000/7000 = 0.71$, respectively). This balancing process removes the association between parity and Maternity Unit level and approximates the exchangeability conditions that would be expected under random allocation.

The weighted frequencies of maternal birth satisfaction (green number in Figure 11) and dissatisfaction (red number in Figure 11) are then recalculated within each exposure and parity stratum using the same weights. Consequently, the estimation of outcome probabilities is no longer driven by the original imbalance in parity between Maternity Unit levels but by the balanced pseudo-population generated through IPTW.

Once the pseudo-population has been built from the IPTW, it is possible to assess again the association between the Maternity Unit level and maternal birth satisfaction with weighted probabilities. The weighting procedure attempts to balance parity across exposure groups and the resulting weighted sample approximates the exchangeability conditions expected under randomization.

Hence, in the pseudo-population, the probability of maternal birth satisfaction is re-estimated using the weighted frequencies instead of the original observed counts. Weighted conditional probabilities are formally expressed as:

$$P_{\omega}(Satisfaction = 1|MaternityUnit = 1) = 5400/7000 = 0.77$$

$$P_{\omega}(Satisfaction = 1|MaternityUnit = 2) = 3800/7000 = 0.54$$

where the subscript ω denotes that the probabilities are estimated in the weighted pseudo-population obtained by IPTW. The association remains after weighting, suggesting that differences in maternal birth satisfaction persist even after parity is balanced between exposure groups.

Unlike the crude probabilities estimated in the original observational sample, these weighted probabilities account for the unequal distribution of parity between exposure groups. The comparison between Maternity Unit levels is therefore less confounded and better captures the causal effect of the exposure on maternal birth satisfaction.

In this example, IPTW was implemented in STATA with the $teffects\ ipw$ command. The treatment model was specified through a linear regression model for binary outcome where the probability of giving birth in an established Maternity Unit level was estimated conditional to parity. The inverse probability weights obtained were then used to estimate the average treatment effect (ATE) of Maternity Unit level on maternal birth satisfaction.

Results showed a marginal causal effect (ATE) of -0.23 ($95\% CI\ -0.26, -0.20$; $p < 0.001$). After IPTW balancing parity between exposure groups, women giving birth in II level Maternity Units still had a lower probability ($\sim -23\%$) of reporting maternal birth satisfaction compared to women giving birth in I level Maternity Units. More in detail, the estimated weighted mean probability of maternal birth satisfaction in the reference group (I level Maternity Unit) was equal to 0.77, indicating that approximately 77% of women would report satisfaction under the corresponding exposure condition

in the weighted pseudo-population. Whereas in II level Midwifery Unit about 51% ($0.77 - 0.23 = 0.51$) are satisfied. It is interesting to note that these estimates are fully consistent with the weighted probabilities computed previously using the probabilistic approach, which confirms the consistency between the direct calculation of weighted probabilities and the regression-based IPTW estimation.

The magnitude of the effect was attenuated after weighting as compared to the crude association before adjustment, indicating that part of the original association was explained by confounding due to parity. However, the association after IPTW indicates that differences in maternal birth satisfaction between Maternity Unit levels are still present after balancing the measured confounder.

2.6.3. Conceptual introduction to G-estimation

While IPTW estimates causal effects by creating a weighted pseudo-population, G-estimation follows a different approach based on stratification and recomposition of the confounder strata.

The general idea is to compare treated and untreated individuals within homogeneous strata of the confounders, in which individuals are more comparable in baseline characteristics. Within each stratum, the conditional probabilities of the outcome given treatment assignment are estimated separately. In a second step, the conditional effects are estimated in each of the confounder strata and then the information from the different strata is recombined to obtain an overall causal estimate less susceptible to confounding. In this context, G-estimation attempts to reproduce the exchangeability conditions expected under randomization by estimating the effects of treatment among individuals with similar baseline characteristics (Hernán & Robins, 2020).

In the motivating example, parity is a confounder because it is associated both with the choice of Maternity Unit and with maternal birth satisfaction. Therefore, the comparison between women giving birth in different Maternity Unit levels should not be performed in the overall sample alone, but also within homogeneous strata of parity. The first step of G-estimation is thus to stratify the population on the levels of the confounder which, in our example, is parity.

The conditional probabilities of maternal birth satisfaction can be estimated separately by level of Maternity Unit within each stratum (for the absolute frequencies needed refer to see Figure 10).

- I level Maternity Unit

$$P(Satisfaction = 1|MaternityUnit = 1\&Pluriparous = 0) = 350/500 = 0.70$$

$$P(Satisfaction = 1|MaternityUnit = 1\&Pluriparous = 1) = 3200/4000 = 0.80$$

- II level Maternity Unit

$$P(Satisfaction = 1|MaternityUnit = 2\&Pluriparous = 0) = 600/1500 = 0.40$$

$$P(Satisfaction = 1|MaternityUnit = 2\&Pluriparous = 0) = 600/1000 = 0.60$$

These conditional probabilities indicate that maternal birth satisfaction varies by parity status within each Maternity Unit level. Thus the crude association observed in the overall population may partly reflect the uneven distribution of parity among exposure groups rather than the causal effect of Maternity Unit level itself.

$$P(Pluriparous = 0) = 2000/7000 = 0.29$$

$$P(Pluriparous = 1) = 5000/7000 = 0.71$$

These probabilities correspond to the proportion of nulliparous and pluriparous women in the whole study population and are used to reconstruct the expected probability of maternal birth satisfaction at each level of Maternity Unit.

- I level Maternity Unit

$$P(Satisfaction = 1|MatenityUnit = 1) = P(Satisfaction = 1|MaternityUnit = 1 \,\&\, Pluriparous = 0) * P(Pluriparous = 0) + P(Satisfaction = 1|MaternityUnit = 1\&Pluriparous = 1) * P(Pluriparous = 1)$$

$$P(Satisfaction = 1|MatenityUnit = 1) = 0.70 * 0.29 + 0.80 * 0.71 = 0.20 + 0.58 = 0.77$$

- II level Maternity Unit

$$P(Satisfaction = 1|MatenityUnit = 2) = P(Satisfaction = 1|MaternityUnit = 2 \,\&\, Pluriparous = 0) * P(Pluriparous = 0) + P(Satisfaction = 1|MaternityUnit = 2 \,\&\, Pluriparous = 1) * P(Pluriparous = 1)$$

$$P(Satisfaction = 1|MatenityUnit = 2) = 0.40 * 0.29 + 0.60 * 0.71 = 0.10 + 0.43 = 0.54$$

After recomposition of the parity strata, the adjusted probabilities of maternal birth satisfaction remain higher among women giving birth in I level Maternity Units as compared with II level Maternity Units.
These recomposed probabilities, as opposed to the crude probabilities estimated in the original observational sample, correct for the imbalance in the parity distribution between exposure groups. Consequently, the comparison between Maternity Unit levels is less confounded and better estimates the causal effect of Maternity Unit level on maternal birth satisfaction.

The probabilistic stratification approach presented in this section aims to illustrate the conceptual logic underlying G-estimation, namely the reconstruction of exchangeability within levels of the confounders. The formal implementation of G-estimation involves the use of statistical models to estimate the treatment effect that makes treated and untreated individuals comparable after having adjusted for measured confounders (Hernán & Robins, 2020).

In the present example, the recomposition of the parity strata was conducted in STATA using the $teffects\ ra$ command, specifying a regression adjustment model for maternal birth satisfaction conditional on parity. The method calculates the expected probability of the outcome for each parity stratum and then re-aggregates these conditional estimates according to the overall distribution of parity in the study population.

The analysis showed an Average Treatment Effect (ATE) equal to -0.23 ($95\%\ CI\ -0.26;\ -0.20;\ p < 0.001$), meaning that the probability of maternal birth satisfaction was lower among women who gave birth in II level Maternity Units compared to I level Maternity Units after adjusting for parity.

More precisely, the estimated adjusted probability of maternal birth satisfaction in the reference group (I level Maternity Unit) was equal to 0.77 ($POmean$). The corresponding adjusted probability in II level Maternity Units using the estimate ATE was approximately 0.54 ($POmean - ATE = 0.77 - 0.23$).

Importantly, these estimates are fully consistent with the recomposed probabilities previously obtained through the probabilistic stratification approach, thus confirming the coherence between the direct recomposition of the confounder strata and the regression model applied in STATA.

More importantly, IPTW and standardization are grounded in the counterfactual framework of causal inference, as both seek to reconstruct the outcome distribution that would be observed under different hypothetical exposure conditions. Although achieved through different analytical procedures, these methods attempt to create a pseudo-population in which exposure groups are comparable with respect to measured confounders. In contrast, traditional regression adjustment addresses confounding by conditioning on covariates within a statistical model rather than by explicitly reconstructing counterfactual populations (Hernán & Robins, 2020).

These methods differ not only in their statistical implementation, but also in the type of causal estimand they target. Regression adjustment typically estimates conditional effects, while IPTW and standardisation target marginal effects. Under correct model specification and absence of unmeasured confounding, both approaches identify the same causal estimand, although they rely on different identifying assumptions.

The fact that regression adjustment, IPTW, and standardization (G-estimation) yielded virtually identical estimates in this simulated example is a consequence of the simplified data-generating structure. The simulated dataset included a single binary confounder (parity), a binary exposure, and correctly specified statistical models without interaction terms. Under these conditions, all three methods effectively reconstruct the same counterfactual comparison and therefore converge on the same estimate of the Average Treatment Effect (ATE). In more complex observational settings involving multiple confounders, continuous variables, non-linear relationships, interaction effects, or model misspecification, differences between these approaches would generally be expected.

2.7. A simulated example mimicking an Randomized Controlled trials (RCTs) in midwifery research

In randomized controlled trials (RCTs), treatment allocation is determined by a random mechanism controlled by the investigator. Under ideal conditions, randomization creates exchangeability between treatment groups, since treated and untreated individuals are expected to have a similar distribution of both measured and unmeasured prognostic factors (i.e. confounders). Consequently, differences in outcomes between groups can be interpreted as causal effects of the intervention (Hernán & Robins, 2020).

In practice, however, randomization does not guarantee perfect balance in every realized sample. Particularly in studies with limited sample sizes or when important prognostic variables are strongly associated with the outcome, baseline imbalances may still occur by chance. In these situations, strategies such as stratified randomization may be used to improve comparability between treatment groups and increase the precision of treatment effect estimates.

To illustrate these concepts, a simulated dataset was generated to reproduce a simplified randomized trial inspired by the clinical context of persistent occiput posterior position (POPP) at birth. The simulation was designed to reflect a realistic clinical scenario based on preliminary evidence available during the planning phase of the ReMaP-POPP trial (Ornaghi et al., 2025). In particular, the distribution of parity and the expected occurrence of POPP were informed by findings from a previously published retrospective cohort study investigating maternal postures and advanced intrapartum techniques in women presenting with occiput posterior position during labour (Fumagalli et al., 2024). The same study provided key assumptions for the sample size calculation subsequently reported in the trial protocol (Chapter 4) (Ornaghi et al., 2025).

A total of 1000 women were generated. Parity was simulated first, with approximately 55% of women classified as nulliparous and 45% as pluriparous, mirroring the distribution observed in the preliminary retrospective cohort (Fumagalli et al., 2024). Treatment allocation and outcome occurrence were then generated according to prespecified probabilistic mechanisms. The intervention was assumed to reduce the probability of POPP, whereas nulliparity was specified as a strong prognostic factor associated with an increased risk of persistent malposition.

Two alternative randomization scenarios were created. The first scenario reproduced a randomized trial in which a moderate imbalance in parity distribution arose between treatment groups, allowing the consequences of residual imbalance to be explored (Section 2.7). The second scenario applied stratified randomization according to parity before treatment allocation, thereby improving baseline comparability between groups (Section 2.8). Together, these examples illustrate how exchangeability may be affected by chance imbalances in important prognostic factors and how design-based strategies can be used to address this issue.

In the simulated trial, women presenting with occiput posterior position during labour were allocated either to an intervention group receiving maternal postures and rebozo techniques or to a standard care group. The primary outcome was persistent occiput posterior position at birth (POPP). Parity was included a key prognostic variable because nulliparous women have constantly shown a higher risk of persistent malposition compared with pluriparous women (Barrowclough, Lin, Kool, Hofmeyr, & Crowther, 2022; Caughey, Sharshiner, & Cheng, 2015; Cheng, Shaffer, & Caughey, 2006; Fumagalli et al., 2024). The hypothesised causal structure underlying the simulated trial is represented in Figure 12.

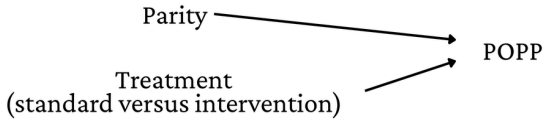


Figure 12

This example will allow visualization of how imbalance in a strong prognostic factor may influence the estimated treatment effect even within an RCT framework, and which strategy could be used in order to improve comparability between groups.

Table 3 describes the distribution of parity and POPP occurrence in the simulated randomized trial population overall and according to treatment allocation (simulated scenario one). Overall, 54% of women were nulliparous, while 46% were multiparous. Random treatment allocation produced groups of similar size: 451 women in the standard care group - $Standard\ group$ - and 549 in the intervention group -$Intervention\ group$.

Table 3: Distribution of parity and persistent occiput posterior position (POPP) in the simulated randomized trial overall and according to treatment allocation under simple randomization

		Overall (n=1000)		Standard group (n=451)		Intervention group (n=549)	
		n	***%***	***n***	***%***	***n***	***%***
Pluriparous	NO	535	54	202	45	333	61
	YES	465	46	249	55	216	39
POPP	NO	610	61	233	52	377	69
	YES	390	39	218	48	172	31

Despite randomization, a modest imbalance in parity distribution was observed between groups, with a slightly lower proportion of pluriparous women in the intervention group compared with the control group (39% *versus* 55%).

The overall probability of POPP was 39%, with a lower proportion of POPP observed in the intervention group compared with the standard care group (31% *versus* 48%).

As a first analytical step, the association between treatment allocation and POPP was evaluated using a generalized linear model for binomial outcome with identity link, estimating RD, without control for parity. This analysis relies on the theoretical assumption underlying randomized controlled trials, namely that randomization creates exchangeability between treatment groups by balancing both measured and unmeasured prognostic factors. Under this assumption, the crude association can be interpreted as an unbiased estimate of the causal effect of the intervention (Hernán & Robins, 2020).

$$P(POPP = 1) = \beta_0 + \beta_1 X$$

whereas β_1 captures the crude RD in POPP associated with treatment allocations. The crude analysis showed that women allocated to the intervention group had a lower probability of POPP compared with women receiving standard care; the intervention was associated with a reduction of 17% of the probability of POPP. This estimate corresponds to a marginal causal effect at the population level, under the assumption that randomization ensures exchangeability by design.

However, although treatment allocation was randomized, the distribution of parity was not perfectly balanced between groups in the simulated sample (Table 3). Because parity is a strong prognostic factor

for fetal malposition persistence, an adjusted model including parity was subsequently fitted to evaluate whether residual imbalance could influence the estimated treatment effect.

To account for the prognostic role of parity, a second generalized linear model with identity link was fitted including both treatment allocation and parity.

$$P(POPP = 1) = \beta_0 + \beta_1 X + \beta_2 X$$

whereas β_1 estimated the adjusted RD in POPP associated with treatment allocations and $\beta_2$the association between POPP and parity, after adjustment for treatment allocation. The intervention remained associated with a lower probability of POPP ($adjusted\, RD = -0.20;\ 95\%CI\ -0.26,\ -0.14;\ p < 0.001$), corresponding to an absolute reduction of risk of POPP ($\sim 20\%$), higher than the previous estimation. Parity showed a strong independent association with the outcome ($POPP = 1$), with pluriparous women having a lower probability of POPP compared with nulliparous women ($adjustedRD = -0.20;\ 95\%CI\ -0.26,\ 0.15;\ p < 0.001$). Parity remained strongly associated with the outcome independently of treatment allocation.

Although adjustment for parity had minimal impact on the estimated treatment effect in this simulated trial, the analysis highlights an important methodological consideration. Even under randomized allocation, finite-sample imbalance in prognostic covariates may occur and influence the precision of effect estimates, particularly when such factors are strongly associated with the outcome.

2.8. Strategies to minimize residual confounding in Randomized Controlled Trials (RCTs)

A possible strategy to reduce residual imbalance in important prognostic variables is stratified randomization (simulated scenario two). In stratified randomization, participants are first classified according to the levels of a clinically relevant baseline variable and are subsequently randomized separately within each stratum. This strategy to minimize residual confounding takes place during the design stage with the aim to ensure exchangeability for the groups within levels of stratification (Hernán & Robins, 2016). In a subsequent simulated example of the RCT, women were stratified according to parity status ($Pluriparous = 0$ *versus* $Pluriparous = 1$) before treatment allocation. This procedure ensures that the proportion of nulliparous and pluriparous women remains comparable between treatment groups, thereby reducing the possibility that imbalance in parity may influence the estimated treatment effect.

Table 4 describes the distribution of parity and POPP occurrence in the simulated trial population after randomization stratified by parity. Unlike the previous simulated trial generated through simple random allocation, parity distribution was highly comparable between treatment groups. The proportion of pluriparous women was similar in the intervention and standard care groups (46% *versus* 47%), indicating that the prognostic factor had been effectively balanced at baseline through the allocation procedure.

Table 4: Distribution of parity and persistent occiput posterior position (POPP) in the simulated randomized trial overall and according to treatment allocation after stratified randomization by parity

		Overall (n=1000)		Standard care (n=515)		Intervention group (n=485)	
		n	***%***	***n***	***%***	***n***	***%***
Pluriparous	NO	535	54	278	54	257	53
	YES	465	47	237	46	228	47
POPP	NO	589	59	250	49	339	70
	YES	411	41	265	51	146	30

As in the previous analysis, the association between treatment allocation and POPP was initially evaluated through a generalized linear model with identity link estimating Risk Difference (RD), without additional adjustment for parity.

$$P(POPP = 1) = \beta_0 + \beta_1 X$$

whereas β_1 captures the crude RD in POPP associated with treatment allocations. The crude analysis showed that women allocated to the intervention group had a lower probability of POPP compared with women receiving standard care ($RD = -0.21;\ 95\%CI\ -0.27,\ -0.15;\ p < 0.001$), corresponding to an absolute reduction of approximately 21% in the probability of persistent occiput posterior position ($POPP = 1$).

To evaluate whether any residual prognostic imbalance remained after stratified randomization, a second generalized linear model with identity link was subsequently fitted including both treatment allocation and parity.

$$P(POPP = 1) = \beta_0 + \beta_1 X + \beta_2 X$$

whereas β_1 estimated the adjusted RD in POPP associated with treatment allocations and $\beta_2$the association between POPP and parity, after adjustment for treatment allocation. The estimated treatment effect remained unchanged ($aRD = -0.21;\ 95\%CI\ -0.27, -0.15;\ p < 0.001$). As expected, parity remained independently associated with the outcome, with pluriparous women showing a lower probability of POPP compared with nulliparous women ($aRD = -0.23;\ 95\%CI\ -0.29,\ 0.17;\ p < 0.001$).

Unlike the previous trial generated through simple randomization, adjustment for parity had no impact on the estimated treatment effect after stratified randomization. This finding suggests that the stratified allocation procedure successfully minimized baseline imbalance in the prognostic variable, thereby improving comparability between treatment groups already at the design stage. Consequently, the crude treatment estimate obtained after stratified randomization was substantially similar to the adjusted estimate, consistent with the expectation of improved exchangeability between groups.

From a causal inference perspective, this example illustrates how stratified randomization may strengthen the internal validity of randomized trials when important prognostic variables are known a priori. By balancing clinically relevant baseline characteristics across treatment groups, stratified randomization reduces the probability that finite-sample imbalance may distort treatment effect estimation. In this sense, stratified randomization can be interpreted as a design-based strategy aimed at improving exchangeability before outcome occurrence and before statistical analysis is performed (Hernán & Robins, 2020).

This logic closely parallels the rationale underlying stratification-based causal inference methods used in observational studies. Methods such as standardization, inverse probability weighting (IPTW), and g-estimation attempt to recreate analytically the same comparability that randomization achieves by design. In both settings, causal inference relies on comparing exposed and unexposed individuals who are similar with respect to important prognostic factors. This conceptual continuity highlights a unifying principle across study designs: whether achieved through design-based strategy (i.e. stratified randomization) or analysis-based strategy (regression, G-estimation, Inverse Probability Treatment Weight), valid causal inference depends on creating or reconstructing exchangeability between comparison groups(Hernán & Robins, 2020).

3. Application of causal inference methods in observational midwifery research

3.1. Maternal birth satisfaction as a maternity health outcomes and care quality indicator

Over the last decades, maternal satisfaction with childbirth has progressively emerged as a relevant indicator of quality of maternity care and a key component of the broader concept of positive childbirth experience promoted by the World Health Organization (WHO) (World Health Organization, 2016, 2018). Within the WHO framework for quality maternal and newborn care, women's experience of care is considered equally important as clinical effectiveness and women safety, recognising that childbirth outcomes should not be evaluated exclusively in physical terms, but also through women's perceptions, emotional wellbeing, autonomy, and involvement in decision-making processes(World Health Organization, 2016).

Maternal birth satisfaction is a multidimensional construct influenced by interpersonal relationships with healthcare professionals, model of care, communication, perceived control during labour, intrapartum interventions, and physical and organisational characteristics of the birth setting (Lewis, Hauck, Ronchi, Crichton, & Waller, 2016; Srivastava, Avan, Rajbangshi, & Bhattacharyya, 2015). In addition to representing an important outcome itself, maternal satisfaction has been associated with several medium- and long-term maternal and neonatal outcomes, including psychological wellbeing, breastfeeding, future reproductive choices, and utilisation of maternity services (Joensuu et al., 2023; Nilvér, Begley, & Berg, 2017). Consequently, evaluating maternal satisfaction may contribute to identify strengths and critical issues within maternity care systems and to informing strategies aimed at improving women-centred care.

In the Italian context, childbirth occurs almost exclusively within hospital settings (Ministero della salute, 2025), which are organised according to different organisational and structural levels of care (Ministero della Sanità, 2010). Although the Italian healthcare system formally distinguishes between I and II level Maternity Units (MUs) based on organisational, technical, and structural criteria (Ministero della Sanità, 2010), limited evidence is available regarding the possible impact of these organisational differences on women's childbirth experience and satisfaction. Moreover, maternal satisfaction is currently not routinely included among the indicators monitored by the Italian maternal and neonatal surveillance systems, despite international recommendations supporting its inclusion as a quality outcome (World Health Organization, 2016).

For these reasons, the first study included in this thesis aimed to investigate the association between the organisational level of Maternity Units and maternal birth satisfaction in a large multicentric Italian cohort. The study represents an application of causal inference methods within an observational framework, addressing a clinically and organisationally relevant question in maternal healthcare research.

3.2. Methodological rationale: addressing confounding through IPTW

The research presented in this chapter was an observational study; therefore, women were not randomly allocated to I or II level Maternity Units. As discussed in Chapter 1, observational studies are particularly vulnerable to confounding because the distribution of prognostic characteristics may differ systematically between exposure groups (Hernán & Robins, 2020). In this context, women giving birth in different organisational settings were characterised by relevant differences in baseline obstetric and socio-demographic characteristics, particularly parity, educational level, and previous caesarean section.

Among these variables, parity represented a particularly important methodological issue (Figure 13). Nulliparous women were more frequently represented in II level Maternity Units, while pluriparous women were more common in I level settings. At the same time, parity is strongly associated with several dimensions of childbirth experience and maternal satisfaction, influencing labour duration, intrapartum interventions, perception of control, stress experienced during labour, and mode of birth (Fumagalli et al., 2021; Johansson & Finnbogadóttir, 2019). Consequently, a crude comparison of satisfaction scores between organisational levels would have risked reflecting differences in women's baseline characteristics rather than differences attributable to the maternity care setting itself.

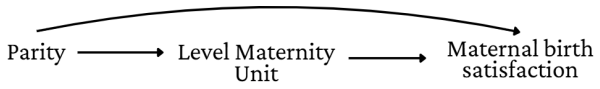


Figure 13

Unlike randomized controlled trials, where exchangeability between treatment groups is theoretically achieved through random allocation, in observational studies comparability between exposed and unexposed groups must be reconstructed analytically (Hernán & Robins, 2020). For this reason, the analysis adopted the Inverse Probability of Treatment Weighting (IPTW) approach based on propensity scores, previously introduced in Chapter 1 as one of the principal causal inference strategies for confounding control.

The IPTW method estimates, for each participant, the probability of receiving the observed exposure conditional on measured baseline covariates. Individuals are then weighted by the inverse of this probability, generating a pseudo-population in which the distribution of measured confounders becomes balanced across exposure groups(Hernán & Robins, 2020). In this simulated balanced population, the exposure assignment can be interpreted as conditionally independent from measured baseline characteristics, approximating some of the properties of randomized allocation.

The choice of IPTW in this study was particularly appropriate for several reasons. First, the multicentric design and large sample size provided adequate statistical power and sufficient overlap between exposure groups to support propensity score modelling. Second, the observed imbalance in key prognostic variables — especially parity — could not be resolved through design-based strategies because participant recruitment had already occurred within a real-world observational setting. Finally, IPTW allowed the inclusion of all participants in the analysis while improving covariate balance between organisational groups, thus enabling a less biased estimation of the association between maternity unit level and maternal satisfaction.

This chapter therefore provides a practical application of the methodological principles discussed in the previous chapter, illustrating how causal inference techniques can be used to strengthen internal validity and improve comparability between groups when randomization is not feasible (Hernán & Robins, 2020).

3.3. Study aim

The aim of this study was to evaluate whether the organisational level of Italian Maternity Units (I versus II level) was causally associated with maternal birth satisfaction among low-risk women giving birth in hospital settings.

3.4. Published article: 'Effect of Maternity Units' organisational levels on Maternal Birth Satisfaction: a multicentric cohort study'

INTRODUCTION

Worldwide, providing high-quality care during childbirth is one of the most important goals for health services[1]. According to the World Health Organization (WHO) all childbearing women should expect a positive birth experience as a recommended outcome, which can be achieved when the birth experience aligns with a woman's personal and sociocultural beliefs and expectations[2–4]. A positive birth experience contributes to maternal satisfaction and should be considered a key quality measure for evaluating and improving maternity healthcare services[2,3].

The WHO defines maternal satisfaction as the extent to which service standards meet expectations[1]. It is a multidimensional and subjective concept, influenced by parents' expectations about childbirth, interpersonal aspects of care (such as the model of care, communication style, and the presence of a birth companion), personal control, and self-efficacy[5–8].

Mother's satisfaction is correlated with long-term maternal and neonatal outcomes. A high satisfaction with the birth experience might have a positive impact on parental attitudes, included future birth choice[9]. While dissatisfaction is linked to increased rates of psychological diseases, breastfeeding difficulties, caesarean sections, and miscarriage in future pregnancies[7,10,11].

Birthplace is a crucial component of birth, which involves physical, emotional, cultural, and social aspects, with an impact on maternal childbirth experience[12]. Therefore, all dimensions of care, including structure, process, and outcome, have an influence on maternal satisfaction[13].

In high-income countries the access to alternative birth places varies within and between countries[14]. In Italy almost all women give birth in hospitals[15]. The Italian NHS provides universal and free of charge maternity care mostly in obstetric-led units and in few midwifery-led units[16,17]. Italian Maternity Units (MU) are classified into two levels of care (I Level and II Level) according to organisational, structural, and technical standards[18], fully described in Table 1. I Level Maternity Unit (MU) provides care to pregnant women over 34 weeks who do not require maternal and neonatal high-complex interventions, and they can choose where to give birth. II Level MUs provide care to all pregnant women regardless of their obstetric risk profile[18]. However, high risk women are usually referred to II level Mus.

Both I Level and II Level MUs are distributed across Italy, but their availability can vary significantly between regions. Urban areas, particularly those with larger hospitals, are more likely to have II Level MUs to handle complex cases. This distribution ensures that all pregnant women have access to appropriate care, though there may be differences in travel distance and accessibility based on location. Efforts are ongoing to balance the distribution of these units to meet regional healthcare demands effectively.

Although recommended by the WHO[1], birth satisfaction is not included in the Italian Maternal and Neonatal Outcomes Reporting System[15].

Several studies have explored how maternal birth experience can be affected by birth environments, including spatial and social dimensions[19,20], as well as midwifery care, and one-to-one care, staffing level, and intrapartum interventions[7,21–25].

Although the role of the quality of maternity services in ensuring a positive birth experience is strongly demonstrated[2,3], there is a gap of knowledge about the effect that the organisational level of MU might have on birth satisfaction.

Our study aimed to investigate if the organisational level (I vs II Level) of the MU had any impact on birth satisfaction in Italy.

METHODS

We conducted a multicentric cohort study. Women who gave birth in eleven maternity units, identified through the educational network of the north-Italian midwifery bachelor courses, were recruited.

In 2010 the Italian Ministry of Health published the State-Regions Agreement[18] containing 10 recommendations, including a re-organisation of care for maternity hospitals and the implementation of different pathways depending on obstetric risk profile. According to organisational, structural, and technical standards defined at national level (Table 1), five I Level MUs and six II Level MUs were enrolled. Supplementary materials describes the characteristics of each research site involved in the study.

Participants were recruited through a quota sampling method among women who gave birth in the MUs involved in the study from 1st January 2022 to 1st July 2022. Each I Level MU recruited 100 women, while each II Level MU enrolled 200 participants.

Table 1: Organisational, Structural and Technical standards that accurately describes I Level MU and II Level MU as defined by law [17]

	I Level MU	**II Level MU**
Target population	Gestational age ≥ 34 gestational week	All gestational age
	Low and Medium obstetric risk: excluding women or foetus with high-complexity interventions	Low, Medium and High obstetric risk: including women or foetus with high-complexity interventions
Number of births per year	500-1000	>1000
Organisational standards	Essential staffing 24/7: almost 2 midwives per shift until 1000 births presence of almost a gynecologist, a pediatrician and an anesthetists	Essential staffing 24/7: almost 3 midwives per shift if births/year<1500; 4 if <2000 and 5 if >2000 presence of almost 2 gynecologists, a pediatrician and a anesthetist
	Essential environment: almost 20 beds/1000 births 3 Labour suits 1 Operating room dedicated to obstetric emergencies 1 Delivery room dedicated to Low risk labour	Essential environment: almost 20 beds/1000 births 3 Labour suits (4 if births/year>2000) and 1 extra Labour suite 1 Operating room dedicated to obstetric emergencies (2 if births/year>1500) 1 Delivery room dedicated to Low risk labour

	Essential services: 1 ultrasound scanner in the Delivery Unit Sub-Intensive care Unit for women and newborns Laboratory tests, blood transfusion and diagnostic imaging available 24/7 Maternal Assisted Transport to Level II MU	Essential services: 1 ultrasound scanner in the Delivery Unit Intensive care Unit for women and newborns Laboratory tests, blood transfusion and diagnostic imaging available 24/7 Neonatal Assisted Transport from Level I MU Availability of additional specialist care to women (as psychological, neurological, ...)
Structural standards	Adequate connection between the Maternity Ward and the Delivery Unit Almost birth-equipment for 2 contemporary births Fetal monitor for each Labour suite Equipment for non pharmacological coping strategies Almost a room dedicated to post-partum women observation	Adequate connection between the Maternity Ward and the Delivery Unit Almost birth-equipment for 3 contemporary births Fetal monitor for each Labour suite Equipment for non pharmacological coping strategies Almost a room dedicated to post-partum women observation Dedicated room to anesthetic's counseling during pregnancy
Technical standards	Periodic evaluation of function and efficiency of every technologies available	Periodic evaluation of function and efficiency of every technologies available

The inclusion criteria were: healthy women with a straightforward pregnancy or pregnancies without major complications (such as cardiac disease, pre-eclampsia, haemoglobinopathies, renal disease, neurological disease, sepsis), aged between 18 and 50 years, with appropriate speaking and reading abilities, who gave birth at term (37-42 gestational weeks).

The exclusion criteria were: planned or pre-labour caesarean section, newborn in poor condition at birth or requiring resuscitation, foetal or neonatal death.

Women who met the inclusion criteria were invited to participate in the study, at least 24 hours after giving birth. Participants were informed about the study's aim and its voluntary and confidential nature. Informed consent was gained from all the participants.

Our expectation on the sample size of this observational study, considering previous data conducted in one of the participating centres[17], was about 500 women who gave birth in I Level MUs and about 1000 who gave birth in II Level MUs.

The observed sample size was a total of 1642 women. Of those, 470 gave birth in I level MUs and 1172 gave birth in II Level MUs. Considering the BSS-R total score ranging from 0 to 40 points, our sample enabled us to detect an effect size of 0.28 standard deviations with a significance level equal to 0.05 and a power equal to 0.90.

Maternal birth satisfaction was measured by the Italian version of Birth Satisfaction Scale Revised (I-BSS-R), an English self-report scale which was translated and adapted within the Italian context [11,26,27]. The I-BSS-R is a reliable and valid tool useful to assess maternal satisfaction with birth in the studied context[17].

The I-BSS-R consists of three sub-scales: Quality of Care Provided (QC), Women's personal Attributes (WA), and Stress Experienced during labour (SE). The QC sub-scale is made up of 4 questions that explore the support offered to women to promote their empowerment during labour, the communication's

quality with health workers and the birth's environment. The WA sub-scale is defined by 2 items regarding the perceptions of personal concerns about labour and birth and the loss of personal-control (both these items are reverse-coding). The SE sub-scale is made up of 4 questions that investigate concerns for herself in both physical and psychological dimension (as possible complications), having experienced long labour and the characteristic of intrapartum experience (stress and concerns during labour). Each of the ten items has 5 points and is scored using a Likert-type scale ranging from 0 to 4 (0 "strongly disagree", 4 "strongly agree"), and four of them are reverse-coded (e.g. "I found childbirth a distressing experience")[26].

The I-BBS-R scores were analysed considering both total score and each sub-scale[28]. The 25° centile of the I-BSS-R distribution (in our sample equal to 24) was used to define a new variable called "low satisfaction" to identify a group of women who on average were less satisfied than the study population.

Socio-demographic characteristics, obstetric history, intrapartum variables and maternal and postnatal outcomes were also explored.

Statistical analysis

The characteristics of the entire sample and the ones within Levels of MUs were described using frequencies and percentages, for categorical and discrete variables, and using summary indicators (mean and standard deviation (SD)), for continuous variables. Distribution's differences within levels were tested using Chi-square test (for categorical variables) and t-test (for continuous variables).

To assess the association between the levels of MUs and maternal satisfaction with birth, it was necessary to eliminate or reduce the confounding effects of confounding variables, which were related to the observational nature of the study. Confounding variables are those who influence the treatment (I Level MU and I Level MU) and outcome (maternal birth satisfaction). To reduce the confounding effects and investigate the association between levels of MUs and maternal satisfaction the inverse probability of treatment weighting method (IPTW)[29] was used: it reduces or eliminates the effects of confounding to estimate treatment effects using non-randomized data. The IPTW allowed us to create a pseudo-population where confounding variables were balanced; the confounding were defined within characteristics that were temporarily antecedent to the treatment, the choice of the place of birth, and that had different distribution within MU's levels: maternal age, university degree, previous caesarean-section and parity. The association between MU's levels and maternal satisfaction at birth (total score, each subscale, and the presence/absence of low satisfaction) was detected using balanced data from the IPWT method.

All tests performed were two-sided and a p-value$\leq$0.05 was considered statistically significant.

Statistical analyses were performed using Stata/MP 18.0.

Ethical consideration

The Local Ethics Committee approved this study before the beginning of the participants enrolment (MSM_2019, December 23, 2019). Written Informed consent was obtained from all participants.

RESULTS

Among the 1642 women recruited, 470 (28.6%) women gave birth in a I Level MU and 1172 (71.4%) in a II Level.

Women who gave birth in a II Level MU had higher levels of education (p-Value 0.01) and were mostly nulliparous (p-Value<0.001), compared to women who had access to a I Level MU. The frequency of the

previous caesarean section was higher across II Level MUs (p-Value 0.012). No differences within the antenatal care were identified between groups. Among nulliparous, women who gave birth in a I Level MU, more often attended antenatal classes (p-Value <0.001). (Table 2)

Women who gave birth in a I level MU received a different intrapartum care compared to women who deliver in a II level MU (Table 2). One-to-one midwifery care, non-pharmacological strategies to cope with pain and skin to skin contact were more likely to be provided in I level MUs. Moreover, lower levels of intrapartum interventions rates (epidural analgesia, oxytocin augmentation and episiotomy) were reported in I level MUs compared to the care provided in II Level MUs.

A total of 1425 women (86.79%) had a spontaneous vaginal birth, 124 (7.55%) women had a caesarean section, and 93 had an instrumental vaginal birth (n=93, 5.66%). No significant differences were identified between levels of MU (Table 2).

Table 2: distribution of socio-demographic variables, obstetric history, antenatal care, maternal outcomes and post-natal care in the entire sample and within MU's levels.

* Percentages calculated within women who had spontaneous or vacuum assisted vaginal birth (n=1518)

			Overall (n=1642)		**I Level MU (n=470)**		**II Level MU (n=1172)**		**p-Value**
			mean	***SD***	***mean***	***SD***	***mean***	***SD***	
Socio demographic variables	Maternal age (years)		32.91	4.96	32.29	4.65	33.17	5.06	0.001
			n	***%***	***n***	***%***	***n***	***%***	
	University degree		827	50.37	213	45.32	614	52.39	0.01
	Employed		1274	77.59	353	75.11	921	78.58	0.127
Obstetric history	Primiparous		855	52.07	199	42.34	656	55.97	<0.001
	Previous Vaginal Instrumental birth		39	2.38	8	1.7	31	2.65	0.257
	Previous Cesarean Section		71	4.32	11	2.34	60	5.12	0.012
Antenatal care	Antenatal class attending		736	44.82	227	48.3	509	43.43	0.073
	Public care in pregnancy		710	43.43	187	39.79	523	44.89	0.059
	Midwife presence during pregnancy care		450	27.41	128	27.23	322	27.47	0.921
Intrapartum care	Spontaneous onset of labour		1133	69.04	345	73.4	788	67.29	0.015
	One to one		1314	80.12	446	95.1	868	74.12	<0.001
	Non Pharmacological analgesia		1100	66.99	347	73.83	753	64.25	<0.001
	Intermittent auscultation of fetal hearth		409	24.94	71	15.11	338	28.89	<0.001
	Prologed active phase of first stage (>=12 hours)		58	3.54	22	4.69	36	3.07	0.109
	Augmentation with Oxytocin		559	34.09	129	27.45	430	36.75	<0.001
	Epidural analgesia		711	43.33	159	33.9	552	47.1	<0.001
Maternal outcomes	Mode of birth	Spontaneous	1425	86.78	409	87.02	1016	86.69	0.84
		Vacuum assisted	93	5.66	28	5.96	65	5.55	
		Caesarean section	124	7.55	33	7.02	91	7.76	
	Episiotomy*		225	14.82	48	10.98	177	16.37	0.007
	Perineal integrity*		353	23.25	119	27.23	234	21.65	0.02

	Spontaneous delivery of placenta	1517	92.44	436	92.96	1081	92.24	0.614
Post-natal Care	Skin to skin	1422	86.6	421	89.57	1001	85.41	0.025
	First breastfeed until 1 hour	1399	85.2	395	84.04	1004	85.67	0.403

The comparison between and within levels of MUs are reported in Figure 1. The mean of the I-BSS-R total score was 27.35 (SD=5.46), with higher values in women who gave birth in a I Level MU (27.92 versus 27.12) (Table 3). Overall, the score of the three sub-scales contributed differently to the total score (Table 3): the mean of the QC sub-scale was 13.98, the maximum achievable is 16; while the mean of the WA sub-scale was 4.72, the maximum achievable is 8 and the mean of the SE sub-scale was 8.65, the maximum achievable is 16.

Table 3: I-BSS-R total score and subscale in the entire sample and within levels

	Item (N)	**Overall (n=1642)**		**I Level MU (n=470)**		**II Level MU (n=1172)**	
		mean	***SD***	***mean***	***SD***	***mean***	***SD***
I-BSS-R	10	27.35	5.46	27.92	5.39	27.12	5.48
Sub-Scale							
Quality of care provision	4	13.98	2.02	14.26	1.86	13.87	2.07
Women's personal attributes	2	4.72	1.95	4.67	1.98	4.74	1.94
Stress experienced	4	8.65	3.24	8.99	3.29	8.51	3.21

After propensity score adjustment, maternal satisfaction was similar within women who gave birth in I Level and II Level MUs (mean 27.70 versus mean 27.20; p-Value 0.096) (Table 4). Of note, the QC sub-scale was significantly higher in women who gave birth in a I Level MU (mean 14.28; CI 95%: 14.1;14.46) compared to women who gave birth in a II Level MU (mean 13.87; CI 95%: 13.65;14.09) (p-Value <0.001) (Table 4). No differences between levels were identified for both the WA (p-Value 0.159) and the SE (p-Value 0.164) sub-scales (Table 4). In the overall sample, 386 women have been included in the "low satisfaction" group, with a BSS-R total score lower than 24. The probability of being less satisfied than the average was higher in women who gave birth in II Level MUs (24.70% versus 20.28%; p-Value 0.057) (Table 4).

Table 4: I-BSS-R (total score and sub-scales) within I and II Level MUs with balanced data (Propensity score weighting)

	I Level MU		**II Level MU**		**p-Value**
	mean	*95% IC*	*mean*	*95% IC*	
I-BSS-R	27.7	[27.19;28.21]	27.2	[26.61; 27.79]	0.096
Sub-Scale					
Quality of care provision	14.28	[14.1;14.46]	13.87	[13.65;14.09]	<0.001
Women's personal attributes	4.6	[4.41;4.78]	4.75	[4.54;4.97]	0.159
Stress experienced	8.83	[8.52; 9.13]	8.58	[8.23;8.93]	0.164

DISCUSSION

Our study is the first exploring the effect of the organisational level of the MU on maternal birth satisfaction within the Italian context, where births occur almost exclusively in hospital settings[18,30]. The impact of birthplaces on maternal experience has been widely explored in literature[31–33]. Our study, which

focused on the role of organisational, structural, and technical aspects of MUs, contributes to increasing knowledge about birth satisfaction in hospital settings.

The main strength of the present study is its multicenter nature, with eleven participating research sites, including a large sample size, which ensures the generalizability of our results for healthy pregnant women. Furthermore, the power of our study is enhanced by the use of the IPTW method[29] that minimises or eliminates the confounders revealed by the literature[6,17,34], resulting in a more accurate evaluation of the impact of the organisational level on maternal birth satisfaction.

One of the main limitations of our study is the exclusion of high-risk women and other vulnerable groups from the target population. This exclusion may affect the generalizability of our findings, especially given the significant representation of these groups in our country[35]. Therefore, future research should aim to include women with higher risk profiles to provide a more comprehensive understanding of birth satisfaction across different populations[2,3,36]. Including these groups in future studies may either reinforce the trends observed in low-risk women or reveal new factors influencing maternal satisfaction, thus offering valuable insights for improving maternity care services for all women.

In addition, future research exploring Italian mothers' experiences within different birth settings, such as MLU and home, is warranted. Another interesting topic might be to investigate determinants of maternal birth satisfaction.

We found an average score of the I-BSS-R that falls within the range reported across similar studies[6,24,37–39], which is between 24.2[24] and 31.94[6]. When focusing on each sub-scale, similarities and differences were found compared to the literature. The average score of our QC sub-scale was consistent with the one reported by previous authors[6,37–39]. When considering the average score of the SE sub-scale, it turned out to be lower than prior evidence[6,37–39]. In contrast, the average score of the WA sub-scale was similar to the one reported by the same cited studies[37–39] but lower than the one evaluated in the USA study by Fleming et al.[6], who recruited women from different birth settings.

Our results showed that MU's organisational level did not affect maternal birth satisfaction, with similar I-BSS-R total scores within I Level and II Level MUs. This finding might be explained by a similar distribution of modes of birth observed in both I and II levels MUs. Mode of birth has been suggested as one of the factors that impact most on maternal birth satisfaction[6,22,34]. On the contrary, we observed that Level I and II MUs adopt different approaches to intrapartum care, with midwifery care in level I MUs being more focused on promoting normal labor and birth [16,17,33,40]. Although this was not the primary aim of our study, our data indicated that women in I Level MUs were more likely to experience spontaneous labor onset, use non-pharmacological pain relief, receive one-to-one midwifery care and were less likely to undergo labor augmentation. These findings align with previous research in similar context[35]. Moreover, existing studies suggest that healthy women giving birth in high birth volume hospitals might be more exposed to interventions such as epidural analgesia and episiotomy. This could be due to staff shortages, time pressure from simultaneous births, or the higher frequency of complicated cases, which may lead to overtreatment of low-risk women[41,42].

Another interesting finding is the one showing that women who gave birth in II level MUs were more likely to be included in the "low satisfaction" group, compared to participants who gave birth in I level MUs. Our hypothesis is that the above mentioned results might be linked. In fact, midwives who routinely look after both low and high-risk women in II level MUs, are exposed to increasing amounts of interventions that might also influence their attitude towards childbirth care. This higher perception of risk perceived by midwives could lead to a higher provision of unnecessary obstetric interventions. Thus, contributing to a medicalisation of birth. This might be the reason why women in II Level MUs, who

were more likely to experience interventions, were also at higher risk of being less satisfied than women who give birth in a I level MU. Midwives working in I level MUs should be more familiar with a health-promoting approach, supporting normal birth. Hence, Midwife-Led-Unit should be considered one of the evidence-based strategies to improve several maternal and neonatal outcomes, including maternal satisfaction with birth.

The aspects mentioned above, related to midwifery practice, had an influence on the three sub-scales of the total I-BSS-R score. The QL, WA and SE sub-scales assess different dimensions related to maternal satisfaction with birth. The comparison of each sub-scale between I and II Level MUs, enables us to deeply explore the impact of the organisational level on birth experience.

When we considered each sub-scale separately, the QC sub-scale had a greater score when the birth occurred in a I level MU. Again, this finding might be explained by the hypothesis that I Level MUs are more likely to promote all aspects investigated by the QC sub-scale, including women's empowerment, support by professional staff, and presence of a comfortable environment[26]. The literature suggests that one-to-one midwifery care, non-pharmacological coping strategies and skin-to-skin contact, have a positive impact on the quality of care and in turn on maternal birth satisfaction[16,33,40]. In this regard, we found that those procedures were offered more often in I level MUs. Thus, contributing to increasing the probability of being in the "low satisfaction" group in women who gave birth in II level MUs. Unlike the QC sub-scale, the MU's organisational level didn't affect the scores of the SE and the WA sub-scales, with mean values that were similar in both I and II Level MUs. The SE sub-scale assesses the physical and psychological stress due to prolonged labour, vacuum-assisted birth and caesarean section[26]. The similar approach and adoption of these intrapartum interventions within settings could explain the absence of difference in the average scores of the SE sub-scale between MUs levels.

Moreover, the WA sub-scale explores maternal competencies, such as women's concerns about labour and perception of control[26]. According to the literature, midwifery antenatal care and continuity models of care have a key role in counselling, they promote an ongoing individual adjustment of maternal expectation of birth that facilitates women's decision-making, women's awareness, emotional support, and, in turn, a higher maternal satisfaction[43–46]. In our study, the characteristics of the antenatal care offered to women were very similar regardless of the level of the MU where they gave birth, including the type of healthcare professionals providing the care, characteristics of the setting where the birth occurred (public or private hospitals) and rate of antenatal education attendance. In addition, the model of care provided to women during the childbearing pathway was not included as one of the criteria to define the difference between organisational levels of maternity units in the Italian State-Regions Agreement[18]. All these considerations might explain why no significant differences were observed in the WA sub-scale average scores between the two levels of MUs.

The average scores of the SE and the WA sub-scales were significantly below the maximum achievable score. Based on these findings, maternity healthcare professionals should consider the items measured by both the SE and the WA sub-scales to improve women's birth experience. For instance, based on the knowledge that intrapartum interventions have an impact on the SE sub-scale[26], the promotion of the normal progress of labour should be supported as a potential strategy to enhance maternal satisfaction. Again, the development of a partnership relationship between women and midwives might positively affect aspects investigated by the WA sub-scale[43–46]. The present study could be an opportunity to identify which components of care deserve attention, thus facilitating the implementation of effective strategies or interventions to promote satisfaction, regardless of the organisational standards. Most of the concepts that are embedded in the BSS-R sub-scales are strongly supported by both the midwifery care models of care

and the philosophy promoted by Midwife Led Units (MLU)[47]. For this reason, MLU should be offered to women with straightforward pregnancies within the Italian health services. In practical terms, strategies such as continuous one-to-one midwifery care throughout labor, the use of non-pharmacological pain relief methods (e.g., breathing techniques, hydrotherapy, or massage), and encouraging women to actively participate in decision-making during labor have been shown to improve maternal satisfaction and overall birth outcomes. Evidence suggests that these approaches promote a more positive birth experience reducing the need for unnecessary interventions. Furthermore, ensuring adequate staffing levels to allow personalized care and training healthcare professionals to adopt a woman-centered approach are crucial for successfully implementing these strategies in everyday clinical practice.

Practical implications

- Provide evidence-based care regardless of the healthcare setting, ensuring that all women receive high-quality, individualized care.
- Ensure adequate staffing levels, access to necessary equipment, and ongoing training for healthcare professionals to maintain high standards of care.
- Utilize validated tools for risk assessment to accurately identify low-risk women and tailor care accordingly.
- Empower women with information about their care options and the potential benefits and risks associated with different interventions.
- Develop and distribute educational materials, offer counselling sessions, and engage women in discussions about their preferences and choices during antenatal visits.
- Advocate for policy changes to integrate MLUs within the healthcare system and ensure they are adequately funded and staffed.

By incorporating these practical strategies into clinical practice, maternal health professionals can contribute to more positive birth experiences and improve overall maternal satisfaction.

CONCLUSION

The present study showed that the level of the MU, defined using organisational, structural, and technical criteria, did not impact on maternal satisfaction with birth. However, the classification used to define the level of an Italian MU, does not consider the model of care provided to the women, which is considered to have an important influence on maternal satisfaction.

This study suggests key strategies to address high quality maternal and newborn healthcare, which might contribute to promote a positive experience of birth for all women. The key role of midwifery models of care in helping women to achieve empowerment, high perception of control and the best health outcomes, should be acknowledged. This study contributes to understanding how the level of the MU, defined by organisational, structural, and technical standards, might impact on women's birth satisfaction. In addition, this research might provide the basis for future investigations regarding the role of variables affecting the I-BSS-R. In particular, we should focus on factors influencing the WA and the SE sub-scales, aiming at identifying strategies or interventions to promote a positive birth experience.

REFERENCES

1. World Health Organisation. *Standard for Improving Quality of Maternal and Newborn Care in Health Facilities.*; 2016.
2. World Health Organization. *WHO Recommendations. Intrapartum Care for a Positive Childbirth Experience.*; 2018.

3. World Health Organization. *WHO Recommendations on Maternal and Newborn Care for a Positive Postnatal Experience.*; 2022.
4. Lazzerini M, Mariani I, Semenzato C, Valente EP. Association between maternal satisfaction and other indicators of quality of care at childbirth: a cross-sectional study based on the WHO standards. *BMJ Open*. 2020;10(9):e037063. doi:10.1136/bmjopen-2020-037063
5. Christiaens W, Bracke P. Assessment of social psychological determinants of satisfaction with childbirth in a cross-national perspective. *BMC Pregnancy Childbirth*. 2007;7. doi:10.1186/1471-2393-7-26
6. Fleming SE, Donovan-Batson C, Burduli E, Barbosa-Leiker C, Hollins Martin CJ, Martin CR. Birth Satisfaction Scale/Birth Satisfaction Scale-Revised (BSS/BSS-R): A large scale United States planned home birth and birth centre survey. *Midwifery*. 2016;41:9-15. doi:10.1016/j.midw.2016.07.008
7. Ford JB, Hindmarsh DM, Browne KM, Todd AL. Are women birthing in New South Wales hospitals satisfied with their care? *BMC Res Notes*. 2015;8(1). doi:10.1186/s13104-015-1067-2
8. Lewis L, Hauck YL, Ronchi F, Crichton C, Waller L. Gaining insight into how women conceptualize satisfaction: Western Australian women's perception of their maternity care experiences. *BMC Pregnancy Childbirth*. 2016;16(1). doi:10.1186/s12884-015-0759-x
9. Joensuu JM, Saarijärvi H, Rouhe H, et al. Effect of the maternal childbirth experience on a subsequent birth: A retrospective 7-year cohort study of primiparas in Finland. *BMJ Open*. 2023;13(3). doi:10.1136/bmjopen-2022-069918
10. Nilvér H, Begley C, Berg M. Measuring women's childbirth experiences: A systematic review for identification and analysis of validated instruments. *BMC Pregnancy Childbirth*. 2017;17(1). doi:10.1186/s12884-017-1356-y
11. Nespoli A, Colciago E, Fumagalli S, Locatelli A, Hollins Martin CJ, Martin CR. Validation and factor structure of the Italian version of the Birth Satisfaction Scale-Revised (BSS-R). *J Reprod Infant Psychol*. 2021;39(5):516-531. doi:10.1080/02646838.2020.1836333
12. Grigg C, Tracy SK, Daellenbach R, Kensington M, Schmied V. An exploration of influences on women's birthplace decision-making in New Zealand: A mixed methods prospective cohort within the Evaluating Maternity Units study. *BMC Pregnancy Childbirth*. 2014;14(1). doi:10.1186/1471-2393-14-210
13. Srivastava A, Avan BI, Rajbangshi P, Bhattacharyya S. Determinants of women's satisfaction with maternal health care: A review of literature from developing countries. *BMC Pregnancy Childbirth*. 2015;15(1). doi:10.1186/s12884-015-0525-0
14. Scarf VL, Rossiter C, Vedam S, et al. Maternal and perinatal outcomes by planned place of birth among women with low-risk pregnancies in high-income countries: A systematic review and meta-analysis. *Midwifery*. 2018;62:240-255. doi:10.1016/j.midw.2018.03.024
15. Ministero della Salute. *Certificato Di Assistenza al Parto (CeDAP). Analisi Dell'evento Nascita. Anno 2022.*; 2023. www.salute.gov.it/statistiche
16. Borrelli S, Fumagalli S, Colciago E, Downey J, Spiby H, Nespoli A. How should a video-call service for early labour be provided? A qualitative study of midwives' perspectives in the United Kingdom and Italy. *Women and Birth*. Published online November 1, 2023. doi:10.1016/j.wombi.2023.06.006

17. Fumagalli S, Colciago E, Antolini L, Riva A, Nespoli A, Locatelli A. Variables related to maternal satisfaction with intrapartum care in Northern Italy. *Women and Birth*. 2021;34(2):154-161. doi:10.1016/j.wombi.2020.01.012
18. Ministero della Salute. Accordo Stato-Regioni 16 dicembre 2010. Published online 2010.
19. Bellini E, Macchi A, Setola N, Lindahl G. Sensory Design in the Birth Environment: Learning from Existing Case Studies. *Buildings*. 2023;13(3). doi:10.3390/buildings13030604
20. Migliorini L, Setola N, Naldi E, Rompianesi MC, Iannuzzi L, Cardinali P. Exploring the Role of Birth Environment on Italian Mothers' Emotional Experience during Childbirth. *Int J Environ Res Public Health*. 2023;20(15). doi:10.3390/ijerph20156529
21. Fenaroli V, Molgora S, Dodaro S, et al. The childbirth experience: Obstetric and psychological predictors in Italian primiparous women. *BMC Pregnancy Childbirth*. 2019;19(1). doi:10.1186/s12884-019-2561-7
22. Johansson C, Finnbogadóttir H. First-time mothers' satisfaction with their birth experience – a cross-sectional study. *Midwifery*. 2019;79. doi:10.1016/j.midw.2019.102540
23. Inversetti A, Fumagalli S, Nespoli A, et al. Childbirth experience and practice changing during COVID-19 pandemic: A cross-sectional study. *Nurs Open*. 2021;8(6):3627-3634. doi:10.1002/nop2.913
24. Mauri PA, Pilloni E, Stillavato S, et al. Role of mode of induction and delivery and women's satisfaction after induction of labour at term: An observational study. *European Journal of Obstetrics and Gynecology and Reproductive Biology*. 2023;286:47-51. doi:10.1016/j.ejogrb.2023.05.012
25. Montori S, Lugli F, Monesi M, et al. Telemedicine in the treatment of gestational diabetes: An observational cohort study on pregnancy outcomes and maternal satisfaction. *Diabetic Medicine*. Published online 2023. doi:10.1111/dme.15201
26. Hollins Martin CJ, Martin CR. Development and psychometric properties of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*. 2014;30(6):610-619. doi:10.1016/j.midw.2013.10.006
27. Nespoli A, Colciago E, Pedroni S, Perego S, Fumagalli S. The birth satisfaction scale-revised (BSS-R): Process of translation and adaptation in an Italian context. *Ann Ist Super Sanita*. 2018;54(4):340-347. doi:10.4415/ANN_18_04_11
28. Martin CR, Hollins Martin CJ, Burduli E, Barbosa-Leiker C, Donovan-Batson C, Fleming SE. The Birth Satisfaction Scale–Revised (BSS-R): should the subscale scores or the total score be used? *J Reprod Infant Psychol*. 2018;36(5):530-535. doi:10.1080/02646838.2018.1490498
29. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424. doi:10.1080/00273171.2011.568786
30. Comitato Percorso Nascita nazionale. *Linee Di Indirizzo per La Definizione e l'organizzazione Dell'assistenza in Autonomia Da Parte Delle Ostetriche Alle Gravidanza a Basso Rischio*.; 2017.
31. Brocklehurst P, Hardy P, Hollowell J, et al. Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: The Birthplace in England national prospective cohort study. *BMJ (Online)*. 2012;343(7840). doi:10.1136/bmj.d7400
32. Olsen O, Clausen JA. Planned hospital birth compared with planned home birth for pregnant women at low risk of complications. *Cochrane Database of Systematic Reviews*. 2023;2023(3). doi:10.1002/14651858.CD000352.pub3

33. Takayama T, Phongluxa K, Nonaka D, et al. Is the place of birth related to the mother's satisfaction with childbirth? A cross-sectional study in a rural district of the Lao People's Democratic Republic (Lao PDR). *BMC Pregnancy Childbirth*. 2019;19(1). doi:10.1186/s12884-019-2483-4
34. Hochman N, Galper A, Stanger V, et al. Risk factors for a negative birth experience using the Birth Satisfaction Scale-Revised. *International Journal of Gynecology and Obstetrics*. Published online 2023. doi:10.1002/ijgo.14884
35. Fumagalli S, Nespoli A, Panzeri M, et al. Intrapartum Quality of Care among Healthy Women: A Population-Based Cohort Study in an Italian Region. *Int J Environ Res Public Health*. 2024;21(5). doi:10.3390/ijerph21050629
36. Colciago E, Merazzi B, Panzeri M, Fumagalli S, Nespoli A. Women's vulnerability within the childbearing continuum: A scoping review. *Eur J Midwifery*. 2020;4(May). doi:10.18332/ejm/120003
37. Romero-Gonzalez B, Peralta-Ramirez MI, Caparros-Gonzalez RA, Cambil-Ledesma A, Hollins Martin CJ, Martin CR. Spanish validation and factor structure of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*. 2019;70:31-37. doi:10.1016/j.midw.2018.12.009
38. Grundström H, Martin CJH, Malmquist A, Nieminen K, Martin CR. Translation and validation of the Swedish version of the Birth Satisfaction Scale-Revised (BSS-R). *Midwifery*. 2023;124. doi:10.1016/j.midw.2023.103745
39. Mollard E, Kupzyk K. *Birth Satisfaction during the Early Months of the COVID-19 Pandemic in the United States*. Vol 47.; 2022. http://links.lww.com/
40. Hodnett ED, Downe S, Walsh D. Alternative versus conventional institutional settings for birth. *Cochrane Database of Systematic Reviews*. 2012;2012(8). doi:10.1002/14651858.CD000012.pub4
41. Pfniss I, Gold D, Holter M, et al. Birth during off-hours: Impact of time of birth, staff´s seniority, and unit volume on maternal adverse outcomes—a population-based cross-sectional study of 87 065 deliveries. *Birth*. 2023;50(2):449-460. doi:10.1111/birt.12663
42. Zinsser LA, Stoll K, Gross MM. Midwives' attitudes towards supporting normal labour and birth - A cross-sectional study in South Germany. *Midwifery*. 2016;39:98-102. doi:10.1016/j.midw.2016.05.006
43. Colciago E, Brown AM, Ornaghi S, et al. The unpredictable nature of labour and birth: A qualitative investigation regarding expectations of Italian parents to-be. *Midwifery*. 2022;108. doi:10.1016/j.midw.2022.103286
44. Sandall J, Soltani H, Gates S, Shennan A, Devane D. Midwife-led continuity models versus other models of care for childbearing women. *Cochrane Database of Systematic Reviews*. 2016;2016(4). doi:10.1002/14651858.CD004667.pub5
45. Perriman N, Davis DL, Ferguson S. What women value in the midwifery continuity of care model: A systematic review with meta-synthesis. *Midwifery*. 2018;62:220-229. doi:10.1016/j.midw.2018.04.011
46. Forster DA, McLachlan HL, Davey MA, et al. Continuity of care by a primary midwife (caseload midwifery) increases women's satisfaction with antenatal, intrapartum and postpartum care: Results from the COSMOS randomised controlled trial. *BMC Pregnancy Childbirth*. 2016;16(1). doi:10.1186/s12884-016-0798-y
47. Rocca-Ihenacho L, Batinelli L, Thaels E, et al. *Midwifery Unit Standards*.; 2018.

4. Application of causal inference principles in randomized midwifery research

4.1. Management of persistent occiput posterior position in midwifery care

Foetal occiput posterior position (OPP) is the most frequent foetal malposition during labour, occurring in approximately $30 - 40\%$ of women at the onset of active labour. Although most fetuses spontaneously rotate to an occiput anterior position (OAP), a proportion persist as persistent occiput posterior position (POPP), particularly during the second stage of labour, when spontaneous rotation becomes unlikely (Barrowclough et al., 2022; Gardberg, Laakkonen, & Sälevaara, 1998). POPP has been consistently associated with prolonged labour, increased rates of operative vaginal birth and caesarean section, higher incidence of severe perineal trauma, and adverse neonatal outcomes, including low Apgar scores, acidemia at birth, neonatal intensive care unit admission, and neonatal encephalopathy (Barrowclough et al., 2022; Cheng et al., 2006).

Several maternal and intrapartum factors have been associated with an increased risk of OPP and POPP, including advanced maternal age, nulliparity, high body mass index, epidural analgesia, fetal macrosomia, anterior placenta, and pelvic anatomical characteristics such as narrow pubic arch angles (Caughey et al., 2015; Cheng et al., 2006; Ghi et al., 2016). These factors may influence fetal head engagement and rotation dynamics during labour progression. In particular, nulliparity represents one of the strongest prognostic factors associated with persistent malposition, likely because of differences in labour progression, uterine dynamics, and pelvic adaptation during the first childbirth experience (Barrowclough et al., 2022; Fumagalli et al., 2024). From a causal inference perspective, parity can be conceptualized as a baseline prognostic factor associated with the outcome independently of treatment allocation. The hypothesized causal structure of the trial can therefore be represented as follows (Figure 14):

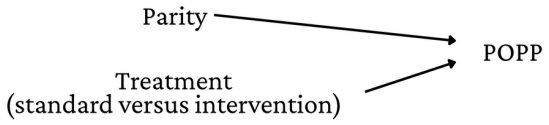


Figure 14

Within this clinical context, several non-pharmacological interventions have been proposed to facilitate fetal rotation from occiput posterior to occiput anterior position during labour. Among these, maternal postural techniques and the use of rebozo manoeuvres have gained increasing attention in midwifery practice because of their non-invasive nature and their potential role in promoting physiological labour progression. Maternal positioning strategies aim to modify pelvic diameters and maternal biomechanics in order to facilitate fetal rotation, while rebozo techniques involve the use of a traditional woven cloth to perform gentle mobilization and relaxation manoeuvres during labour(Cohen & Thomas, 2015; Iversen, Midtgaard, Ekelin, & Hegaard, 2017). However, despite growing clinical interest, evidence regarding the effectiveness of these interventions remains limited and heterogeneous (Barrowclough et al., 2022).

For this reason, the second study included in this thesis focused on evaluating the effectiveness of maternal postures and rebozo techniques in reducing the risk of persistent occiput posterior position at birth among women presenting with OPP during labour. Unlike the observational study presented in the

previous section, this research question was investigated through a randomized controlled trial (RCT), allowing a different application of causal inference principles within experimental research.

4.2. Methodological rationale: exchangeability and residual imbalance in randomized studies

Randomized controlled trials are considered the gold standard for causal inference because random allocation of treatment theoretically ensures exchangeability between intervention groups (Hernán & Robins, 2020). Under ideal conditions, randomization distributes both measured and unmeasured prognostic factors equally across treatment arms, thereby minimizing confounding and allowing observed differences in outcomes to be interpreted causally. In contrast to observational studies, where causal analyses attempt to emulate a hypothetical target trial, RCTs directly implement the target trial through prospective treatment allocation and protocol-defined follow-up.

Nevertheless, as discussed in the previous section of this chapter, randomization does not guarantee perfect balance in every realized sample. Especially in studies with relatively limited sample size, important prognostic variables may remain unevenly distributed between groups purely by chance. When such variables are strongly associated with the outcome, residual imbalance may influence the precision and interpretation of treatment effect estimates.

In the context of persistent occiput posterior position, parity represents a particularly important prognostic variable. Nulliparous women have consistently shown higher risks of persistent malposition, prolonged labour, operative birth, and intrapartum interventions compared with pluriparous women (Barrowclough et al., 2022; Caughey et al., 2015; Cheng et al., 2006; Fumagalli et al., 2024). Consequently, even within a randomized framework, imbalance in parity distribution between intervention groups may affect the estimation of the intervention effect.

From a causal inference perspective, parity can therefore be represented as a baseline prognostic factor associated with the outcome independently of treatment allocation. Although randomization blocks systematic confounding by ensuring that treatment assignment is independent of baseline covariates, parity may still contribute to random variability in outcome occurrence across groups if not adequately balanced.

The causal structure underlying this research question can be formalised through a Directed Acyclic Graph (DAG), in which parity acts as a prognostic factor influencing the probability of persistent occiput posterior position, while remaining independent of treatment assignment because of randomization. Within the causal inference framework, the DAG serves not only as a graphical representation of assumptions, but also as a tool to align the design, conduct and analysis of the study. The identification of parity as a clinically relevant determinant of the outcome informed the choice of stratified randomization at the design stage and guided the interpretation of treatment effect estimates during the analysis. This illustrates how valid causal inference relies on the coherence between the hypothesised causal model, the study design and the analytical strategy adopted.

For this reason, the trial presented in this thesis adopted stratified randomization according to parity status. The stratified allocation sequence was generated before participant enrolment and implemented

separately within parity strata. The code used to generate the randomization schedule is reported in Appendix 1. Women were first stratified as nulliparous or pluriparous and subsequently randomized within each stratum to either the intervention group or standard care. This allocation strategy aimed to improve comparability between treatment groups already at the design stage, minimizing the probability that imbalance in parity could distort treatment effect estimation (Hernán & Robins, 2020).

As discussed earlier in this chapter, stratified randomization can be interpreted as a design-based analogue of stratification methods used in observational causal inference. Whereas methods such as IPTW and standardization attempt to reconstruct exchangeability analytically after data collection, stratified randomization seeks to guarantee comparability prospectively before outcome occurrence. In both settings, the underlying causal principle remains the same: valid causal inference depends on comparing groups that are exchangeable with respect to relevant prognostic characteristics (Hernán & Robins, 2020).

4.3. Study aim

The aim of this randomized controlled trial was to evaluate whether maternal postural techniques combined with rebozo manoeuvres during labour could reduce the risk of persistent occiput posterior position at birth compared with standard intrapartum care among women presenting with fetal occiput posterior position during labour.

4.4. Published article: 'Rebozo and maternal postures to prevent persistent occiput posterior position of the fetal head: protocol for a randomized clinical trial - the ReMaP-POPP RCT -

INTRODUCTION

Background and rationale

Occiput posterior position (OPP), i.e., when the back of the fetal head lies posteriorly in the mother's pelvis, is the most common fetal malposition, being identified in 35% of laboring women[1,2].

Probabilities of OPP in labor vary among studies on the topic due to differences in OPP assessment, with research employing the gold standard technique of transabdominal suprapubic sonography alongside or alternatively to vaginal examination having the highest accuracy[3-5].

Overall, the probability of OPP has been estimated to be 35% in the early phase of active labor, 20-30% at full cervical dilatation, and 5-13% at delivery[1,2].

The decrease of OPP probability throughout labor till delivery is due to spontaneous rotation of OPP fetuses in an occiput anterior position (OAP), the most favorable for labor and birth, usually before the start or at the beginning of the second stage of labor[5-8]. Spontaneous rotations are unlikely once the second stage has begun. The vast majority of OPP deliveries results from the failure of rotation from this position[3-5].

Persistent OPP (POPP) during the second stage of labor has been shown to increase the risk of both maternal and neonatal adverse outcomes.

It exposes to prolonged labor, operative abdominal and vaginal delivery, high-degree perineal lacerations, and postpartum hemorrhage. Also, POPP has been related to low Apgar scores and cord blood pH values at birth, admission to neonatal intensive care unit, and neonatal encephalopathy[2,9-17]. Importantly, traumatic births, such as those that can be more frequently observed among women with POPP, have been associated with decreased birth satisfaction and higher rates of impaired post-partum emotional health[18].

Several factors have been recognized to associate with heightened odds of POPP, including loose abdomen (i.e. reduced muscular tone or laxity of the abdominal wall), nulliparity, late- and post-term pregnancy, fetal macrosomia, high maternal body mass index, epidural analgesia, untimely amniotomy, anterior positioned placenta, and narrow pubic arch angle typical of the android and anthropoid maternal pelvis[5,16,19-25].

Different interventions have been evaluated to favor anterior rotation of an OPP fetus during labor, such as maternal posturing[26-28] and manual or instrumental rotation of the fetal head[29-35].

Both manual and instrumental rotation are associated with high success rates but can be performed only at the end of the first or during the second stage of labor and may be associated with fetal and maternal complications. Also, they require the provider's skill and confidence[30-39]. On the other hand, maternal postures, such as hands-and-knees and lateral recumbent position, seek to non-invasively promote flexion of the fetal head to favor its spontaneous rotation into the OAP through changes in the pelvic angles and diameters and forces of gravity and buoyancy[40-42].

Several randomized clinical trials investigating hands-and-knees and lateral recumbent positions for increasing the probability of OAP in laboring women with an OPP fetus have been completed in the last twenty years, with contrasting results[18,27,28,40,43-51]. In a few studies, a positive effect of maternal postures was observed[46-48,50], whereas in the majority of them findings were negative[27,28,40,43-45]. However, in most negative trials substantial limitations can be observed, including a clinical instead of a sonographic evaluation of OPP[45], an inadequate sample size[43], and the maintenance of the investigated maternal posture for a too short period of time (<30 minutes)[43-45]. It is likely that fetal head rotation is induced by the combination of maternal posture and uterine contractility, thus making maintenance of a certain posture for a sufficient period of time (at least one hour) pivotal for its effectiveness[18,28]. In addition, in all but one trial, only one specific maternal posture represented the intervention under investigation[28,43-50]. Yet, “active births”, characterized by at least three changes in maternal posture during the active phase of labor, had been shown to associate with shorter labor and lower cesarean delivery rate than those observed in “inactive” laboring women[52].

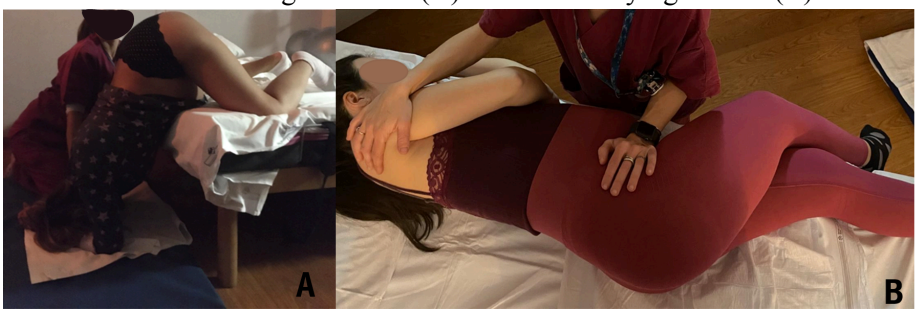
Of note, little or no effect on health outcomes of women and their infants and on maternal satisfaction has been reported by a recent Cochrane review on maternal postures in labor in case of fetal malposition, and the need of further research on the topic has been highlighted[18].

Although no clear benefits of maternal postures have been identified[18,27,28,45], some trials have shown an improvement in maternal back pain and comfort[43,44,49]. Considering this, and the fact that maternal postures are safe, simple to execute, non-invasive, and widely accepted by women, they are a cornerstone of the current obstetric practice employed by midwives in assisting laboring women with an OPP fetus in Italy as well as worldwide.

More recently, additional interventions have been proposed as non-invasive alternatives to classical maternal postures for favoring anterior rotation of an OPP fetus: the forward-leaning inversion (FLI) and the side-lying release (SLR) postures and the Rebozo technique[53,54].

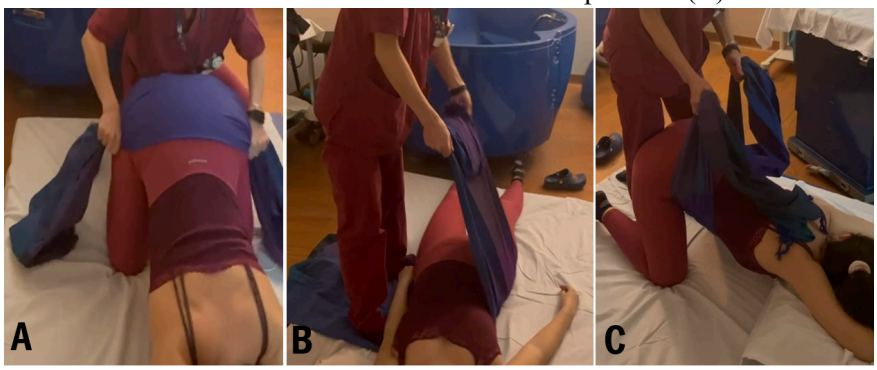
The FLI and SLR postures rely on the importance of the role of soft tissues, including ligaments, muscles, and connective tissue, in fetal head's correct positioning (i.e., anterior) and labor onset and progress (Fig. 1, A and B). Briefly, the FLI uses stretch receptors to untwist and lengthen the utero-sacral ligament for increased maternal comfort, dilation ease, and improved fetal position, whereas the SLR allows uterus repositioning, pelvic floor softening, and buttock and hip muscles release[55].

Figure 1. The forward-leaning inversion and the side-lying release procedures. Pictures show the forward-leaning inversion (A) and the side-lying release (B).



The Rebozo technique is based on the use of the rebozo (a woven shawl) to massage or shift the woman's pelvis or uterus (Fig. 2, A-D), thereby encouraging fetal rotation and optimum positioning, as well as promoting maternal comfort during labor. It is a long-standing traditional practice from central and southern Mexico, which, in recent years, has spread in other countries, including the United States, Denmark, and Italy[53,54].

Figure 2. The Rebozo technique.Pictures display the use of the Rebozo to massage the maternal pelvis in a woman in hands-and-knees position (A), the maternal side ipsilateral to the fetal spine (B), and the uterus in a woman in hands-and-knees position (C).



Neither the FLI and SLR postures nor the Rebozo technique have been shown to harm the mother or the fetus[53,55,56]. Also, they are non-invasive and have been associated with a positive childbirth experience[54,55]. Yet, neither intervention has been rigorously evaluated in a research study as to its effectiveness in determining anterior rotation of an OPP fetus.

Given the complications that are associated with a POPP, it is important to rigorously evaluate interventions that may help fetuses to rotate to the OAP.

There are currently two trials registered on Clinicaltrials.gov and one registered on ANZCTR (Australian New Zealand Clinical Trial Register), not yet recruiting and located in the United States and New Zealand, respectively, aiming to investigate lateral recumbent position to promote anterior rotation of the fetal head in case of OPP (NCT05307393, "Maternal Positioning to Correct Fetal Occiput Posterior", NCT05881629, "Early Diagnosis and Intervention for Fetal Malposition in Active Labor and Its Impact on Mode of Delivery", and "Evaluate the design of a randomised trial of posture for occiput posterior position in labour (POPPIL) to reduce operative births. A feasibility study", ACTRN12624000375550)[57,58].

It is well known in the field of medicine and public health that a bundle of interventions, grouped together in a single protocol, for improving a specific outcome is usually more effective than any single intervention delivered alone[59]. This concept is referred to as the "care bundle"[60,61]. By acting on uterus repositioning and pelvic floor muscle and uterine ligaments and fascia release, the supposed mechanisms of action underlying both the FLI and SLR postures and the Rebozo technique are similar. Such mechanisms can possibly create the most favorable conditions to allow anterior rotation of the OPP fetal head[53,55,56]. Also, it is known that frequent changes in maternal postures during active labor can reduce labor length and operative delivery rate[52].

Therefore, we hypothesized that a combination of FLI and SLR postures and Rebozo technique in a pre-specified sequence during the first stage of labor of women with an OPP fetus would favor fetal head

flexion and, in turn, its anterior rotation, thus leading to lower probabilities of OPP at three hours and thirty minutes from randomization compared to women using only maternal postures[28,40].

Objectives

The aim of our study is to assess whether a sequenced approach including the FLI and SLR postures and the Rebozo technique favors fetal head rotation from OPP to OAP in a randomized controlled trial.

METHODS AND ANALYSIS

Patient and public involvement

Patients were not directly involved in the design or conduct of the study; however, a thorough literature review, which included evaluation of both quantitative and qualitative studies on patients' experience and preferences in case of labor with a posterior fetal head position, has been conducted before defining the research question and outcome measures[18]. The burden of the intervention will be assessed by evaluating maternal satisfaction of the childbirth event.

Study's results will be disseminated to participants by means of dedicated in-person events and online communications (Fondazione IRCCS San Gerardo dei Tintori Facebook page).

Trial design and setting

Study Design

The ReMaP-POPP (Rebozo and maternal postures to prevent persistent occiput posterior position of the fetal head) trial will be a pragmatic, open-label, single-center randomized controlled trial with two parallel groups.

Maternal postures, including upright, lateral recumbent, and hands-and-knees, have been commonly used at our Institution by midwives assisting laboring women with OPP fetuses. Also, in the last couple of years, training sessions for both the FLI and SLR postures and the Rebozo technique have been undertaken by midwives at our Institution to improve their knowledge and operational confidence in both interventions, which, in turn, have been progressively introduced in clinical practice. However, their use in laboring women with an OPP fetus is still inconsistent.

Study Setting

The study will be conducted at the Labor & Birth (L&B) unit, Fondazione IRCCS San Gerardo dei Tintori, Monza, Italy. The research center is an academic maternity unit with approximately 2,600 births per year.

Eligibility criteria

Eligible women will be ≥18 years old, in labor, and with a singleton term fetus (>$37^{0/7}$ weeks of gestation) in an OPP clinically diagnosed between 3 and 8 cm of cervical dilation and confirmed by transabdominal sonography[62]. Ultrasound will be performed by a trained clinician, who will be different

from the clinician performing Leopold's maneuvers and vaginal examination and blinded to the findings of these two steps and to the randomization group.

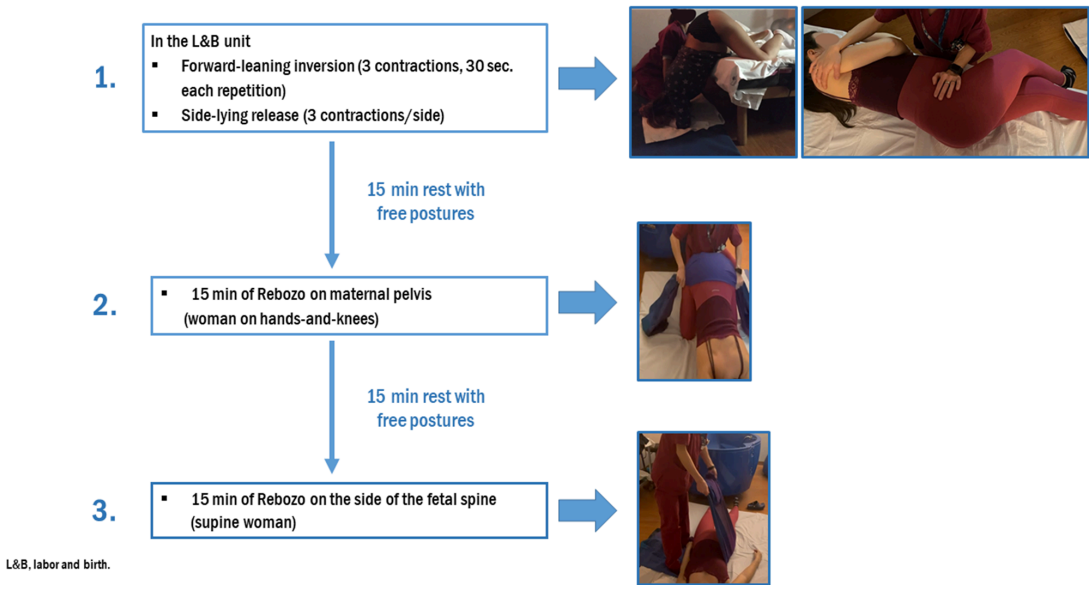
Women with growth restricted fetuses (FGR), defined according to the Delphi consensus, will be excluded[63]. Early FGR (<32 weeks) includes estimated fetal weight (EFW) or fetal abdominal circumference (AC) <3rd centile, absent end-diastolic flow in the umbilical artery (UA), or EFW/AC <10th centile with UA or uterine artery pulsatility index (PI) >95th centile. Late FGR (≥32 weeks) includes EFW or AC <3rd centile, or EFW/AC <10th centile, crossing >2 growth chart quartiles, cerebroplacental ratio (CPR) <5th centile, or UA-PI >95th centile[63]. Fetuses with congenital anomalies or infections or with chromosomal abnormalities, intrauterine fetal demise, and fetal and/or maternal conditions requiring urgent or emergent delivery or impeding the use of the Spinning Babies® procedures and/or the Rebozo technique (e.g., non-reassuring fetal heart rate - FHR, abnormal vaginal bleeding, epidural analgesia resulting in reduced mobility, polyhydramnios, extrapelvic fetal head, body mass index ≥35 Kg/m^2, hypertensive disorders of pregnancy - HDP with inadequate control of blood pressure[64], maternal heart disease in class III to V according to the modified World Health Organization – mWHO classification,[65] glaucoma or ocular surgery, esophageal reflux disease, hypermobile sacro-ileac joint, and severe symphysis dysfunction) will be excluded. Also, we will not enroll women in whom cesarean section is planned as mode of birth, who did not receive information regarding the study before onset of active labor, who do not understand Italian if an official translator cannot be present at the time of explaining the study and proposing enrollment, as well as women in whom the Spinning Babies® procedures and/or the Rebozo technique have been already used before enrollment.

Intervention and comparator

Specifically, the intervention group will receive a combination of FLI and SLR postures and Rebozo technique in a pre-specified order, as displayed in Figure 3. Duration of the intervention ranges between 90 and 105 minutes, according to the rhythm of uterine contractions. The range of cervical dilation for eligibility (3-8 cm) has been chosen to allow completion of the intervention. The intervention will start within thirty minutes after randomization; it can be interrupted at any time if needed, e.g., placement of epidural analgesia or FHR abnormalities; it will be considered performed when the proposed sequence is completed (Fig. 3).

Women allocated to the control group will receive the standard of care, including upright, semi-recumbent, lateral recumbent, and hands-and-knees position. They will not receive FLI or SLR or Rebozo.

Figure 3. Sequence of forward-leaning inversion and side-lying release procedures and the Rebozo technique in the intervention group. L&B, labor and birth.



After completion of the sequence, women in the intervention group will be allowed to move freely and adopt different postures, such as upright, semi-recumbent, lateral recumbent, or hands-and-knees. No FLI or SLR postures or the Rebozo technique will be further allowed until sonographic assessment of the fetal head is performed, at three hours and thirty minutes after randomization. Similarly, the control group will continue to adopt maternal postures with no possibility of performing the FLI or SLR postures or the Rebozo technique until sonographic assessment of fetal head position is completed.

The choice of this time frame for primary outcome assessment was based on several considerations. First, it ensures adequate time for transferring the woman to the labor room and for the midwife's initial assessment, including evaluation of pain intensity and coping ability. The subsequent three hours are expected to allow for full implementation of the intervention sequence in most women in the intervention group—even in the presence of low-frequency uterine contractions or temporary interruptions due to epidural placement or FHR abnormalities. Moreover, this time frame provides a reasonable balance between protocol adherence and clinical flexibility, while also minimizing the risk of contamination in the control group: after this period, FLI, SLR, and Rebozo may be used in the control group as well, at the midwife's discretion.

Once sonography is performed at three hours and thirty minutes after randomization, findings will be communicated to the midwife and obstetrician in charge to allow proper labor management among women in both groups, with the possibility of using maternal postures, FLI and SLR postures , and Rebozo technique, according to the L&B unit's practices.

An additional ultrasound to assess fetal head and spine position will be performed once full cervical dilation is reached. Also, sonography will be performed to confirm fetal head and spine position before performing operative vaginal or abdominal delivery.

Before the beginning of the trial, all participating midwives and obstetricians will receive didactic instructions on the protocol procedures and review the proposed sequenced approach to ensure consistency. A graphical abstract of the study protocol will also be available as a poster in the L&B unit. One senior midwife and one obstetrician per shift will be identified as the clinicians in charge of coordinating and overseeing the trial procedures during the shift. In addition, a PhD candidate, who has a strong background in midwifery (M.P.), will be dedicated to the study and, thus, present daily in the research setting to monitor implementation, support the clinical team, and ensure adherence to the protocol throughout the study period.

Outcomes

The primary outcome is the probability of OPP of the fetal head three hours and thirty minutes after randomization, diagnosed by suprapubic transabdominal sonography by a trained clinician blinded to the randomization group. The fetal head position will be classified into one of three categories (Fig. 4 and 5): occiput anterior (OA) (right and left), occiput transverse (OT) (right and left), and OP (right and left).[62] If the fetal head position cannot be clearly defined, sonography will be repeated by another trained, senior clinician. For women who will give birth before the end of the three hours and a half after randomization, fetal head position at birth will be considered as the fetal head position for this outcome.

The main secondary outcomes are: 1) probability of OPP confirmed by ultrasound examination at full cervical dilation (before the beginning of active maternal pushing efforts); for women who will give birth before sonography at full dilation can be performed, information regarding fetal head position at this timing will be considered as "missing"; and 2) probability of OPP at delivery (in women undergoing operative delivery, the fetal head position diagnosed by sonography right before delivery will be considered as the position at birth).

Additional secondary outcomes are: 3) duration of of labor; 4) mode of delivery; 5) perineal lacerations (episiotomy or third- and fourth-degree perineal tears); 6) primary postpartum hemorrhage (quantitative blood loss >1000 mL[66,67]) 7)woman's pain intensity and ability to cope with pain, assessed using the Pain Intensity Scale and the Pain Coping Scale at two time points (within the first 30 minutes and at 3 hours and 30 minutes after randomization) (Fig. 6, A and B); presence of epidural analgesia will also be

recorded;[68] 8) maternal birth satisfaction level, assessed using the 10-item Birth Satisfaction Scale-Revised (BSS-R) at least 24 hours after birth and before hospital discharge;[69] 9) admission to neonatal intensive care unit; 10) signs and symptoms of pelvic floor and sexual dysfunction at 6–9 months postpartum, including presence of urinary/fecal incontinence, pelvic pressure, and dyspareunia, assessed through clinical interview, validated questionnaires (Pelvic Floor Distress Inventory -PFDI-20- and Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire -PISQ-12), vaginal manometry, electromyography, the Modified Oxford Grading System, Pelvic Organ Prolapse Quantification (POP-Q) score, stress test for incontinence, and 2D/3D ultrasound assessment of bladder neck descent and/or genital ballooning.[70-76]

Figure 4. Fetal head fontanels and the potential location of fontanels and spine in the maternal pelvis.
Figure displays the occipital and the frontal fontanel of the fetal head (A), as well as their location inside the maternal pelvis (B), alongside the location of the fetal spine (C).
OP, occiput posterior; LOT, left occiput transverse. OA, occiput anterior; ROT, right occiput transverse.

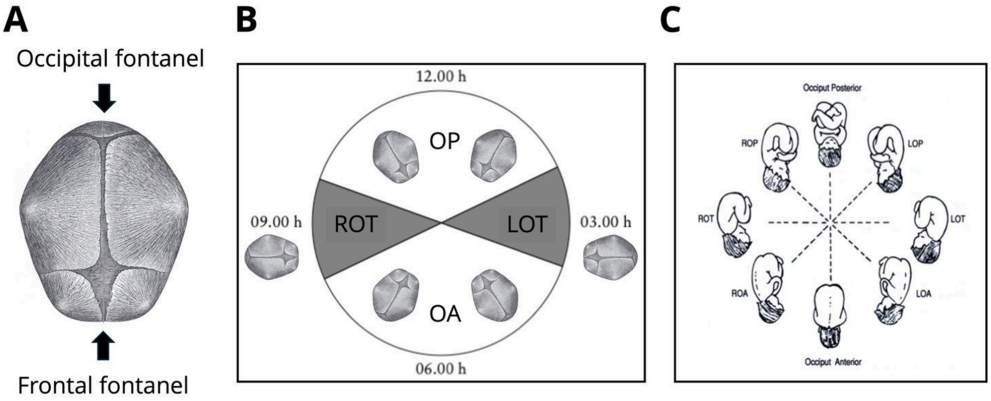
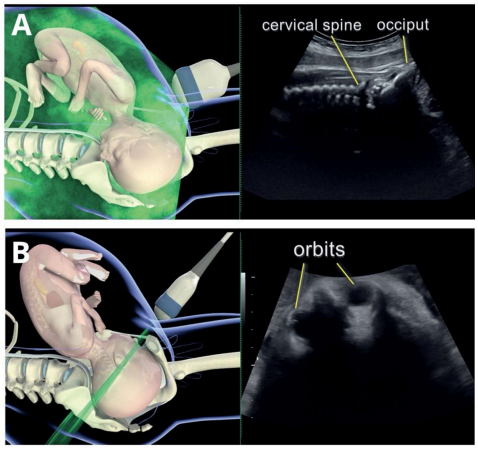


Figure 5. Sonography for diagnosing fetal spine and head position during labor.
A suprapubic transabdominal sonography identifies an anterior fetal spine and occiput in an occiput anterior fetus (A) and anterior fetal orbits in an occiput posterior fetus (B).
Adapted from Ghi et al. [61]
Reproduced with permission from: Ghi T, Eggebø T, Lees C, et al. ISUOG Practice Guidelines: intrapartum ultrasound. Ultrasound in Obstetrics & Gynecology. 2018;52(1):128-139.



Data collection methods

Data collection process

Midwives and obstetricians will perform Leopold's maneuvers, vaginal examination, and sonography at the time of woman's assessment for eligibility. Sonography will be performed by a physician blinded to the randomization group. Midwives and obstetricians involved in eligibility assessment will enroll eligible women and obtain the signed informed consent. The senior midwife or obstetrician in charge of coordinating and overseeing the trial procedures during the shift will randomize the eligible women.

Women will be assessed sonographically at enrollment, at the end of the three and a half hours after randomization, at full cervical dilation, and before any operative delivery to define fetal head and spine position. In women undergoing operative delivery, the fetal head position diagnosed by sonography right before delivery will be considered as the position at birth. Position of the fetal head will be also prospectively collected at each vaginal examination during labor, alongside with the fetal head level, and at birth by the physicians managing the laboring woman, and recorded in a dedicated study's data collection chart. Also, details of the postures employed by the women and their duration will be recorded, as well as the number and duration of interruptions in the intervention sequence, if any, and its completion. Further, deviations to the protocol (e.g., use of Rebozo in a woman in the control group) will be recorded.

Two research assistants independent of the local medical team will assess the collected study's data collection chart and record the included data in the Research Electronic Data Capture (REDCap)

database; all paper printouts of the ultrasound assessment will be collected, scanned, and safely stored in the system. Another independent research assistant will monitor and assess the quality of data for all participants.

Race and ethnicity will be classified according to the method recommended by the National Institute of Health, notice number NOT-OD-15-089. The following categories will be considered: American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, and White.[82] Immigrant status will be defined as any person who is not a citizen or national of Italy but who is living permanently in Italy. This group will also include refugee claimants.

The data will be treated confidentially, stored securely on laptops that will be password protected and disposed of properly at the completion of the study (Data Protection Code, 2003; GDPR, 2018).

Data collection tool

Data will be collected online using the REDCap system provided by the University of Milan-Bicocca Clinical Research Office (BiCRO).

REDCap assures maintenance of data confidentiality by the automatic generation of a research code for each enrolled woman by means of an electronic document.

A paper logbook containing information on all women screened for eligibility and the reasons for ineligibility or lack of inclusion if eligibility is met will be kept separately, in a locked cabinet accessible only by the principal investigator and the research assistants. This logbook will also contain research codes and personal data of enrolled women.

The REDCap database is constituted by different modules to collect data regarding maternal general characteristics, medical and obstetric history, ultrasound scans and clinical course of the pregnancy, labor and birth outcomes, and neonatal outcomes (Supplementary Material 1).

The research assistants involved in data collection and recording in the REDCap database will be trained to use the system before the commencement of the study; only these research assistants and the principal investigator will be granted access to the system.

Before the beginning of the trial, all participating midwives and obstetricians will be specifically trained on the execution of transabdominal ultrasound to assess fetal head and spine position to ensure adequacy and consistency in outcome assessment.

For participants who discontinue or deviate from the intervention protocol, all outcome data specified above will be collected. Reasons for non-adherence (e.g., discomfort, perceived inefficacy, adverse effects, or other personal or clinical factors) will be systematically recorded in the case report forms. Similarly, reasons for non-retention (e.g., withdrawal of consent, loss to follow-up) will be documented when available. To support retention at the 6-month follow-up, participants will be offered a

free-of-charge gynecological assessment, including clinical evaluation and counselling, as an incentive and opportunity for continuity of care. Additional periodic reminders (via phone or email) will be sent to participants to minimize attrition.

Harms

Safety outcomes will include 1) rate of FHR assessments that are considered to be of poor quality (satisfactory/acceptable/unsatisfactory) in women with both intermittent and external continuous FHR monitoring; 2) frequency of FHR signal loss (<10%, 10-25%, 25-50%, >50%) in women with external continuous FHR monitoring; and 3) need of an internal electrode to allow proper continuous FHR monitoring because of maternal positions during the first three hours and a half after randomization. These outcomes will be assessed systematically by the attending midwives or obstetricians managing the randomized women's labor; these professionals will not be blinded to group allocation. Data will be recorded in real time using predefined criteria and structured forms integrated in the Case Report Form (CRF). Harms will be monitored continuously from randomization until birth, and any unanticipated adverse event will be documented according to clinical records and reported to the Data and Safety Monitoring Board (DSMB) as per trial procedures. All adverse events will be coded by the clinical research team using standard obstetric definitions and graded by severity (mild, moderate, severe) and causality (related or unrelated to the intervention). Coding and grading will be reviewed by an independent member of the research team not involved in participant care and blinded to allocation. Serious adverse events or events leading to early discontinuation of the intervention will be promptly reported to the DSMB.

Participant timeline

A detailed flowchart outlining the trial timeline has been developed to support clinical staff in the consistent implementation of the study procedures. This visual tool is available as the Supplementary Material 2.

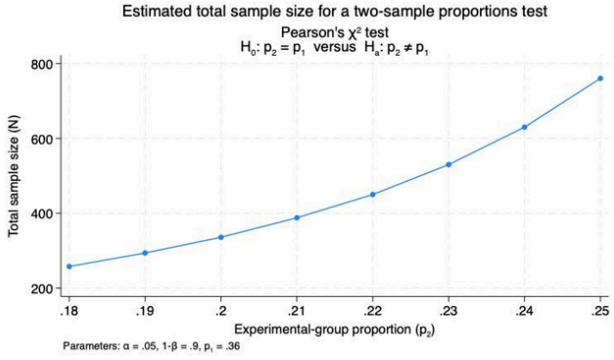
Sample size

Power analysis was conducted by means of Pearson's Chi Squared test comparing two independent proportions.

According to a prior retrospective analysis performed on laboring women managed at our Institution in 2017 (i.e., when only maternal postures, including upright, semi-recumbent, lateral recumbent, and hands-and-knees, were used) and in 2023 (i.e., after the introduction of the FLI and SLR procedures and the Rebozo technique into clinical practice), we observed a probability of POPP at complete dilation of 35.8% (95% confidence interval - CI: 0.31;0.40) and 27.7% (95% CI: 0.22;0.34), respectively (Risk Difference=-0.081, 95% CI: -0.15;-0.008; p=0.031)[77]. The 2017 value is in line with findings from prior trials on maternal postures in fetal malposition.[28,40]

Considering the potential benefits of an approach based on a pre-specified, >1 hour-long sequence of FLI, SLR, and Rebozo procedures, we assume that women in the intervention group can reach a probability of OPP of the fetal head at sonography assessment three hours and a half after randomization of 22% (lower limit of the 95% CI of the probability observed in our 2023 retrospective analysis, 27.7%, when FLI, SLR, and Rebozo procedures had been introduced into clinical practice, although inconsistently used). This leads to a clinically meaningful Δ between the intervention (22%) and the control group (35.8%) of 13.8%. To show this difference with a power of 0.90 (1-β) and a two-sided α value of 0.05, the study will require the inclusion of 231 women in each group, for a total of 462 participants with a 1:1 ratio (Fig. 7). Considering a potential attrition rate of 20%, the final estimated sample size will be n=462/(1-0.20)=578 participants.

Figure 7. Pearson's Chi-Squared test for two independent proportions for sample size estimation. Graph displays the different total sample sizes needed (y axis) according to 1% increment (from 18% to 25%, x axis) of the rate of POPP in the intervention group, as assessed at three hours and thirty minutes after randomization, considering a baseline POPP rate (control group) of 36%.

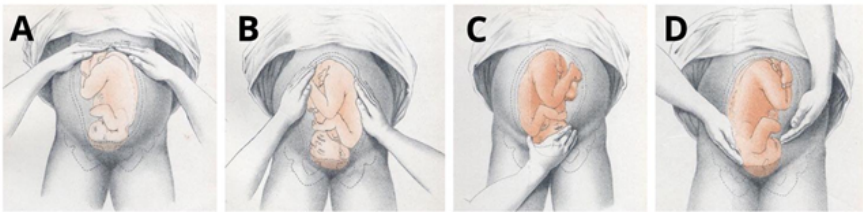


According to our prior retrospective analysis, approximately n=450 women per year will be eligible in the study. Thus, we anticipate a total duration of the enrollment period of ~16 months.

Recruitment

Participants will be recruited at a single academic maternity center located in Northern Italy. The enrollment phase is expected to begin in October 2025, with study completion anticipated by November 2027. All women accessing our L&B unit will be assessed to 1) define fetal spine position by the second Leopold's maneuver (Fig. 8, A-D), 2) diagnose cervical dilation and fetal head position and level by vaginal examination, and 3) confirm fetal head and spine position by transabdominal sonography.

Figure 8. The Leopold's maneuvers.The four Leopold's maneuvers are shown: the first maneuver or fundal grip, assessing the uterine fundus to determine its height and which fetal pole occupies the fundus (A), the second maneuver or umbilical grip, for identifying the location of the fetal spine (B), the third maneuver or the Pawlik's grip, to aid in confirmation of fetal presentation (C), and the fourth maneuver, to define confrontation of the cephalic extremity with the birth canal (D).



Vaginal examination can diagnose fetal OPP by finding the junction of the sagittal and lambdoidal sutures of the fetal head (occipital fontanel) located at least 45 degrees posterior to the transverse diameter of the maternal pelvis (Fig. 4, A-C), whereas sonography can locate the position of the fetal orbits anteriorly or the cervical spine posteriorly to the maternal pelvis (Fig. 5, A and B).[62] Training for performing transabdominal sonography will be completed by all the participating midwives and obstetricians before starting the trial. The training includes both a theoretical component and a practical hands-on session. During the practical part, participants (midwives and obstetricians) perform supervised ultrasound assessments on pregnant women admitted to the maternity unit where the trial will take place. This training model will ensure an adequate level of theoretical and practical expertise among clinicians, supporting the reliability and consistency of the ultrasound-based assessment of the primary outcome.[78]

To increase acceptability and participation, information regarding the study will be provided to all eligible women and their partners during an antenatal visit in the third trimester of pregnancy by clinicians.[79] Fliers about the study will be also posted in the outpatient clinic, the antenatal ward, and the L&B unit. Importantly, all women who plan to give birth at the study centre—regardless of where they receive their antenatal care—are routinely offered a dedicated pre-labour consultation around 36 weeks' gestation. This appointment provides an opportunity for early engagement and personalized explanation of the study. Such information will be repeated in the L&B unit to women in labor and their partners after sonographic confirmation of fetal OPP.[79] Women will be asked to confirm their participation and provide written consent. After inclusion, they will be randomly assigned to the intervention or the control group.

Active engagement of women will be pursued through public dedicated website communications (Fondazione IRCCS San Gerardo dei Tintori Facebook page), digital content, and leaflets reporting on the implications of the study and webinars. The involvement of relevant professional categories (e.g., midwives and obstetricians) will be achieved by means of online and in-person meetings.

Randomisation: sequence generation, allocation concealment mechanism, and implementation

As soon as the written consent is signed, the randomization will be performed by an automated web-based system to ensure allocation concealment (24-hour accessibility with personal login and password), provided by BiCRO. The senior midwife or obstetrician in charge for coordinating and overseeing the trial procedures during the shift will be responsible for randomization.

Considering the unblinded nature of our study and that parity is an important predictor of our primary outcome, randomization will be performed using randomly permuted blocks of varying sizes (4, 6, and 8), stratified by parity (nulliparous/multiparous). This strategy will not be known by local investigators. A blocked randomization with fixed block size will not be employed since the reduced suitability of this type of randomization in an unblinded trial due to potential anticipation and manipulation of treatment assignment of the participant at the end of each block.

The ratio for intervention versus control will be 1:1.

Blinding

Both the enrolled women and their managing midwives and physicians will be unblinded to the randomization group, due to the nature of the intervention. In turn, blinding will be guaranteed for the physician assessing the fetal head position by suprapubic transabdominal sonography, the gold standard technique to objectively define fetal head and spine position.[3-5]

Data management

Trial data will be initially recorded on dedicated paper charts during clinical care and later entered into a secure REDCap database by trained research assistants. REDCap includes built-in range and logic checks, as well as standardized coding, to ensure data accuracy and quality. Paper source documents (e.g. monitoring charts, ultrasound reports) will be securely stored in locked cabinets, and scanned copies will be uploaded to REDCap when needed. Access to both paper and digital data will be restricted to authorized personnel only. Data will be stored securely for a minimum of 10 years, in compliance with institutional and regulatory requirements. A full data management plan is available upon request.

Statistical methods

Data reporting will be performed according to the Consolidated Standards of Reporting Trials (CONSORT) guidelines.[80,81]

A descriptive table of the baseline characteristics will be reported for the participants for both groups. Primary and secondary outcomes will be analyzed on an intention-to-treat basis. Sub-group analyses for parity and cervical dilation at inclusion (3-4 cm or 5 to 8 cm) will be conducted.

For continuous variables, distribution will be visually assessed and means and standard deviations or medians and interquartile ranges will be calculated for normally and not normally distributed variables, respectively; Student's t-test and Mann Whitney U-test will be used to compare the outcomes between groups. For dichotomous variables, proportions will be calculated, and Chi-Square and Fisher exact tests

will be used as appropriate to assess differences in outcomes between groups. The effects of the intervention will be estimated by relative risks and their 95% CI.

Probabilities of episiotomy and perineal lacerations and the total duration of the second stage of labor will be calculated only among women with vaginal deliveries. The speed of cervical dilation will be calculated only on women who reach complete dilation; it will be defined as mean cervical dilation measured in centimeters per hour of labor and calculated with the following formula: (10 – cervical dilation at randomization)/(time at complete dilation – time at randomization).

A Cochran-Mantel-Haenszel homogeneity test will be performed to assess the consistency of the primary outcome according to parity.

Analyses will be performed with the open-source R software v.4.4.2 (R Foundation for Statistical Computing) and STATA software v. 16 (College Station, TX: StataCorp LLC). Significance will be set at <.05.

Data monitoring committee

An independent DSMB will be established prior to trial initiation. The board will consist of three members with expertise in obstetrics, biostatistics, and clinical epidemiology. The DSMB will operate independently from the study investigators, and funders, and all members will be required to declare any conflict of interest before appointment and throughout the study. The DSMB will be responsible for overseeing participant safety, assessing protocol adherence and data quality, and reviewing unblinded interim data. It will meet after the enrolment of the first participant and subsequently when approximately 25%, 50%, and 75% of the target sample has been recruited. After each review, the DSMB will provide written recommendations to the trial Steering Committee. A detailed DSMB charter outlining roles, procedures, and governance is available upon request.

Interim analyses will be conducted by an independent statistician and will be accessible only to the DSMB. Trial investigators and staff will remain blinded to group allocation. Interim reviews will be descriptive in nature and focused on safety, recruitment progress, and data quality. Formal statistical testing will be performed only if serious concerns arise. The DSMB may recommend early termination or protocol modifications in the event of significant safety issues, clear evidence of benefit or futility, or other ethical or feasibility concerns. Final decisions regarding trial continuation will be made by the Steering Committee, based on DSMB recommendations.

Trial monitoring

A risk-based monitoring strategy will be implemented to ensure trial integrity and adherence to the study protocol. Monitoring activities will include regular on-site visits and centralized data checks. Throughout the study period, biweekly audit meetings will be conducted to review recruitment progress, assess protocol compliance, and promptly address any operational issues, including deviations or potential

contamination. Monitoring will be coordinated by the study's core research team, independently from clinical staff involved in day-to-day patient care.

ETHICS AND DISSEMINATION:

Research ethics approval

The study has been approved by the Lombardy Ethic Committee n.3 (n. 5499, 20 December 2024) before starting enrollment.

Dissemination policy

The dissemination plan includes the presentation of the findings at national and international scientific meetings, and the publication in peer-reviewed scientific journals in the field of midwifery and obstetrics.

The findings will also be disseminated to the public through reach-out activities involving families and healthcare specialists in order to promote a culture of a positive experience of childbirth that benefits from evidence-based research and further contribute to the long-term promotion of a woman-centered approach to childbirth.

Protocol amendments

Any important protocol amendments will be decided by the principal investigator in consultation with the Steering Committee. Substantive changes will be promptly communicated to the relevant ethics committee, regulatory authorities, and trial registries. Investigators and study staff will be informed through formal communications, and updated documents will be distributed accordingly.

Consent or assent

Written informed consent will be obtained from all participants (Supplementary Material 3). Women will be free to decline participation or to withdraw at any time.

All participants will be provided the name, telephone number, and email address of the principal investigator, in case any question about the study should arise.

Confidentiality

All the procedures will be consistent with the Declaration of Helsinki ethical principles for research involving human subjects(World Medical Association, 2013). The procedures do not imply any change to mother-infant care programmes in place at our Institution.

TRIAL REGISTRATION

The study was registered in the Clinical Trials database (ClinicalTrials.gov, NCT06887634) on March 20, 2025 (https://clinicaltrials.gov/study/NCT06887634?term=id%205499_18.12.2024_m%20bis&rank=1)

Protocol and statistical analysis plan {5}

The full trial protocol was registered in the Clinical Trials database (ClinicalTrials.gov, NCT06887634), where a summary will be accessible.The detailed statistical analysis plan (SAP) will be finalized prior to database lock and uploaded as a supplementary file in the trial registry and journal repository. Both documents will also be available upon reasonable request from the corresponding author.

Data sharing

De-identified individual participant data, the data dictionary, and statistical analysis code will be made available upon reasonable request after publication of the primary results. Materials related to the intervention (including any instructional tools used in the trial) will also be shared if applicable. Data access will require submission of a brief proposal outlining the intended use, which will be reviewed by the study team. Approved data will be shared via secure data transfer platforms. Consent for future data sharing will be obtained from all participants at the time of enrollment. No commercial use of the data will be permitted.

Funding and conflicts of interest

This study is supported by external funding from the University of Milano-Bicocca (grant codes: 2025-ATE-0205 to S.O.; 2025-ATEQC-0028 to S.F.). The funder didn't influence the results/outcomes of the study despite author affiliations with the funder.

REFERENCES

1. Gardberg M, Laakkonen E, Sälevaara M. Intrapartum sonography and persistent occiput posterior position: a study of 408 deliveries. Obstet Gynecol [Internet]. 1998 May;91(5 Pt 1):746–9. Available from: http://dx.doi.org/10.1016/s0029-7844(98)00074-x

2. Gardberg M, Tuppurainen M. Persistent occiput posterior presentation--a clinical problem. Acta Obstet Gynecol Scand [Internet]. 1994 Jan;73(1):45–7. Available from: http://dx.doi.org/10.3109/00016349409013392

3. Sherer DM, Miodovnik M, Bradley KS, Langer O. Intrapartum fetal head position II: comparison between transvaginal digital examination and transabdominal ultrasound assessment during the second stage of labor. Ultrasound Obstet Gynecol [Internet]. 2002 Mar;19(3):264–8. Available from: http://dx.doi.org/10.1046/j.1469-0705.2002.00656.x

4. Sherer DM, Miodovnik M, Bradley KS, Langer O. Intrapartum fetal head position I: comparison between transvaginal digital examination and transabdominal ultrasound assessment during the active stage of labor. Ultrasound Obstet Gynecol [Internet]. 2002 Mar;19(3):258–63. Available from: http://dx.doi.org/10.1046/j.1469-0705.2002.00641.x

5. Akmal S, Tsoi E, Howard R, Osei E, Nicolaides KH. Investigation of occiput posterior delivery by intrapartum sonography. Ultrasound Obstet Gynecol [Internet]. 2004 Sep;24(4):425–8. Available from: http://dx.doi.org/10.1002/uog.1064

6. Hjartardóttir H, Lund SH, Benediktsdóttir S, Geirsson RT, Eggebø TM. When does fetal head rotation occur in spontaneous labor at term: results of an ultrasound-based longitudinal study in nulliparous women. Am J Obstet Gynecol [Internet]. 2021 May;224(5):514.e1–514.e9. Available from: http://dx.doi.org/10.1016/j.ajog.2020.10.054

7. Vitner D, Paltieli Y, Haberman S, Gonen R, Ville Y, Nizard J. Prospective multicenter study of ultrasound-based measurements of fetal head station and position throughout labor. Ultrasound Obstet Gynecol [Internet]. 2015 Nov;46(5):611–5. Available from: http://dx.doi.org/10.1002/uog.14821

8. Parsy T, Bettiol C, Vidal F, Allouche M, Loussert-Chambre L, Guerby P. Persistent occiput posterior position: predictive factors of spontaneous rotation of the fetal head. J Matern Fetal Neonatal Med [Internet]. 2023 Dec;36(1):2192854. Available from: http://dx.doi.org/10.1080/14767058.2023.2192854

9. Fitzpatrick M, McQuillan K, O'Herlihy C. Influence of persistent occiput posterior position on delivery outcome. Obstet Gynecol [Internet]. 2001 Dec;98(6):1027–31. Available from: http://dx.doi.org/10.1016/s0029-7844(01)01600-3

10. Ponkey SE, Cohen AP, Heffner LJ, Lieberman E. Persistent fetal occiput posterior position: obstetric outcomes. Obstet Gynecol [Internet]. 2003 May;101(5 Pt 1):915–20. Available from: http://dx.doi.org/10.1016/s0029-7844(03)00068-1

11. Senécal J, Xiong X, Fraser WD, Pushing Early Or Pushing Late with Epidural study group. Effect of fetal position on second-stage duration and labor outcome. Obstet Gynecol [Internet]. 2005 Apr;105(4):763–72. Available from: http://dx.doi.org/10.1097/01.AOG.0000154889.47063.84

12. Cheng YW, Shaffer BL, Caughey AB. The association between persistent occiput posterior position and neonatal outcomes. Obstet Gynecol [Internet]. 2006 Apr;107(4):837–44. Available from: http://dx.doi.org/10.1097/01.AOG.0000206217.07883.a2

13. Cheng YW, Hubbard A, Caughey AB, Tager IB. The association between persistent fetal occiput posterior position and perinatal outcomes: an example of propensity score and covariate distance matching. Am J Epidemiol [Internet]. 2010 Mar 15;171(6):656–63. Available from: http://dx.doi.org/10.1093/aje/kwp437

14. Foggin HH, Albert AY, Minielly NC, Lisonkova S, Koenig NA, Jacobs EN, et al. Labor and delivery outcomes by delivery method in term deliveries in occiput posterior position: a population-based retrospective cohort study. AJOG Glob Rep [Internet]. 2022 Nov;2(4):100080. Available from: http://dx.doi.org/10.1016/j.xagr.2022.100080

15. Pergialiotis V, Bellos I, Fanaki M, Vrachnis N, Doumouchtsis SK. Risk factors for severe perineal trauma during childbirth: An updated meta-analysis. Eur J Obstet Gynecol Reprod Biol [Internet]. 2020 Apr;247:94–100. Available from: http://dx.doi.org/10.1016/j.ejogrb.2020.02.025

16. Cheng YW, Shaffer BL, Caughey AB. Associated factors and outcomes of persistent occiput posterior position: A retrospective cohort study from 1976 to 2001. J Matern Fetal Neonatal Med [Internet]. 2006 Sep;19(9):563–8. Available from: http://dx.doi.org/10.1080/14767050600682487

17. Badawi N, Kurinczuk JJ, Keogh JM, Alessandri LM, O'Sullivan F, Burton PR, et al. Intrapartum risk factors for newborn encephalopathy: the Western Australian case-control study. BMJ [Internet]. 1998 Dec 5;317(7172):1554–8. Available from: http://dx.doi.org/10.1136/bmj.317.7172.1554

18. Barrowclough JA, Lin L, Kool B, Hofmeyr GJ, Crowther CA. Maternal postures for fetal malposition in labour for improving the health of mothers and their infants. Cochrane Database Syst Rev [Internet]. 2022 Aug 31;8(8):CD014615. Available from: http://dx.doi.org/10.1002/14651858.CD014615

19. Lieberman E, Davidson K, Lee-Parritz A, Shearer E. Changes in fetal position during labor and their association with epidural analgesia. Obstet Gynecol [Internet]. 2005 May;105(5 Pt 1):974–82. Available from: http://dx.doi.org/10.1097/01.AOG.0000158861.43593.49

20. Le Ray C, Carayol M, Jaquemin S, Mignon A, Cabrol D, Goffinet F. Is epidural analgesia a risk factor for occiput posterior or transverse positions during labour? Eur J Obstet Gynecol Reprod Biol [Internet]. 2005 Nov 1;123(1):22–6. Available from: http://dx.doi.org/10.1016/j.ejogrb.2005.02.009

21. Caughey AB, Sharshiner R, Cheng YW. Fetal malposition: impact and management. Clin Obstet Gynecol [Internet]. 2015 Jun;58(2):241–5. Available from: http://dx.doi.org/10.1097/GRF.0000000000000106

22. Caldwell WE, Moloy HC. Anatomical variations in the female pelvis and their effect in labor with a suggested classification. Am J Obstet Gynecol [Internet]. 1933 Oct;26(4):479–505. Available from: http://dx.doi.org/10.1016/s0002-9378(33)90194-5

23. Ghi T, Youssef A, Martelli F, Bellussi F, Aiello E, Pilu G, et al. Narrow subpubic arch angle is associated with higher risk of persistent occiput posterior position at delivery. Ultrasound Obstet Gynecol [Internet]. 2016 Oct;48(4):511–5. Available from: http://dx.doi.org/10.1002/uog.15808

24. Gardberg M, Tuppurainen M. Anterior placental location predisposes for occiput posterior presentation near term. Acta Obstet Gynecol Scand [Internet]. 1994 Feb;73(2):151–2. Available from: http://dx.doi.org/10.3109/00016349409013418

25. D'Esopo D. The occipito posterior position. Its mechanism and treatment. Am J Obstet Gynecol. 1941;

26. Hunter S, Hofmeyr GJ, Kulier R. Hands and knees posture in late pregnancy or labour for fetal malposition (lateral or posterior). Cochrane Database Syst Rev [Internet]. 2007 Oct 17;2009(4):CD001063. Available from: http://dx.doi.org/10.1002/14651858.CD001063.pub3

27. Levy AT, Weingarten S, Ali A, Quist-Nelson J, Berghella V. Hands-and-knees posturing and fetal occiput anterior position: a systematic review and meta-analysis. Am J Obstet Gynecol MFM [Internet]. 2021 Jul;3(4):100346. Available from: http://dx.doi.org/10.1016/j.ajogmf.2021.100346

28. Le Ray C, Lepleux F, De La Calle A, Guerin J, Sellam N, Dreyfus M, et al. Lateral asymmetric decubitus position for the rotation of occipito-posterior positions: multicenter randomized controlled trial EVADELA. Am J Obstet Gynecol [Internet]. 2016 Oct;215(4):511.e1–7. Available from: http://dx.doi.org/10.1016/j.ajog.2016.05.033

29. Elmore C, McBroom K, Ellis J. Digital and manual rotation of the persistent occiput posterior fetus. J Midwifery Womens Health [Internet]. 2020 May;65(3):387–94. Available from: http://dx.doi.org/10.1111/jmwh.13118

30. Shaffer BL, Cheng YW, Vargas JE, Caughey AB. Manual rotation to reduce caesarean delivery in persistent occiput posterior or transverse position. J Matern Fetal Neonatal Med [Internet]. 2011 Jan;24(1):65–72. Available from: http://dx.doi.org/10.3109/14767051003710276

31. Phipps H, Hyett JA, Kuah S, Pardey J, Matthews G, Ludlow J, et al. Persistent occiput posterior position outcomes following manual rotation: a randomized controlled trial. Am J Obstet Gynecol MFM [Internet]. 2021 Mar;3(2):100306. Available from: http://dx.doi.org/10.1016/j.ajogmf.2021.100306

32. Masturzo B, Farina A, Attamante L, Piazzese A, Rolfo A, Gaglioti P, et al. Sonographic evaluation of the fetal spine position and success rate of manual rotation of the fetus in occiput posterior position: A randomized controlled trial. J Clin Ultrasound [Internet]. 2017 Oct;45(8):472–6. Available from: http://dx.doi.org/10.1002/jcu.22477

33. Le Ray C, Serres P, Schmitz T, Cabrol D, Goffinet F. Manual rotation in occiput posterior or transverse positions: risk factors and consequences on the cesarean delivery rate. Obstet Gynecol [Internet]. 2007 Oct;110(4):873–9. Available from: http://dx.doi.org/10.1097/01.AOG.0000281666.04924.be

34. Blanc J, Castel P, Mauviel F, Baumstarck K, Bretelle F, D'Ercole C, et al. Prophylactic manual rotation of occiput posterior and transverse positions to decrease operative delivery: the PROPOP randomized clinical trial. Am J Obstet Gynecol [Internet]. 2021 Oct;225(4):444.e1–444.e8. Available from: http://dx.doi.org/10.1016/j.ajog.2021.05.020

35. Bertholdt C, Morel O, Zuily S, Ambroise-Grandjean G. Manual rotation of occiput posterior or transverse positions: a systematic review and meta-analysis of randomized controlled trials. Am J Obstet Gynecol [Internet]. 2022 Jun;226(6):781–93. Available from: http://dx.doi.org/10.1016/j.ajog.2021.11.033

36. Bahl R, Van de Venne M, Macleod M, Strachan B, Murphy DJ. Maternal and neonatal morbidity in relation to the instrument used for mid-cavity rotational operative vaginal delivery: a prospective cohort study. BJOG [Internet]. 2013 Nov;120(12):1526–32. Available from: http://dx.doi.org/10.1111/1471-0528.12398

37. Le Ray C, Deneux-Tharaux C, Khireddine I, Dreyfus M, Vardon D, Goffinet F. Manual rotation to decrease operative delivery in posterior or transverse positions. Obstet Gynecol [Internet]. 2013 Sep;122(3):634–40. Available from: http://dx.doi.org/10.1097/AOG.0b013e3182a10e43

38. Stock SJ, Josephs K, Farquharson S, Love C, Cooper SE, Kissack C, et al. Maternal and neonatal outcomes of successful Kielland's rotational forceps delivery. Obstet Gynecol [Internet]. 2013 May;121(5):1032–9. Available from: http://dx.doi.org/10.1097/AOG.0b013e31828b72cb

39. Tempest N, Hart A, Walkinshaw S, Hapangama DK. A re-evaluation of the role of rotational forceps: retrospective comparison of maternal and perinatal outcomes following different methods of birth for malposition in the second stage of labour. BJOG [Internet]. 2013 Sep;120(10):1277–84. Available from: http://dx.doi.org/10.1111/1471-0528.12199

40. Desbriere R, Blanc J, Le Dû R, Renner JP, Carcopino X, Loundou A, et al. Is maternal posturing during labor efficient in preventing persistent occiput posterior position? A randomized controlled trial. Am J Obstet Gynecol [Internet]. 2013 Jan;208(1):60.e1–8. Available from: http://dx.doi.org/10.1016/j.ajog.2012.10.882

41. Andrews CM, Andrews EC. Physical theory as a basis for successful rotation of fetal malpositions and conversion of fetal malpresentations. Biol Res Nurs [Internet]. 2004 Oct;6(2):126–40. Available from: http://dx.doi.org/10.1177/1099800404268318

42. Barth WH. Persistent occiput posterior. Obstet Gynecol [Internet]. 2015 Mar;125(3):695–709. Available from: http://dx.doi.org/10.1097/aog.0000000000000647

43. Stremler R, Hodnett E, Petryshen P, Stevens B, Weston J, Willan AR. Randomized controlled trial of hands-and-knees positioning for occipitoposterior position in labor. Birth [Internet]. 2005 Dec;32(4):243–51. Available from: http://dx.doi.org/10.1111/j.0730-7659.2005.00382.x

44. Guittier MJ, Othenin-Girard V, de Gasquet B, Irion O, Boulvain M. Maternal positioning to correct occiput posterior fetal position during the first stage of labour: a randomised controlled trial. BJOG [Internet]. 2016 Dec;123(13):2199–207. Available from: http://dx.doi.org/10.1111/1471-0528.13855

45. Molina-Reyes CY, Muñoz-Martínez AL, Martínez-García E, Moore TL, Burgos-Sánchez JA. Efficacy of hands and knees maternal posture to correct the occipitoposterior fetal position during labor. Index Enferm [Internet]. 2014 Jan 1;23(1):15–20. Available from: http://dx.doi.org/

46. Wu X, Fan L, Wang Q. Correction of occipito-posterior by maternal postures during the process of labor. Zhonghua Fu Chan Ke Za Zhi [Internet]. 2001 Aug;36(8):468–9. Available from: https://pubmed.ncbi.nlm.nih.gov/11758180/

47. Ou X, Chen X, Su J. Correction of occipito-posterior position by maternal posture during the process of labor. Zhonghua Fu Chan Ke Za Zhi [Internet]. 1997 Jun;32(6):329–32. Available from: https://pubmed.ncbi.nlm.nih.gov/9596909/

48. Bueno-Lopez V, Fuentelsaz-Gallego C, Casellas-Caro M, Falgueras-Serrano AM, Crespo-Berros S, Silvano-Cocinero AM, et al. Efficiency of the modified Sims maternal position in the rotation of persistent occiput posterior position during labor: A randomized clinical trial. Birth [Internet]. 2018 Dec;45(4):385–92. Available from: http://dx.doi.org/10.1111/birt.12347

49. Bahmaei K, Iravani M, Moosavi P, Haghighizadeh M. Effect of maternal positioning with occipito-posterior fetal position during labor on pain intensity and satisfaction of mothers. Iranian Journal of Obstetrics, Gynecology and Infertility [Internet]. 2018;21(5):66–73. Available from: http://dx.doi.org/10.22038/ijogi.2018.11386

50. Liu LP, Chen JH, Yang ZJ, Zhu J. Corrective effects of maternal extreme flexure and hip abduction combined with contralateral side-lying on persistent foetal occipito-posterior position. Int J Nurs Pract [Internet]. 2018 Oct;24(5):e12663. Available from: http://dx.doi.org/10.1111/ijn.12663

51. Lee N, Munro V, Oliver K, Flynn J. Maternal positioning with flexed thighs to correct foetal occipito-posterior position in labour: A systematic review and meta-analysis. Midwifery [Internet]. 2021 Aug;99(103008):103008. Available from: http://dx.doi.org/10.1016/j.midw.2021.103008

52. Lepleux F, Hue B, Dugué AE, Six T, Riou C, Dreyfus M. Données obstétricales dans une population bénéficiant de variations posturales en cours de travail et d'accouchement. J Gynecol Obstet Biol Reprod (Paris) [Internet]. 2014 Sep;43(7):504–13. Available from: http://dx.doi.org/10.1016/j.jgyn.2013.06.013

53. Cohen SR, Thomas CR. Rebozo technique for fetal malposition in labor. J Midwifery Womens Health [Internet]. 2015 Jul;60(4):445–51. Available from: http://dx.doi.org/10.1111/jmwh.12352

54. Iversen ML, Midtgaard J, Ekelin M, Hegaard HK. Danish women's experiences of the rebozo technique during labour: A qualitative explorative study. Sex Reprod Healthc [Internet]. 2017 Mar;11:79–85. Available from: http://dx.doi.org/10.1016/j.srhc.2016.10.005

55. Tully G. Posterior perspective. Midwifery Today Int Midwife [Internet]. 2015 Summer;(114):8–11. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26281517

56. de Wolff MG, Ladekarl M, Sparholt L, Lykke JA. Rebozo and External Cephalic Version in breech presentation (RECEIVE): A randomised controlled study. BJOG [Internet]. 2022 Sep;129(10):1666–75. Available from: http://dx.doi.org/10.1111/1471-0528.17111

57. https://clinicaltrials.gov/search?cond=fetal%20occiput%20posterior.

58. https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=387386&showOriginal=true&isReview=true.

59. Donati S, Buoncristiano M, Lega I, D'Aloja P, Maraschini A, ItOSS working group. The Italian Obstetric Surveillance System: Implementation of a bundle of population-based initiatives to reduce haemorrhagic maternal deaths. PLoS One [Internet]. 2021 Apr 23;16(4):e0250373. Available from: http://dx.doi.org/10.1371/journal.pone.0250373

60. Crunden E, Boyce C, Woodman H, Bray B. An evaluation of the impact of the ventilator care bundle. Nurs Crit Care [Internet]. 2005 Sep;10(5):242–6. Available from: http://dx.doi.org/10.1111/j.1362-1017.2005.00134.x

61. Fulbrook P, Mooney S. Care bundles in critical care: a practical approach to evidence-based practice. Nurs Crit Care [Internet]. 2003 Nov;8(6):249–55. Available from: http://dx.doi.org/10.1111/j.1362-1017.2003.00039.x

62. Ghi T, Eggebø T, Lees C, Kalache K, Rozenberg P, Youssef A, et al. ISUOG Practice Guidelines: intrapartum ultrasound. Ultrasound Obstet Gynecol [Internet]. 2018 Jul;52(1):128–39. Available from: http://dx.doi.org/10.1002/uog.19072

63. Gordijn SJ, Beune IM, Thilaganathan B, Papageorghiou A, Baschat AA, Baker PN, et al. Consensus definition of fetal growth restriction: a Delphi procedure. Ultrasound Obstet Gynecol [Internet]. 2016 Sep;48(3):333–9. Available from: http://dx.doi.org/10.1002/uog.15884

64. Brown MA, Magee LA, Kenny LC, Karumanchi SA, McCarthy FP, Saito S, et al. The hypertensive disorders of pregnancy: ISSHP classification, diagnosis & management recommendations for international practice. Pregnancy Hypertens [Internet]. 2018 Jul;13:291–310. Available from: http://dx.doi.org/10.1016/j.preghy.2018.05.004

65. Ornaghi S, Bellante N, Abbamondi A, Maini M, Cesana F, Trabucchi M, et al. Cardiac and obstetric outcomes in pregnant women with heart disease: appraisal of the 2018 mWHO classification. Open Heart [Internet]. 2022 Mar;9(1):e001947. Available from: http://dx.doi.org/10.1136/openhrt-2021-001947

66. Menard MK, Main EK, Currigan SM. Executive summary of the reVITALize initiative: standardizing obstetric data definitions. Obstet Gynecol [Internet]. 2014 Jul;124(1):150–3. Available from: http://dx.doi.org/10.1097/AOG.0000000000000322

67. Committee on Practice Bulletins-Obstetrics. Practice Bulletin no. 183: Postpartum hemorrhage. Obstet Gynecol [Internet]. 2017 Oct;130(4):e168–86. Available from: http://dx.doi.org/10.1097/AOG.0000000000002351

68. Simkin P. Pain, suffering, and trauma in labor and prevention of subsequent posttraumatic stress disorder. J Perinat Educ [Internet]. 2011 Summer;20(3):166–76. Available from: http://dx.doi.org/10.1891/1058-1243.20.3.166

69. Nespoli A, Colciago E, Fumagalli S, Locatelli A, Hollins Martin CJ, Martin CR. Validation and factor structure of the Italian version of the Birth Satisfaction Scale-Revised (BSS-R). J Reprod Infant Psychol [Internet]. 2021 Nov;39(5):516–31. Available from: http://dx.doi.org/10.1080/02646838.2020.1836333

70. Barba M, Cola A, Melocchi T, Braga A, Castronovo F, Manodoro S, et al. Italian validation of the Pelvic Floor Distress Inventory (PFDI-20) questionnaire. Int Urogynecol J [Internet]. 2023 Oct;34(10):2459–65. Available from: http://dx.doi.org/10.1007/s00192-023-05572-8

71. Ferreira CHJ, Barbosa PB, de Oliveira Souza F, Antônio FI, Franco MM, Bø K. Inter-rater reliability study of the modified Oxford Grading Scale and the Peritron manometer. Physiotherapy [Internet]. 2011 Jun;97(2):132–8. Available from: http://dx.doi.org/10.1016/j.physio.2010.06.007

72. Volonte' S, Barba M, Cola A, Marino G, Frigerio M. Italian validation of the short form of the Pelvic Organ Prolapse/Urinary Incontinence Sexual Questionnaire (PISQ-12). Int Urogynecol J [Internet]. 2022 Nov;33(11):3171–5. Available from: http://dx.doi.org/10.1007/s00192-022-05235-0

73. Dietz HP. Pelvic floor ultrasound: A review. Clin Obstet Gynecol [Internet]. 2017 Mar;60(1):58–81. Available from: http://dx.doi.org/10.1097/GRF.0000000000000264

74. Dietz HP. Ultrasound imaging of the pelvic floor. Part I: two-dimensional aspects. Ultrasound Obstet Gynecol [Internet]. 2004 Jan;23(1):80–92. Available from: http://dx.doi.org/10.1002/uog.939

75. Santoro GA, Wieczorek AP, Dietz HP, Mellgren A, Sultan AH, Shobeiri SA, et al. State of the art: an integrated approach to pelvic floor ultrasonography. Ultrasound Obstet Gynecol [Internet]. 2011 Apr;37(4):381–96. Available from: http://dx.doi.org/10.1002/uog.8816

76. Frawley H, Shelly B, Morin M, Bernard S, Bø K, Digesu GA, et al. An International Continence Society (ICS) report on the terminology for pelvic floor muscle assessment. Neurourol Urodyn [Internet]. 2021 Jun;40(5):1217–60. Available from: http://dx.doi.org/10.1002/nau.24658

77. Fumagalli S, Antolini L, Nespoli A, Panzeri M, Terenghi T, Ferrini S, et al. Rebozo and advanced maternal postures: A promising set of intrapartum interventions to reduce persistent occiput posterior

position of the fetal head. Eur J Midwifery [Internet]. 2024 Oct 18;8. Available from: http://dx.doi.org/10.18332/ejm/191511

78. Rozenberg P, Porcher R, Salomon LJ, Boirot F, Morin C, Ville Y. Comparison of the learning curves of digital examination and transabdominal sonography for the determination of fetal head position during labor. Ultrasound Obstet Gynecol [Internet]. 2008 Mar;31(3):332–7. Available from: http://dx.doi.org/10.1002/uog.5267

79. Barrowclough J, Crowther C, Kool B. Midwives' views on the acceptability of a future trial of the Sims posture for fetal malposition in labor in the context of their knowledge and practice: A mixed-methods study. Eur J Midwifery [Internet]. 2022 Aug 1;6(August):50. Available from: http://dx.doi.org/10.18332/ejm/150377

80. Schulz KF, Altman DG, Moher D, CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. Obstet Gynecol [Internet]. 2010 May;115(5):1063–70. Available from: http://dx.doi.org/10.1097/AOG.0b013e3181d9d421

81. Butcher NJ, Monsour A, Mew EJ, Chan AW, Moher D, Mayo-Wilson E, et al. Guidelines for reporting outcomes in trial reports: The CONSORT-outcomes 2022 extension. JAMA [Internet]. 2022 Dec 13;328(22):2252–64. Available from: http://dx.doi.org/10.1001/jama.2022.21022

82. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA [Internet]. 2013 Nov 27;310(20):2191–4. Available from: http://dx.doi.org/10.1001/jama.2013.281053

83. National Institute of Health. Racial and Ethnic Categories and Definitions for NIH Diversity Programs and for Other Reporting Purposes. 2015; Available from: https://case.edu/nursing/sites/default/files/2018-06/Racial_Ethnic_Categories_Definitions_NIH_Diversity.pdf

5. ***Conclusion***

This thesis explored the use of principles of causal inference in the context of research in the field of midwifery and maternity care, combining methodological reasoning with clinically relevant questions that arise from midwifery practice. We have demonstrated throughout the different chapters that causal inference is not only a statistical framework for the analysis of data, but also a structured way of thinking for generating valid clinical knowledge from observational and experimental research designs.

The concept of exchangeability, which is a central theme throughout the thesis, supports causal interpretation of associations and needs to be explicitly considered in how comparability between groups is achieved. In observational studies where treatment assignment is not controlled by design, differences in baseline prognostic characteristics create confounding and threaten causal identification. In randomised studies, randomisation provides a theoretical guarantee of exchangeability, but finite-sample variability may still lead to residual imbalance in important prognostic factors. Within this framework, Directed Acyclic Graphs (DAGs), counterfactual reasoning and modern causal methods were used as complementary tools to formalise assumptions, clarify sources of bias and guide analytical choices. Importantly, the thesis highlights that valid causal inference cannot rely on statistical analysis alone. Rather, it emerges from the coherent integration of the causal question, the underlying causal structure represented through DAGs, the study design adopted to address that question, and the analytical methods used to estimate the effect of interest.

The thesis shows that the causal inference framework enables a connection between methodological rigour and clinically relevant questions in the context of midwifery research. Its use allows a more disciplined interpretation of maternal and intrapartum outcomes, especially in settings where randomised trials are not always practical or where care processes are affected by complex clinical and organisational structures. Causal inference is described not as a collection of statistical techniques, but as a broad scientific methodology framework that links the design of a study, the analytical approach and clinical reasoning. Across both the observational and randomized examples presented in this thesis, the estimation and interpretation of outcomes depended on maintaining coherence between these components. DAGs informed the identification of relevant prognostic factors and potential sources of bias, study design determined the degree of exchangeability achievable by design, and analytical methods were selected accordingly to address the remaining threats to causal validity.

In midwifery research where interventions, models of care and organisational contexts cannot be separated from women's physiological and experiential aspects, this framework supports the generation of evidence that is methodologically valid and clinically meaningful. In this sense, the thesis advocates a shift from purely associational thinking towards explicitly causal reasoning in the evaluation of maternity care. More broadly, it argues that robust evaluation of maternal and neonatal outcomes requires an integrated approach in which causal assumptions, study design and statistical analysis are considered as interconnected elements of the same inferential process rather than as independent methodological steps.

6. Research project during PhD

6.1. Publications

- ***Panzeri, M.****, Terenghi T, Ornaghi S, Borrelli S, Nespoli A, Locatelli A, Fumagalli S. Attitudes, interest and involvement of midwives in an RCT: a qualitative study with a phenomenological approach. Accepted in European Journal of Midwifery on 03/06/2026.

- Zagra L[#], ***Panzeri M***[#], Fumagalli S, Gramegna T, Nespoli A, Ornaghi S, Locatelli A. Exploring the Passive Second Stage of Labour and Related Perinatal Outcomes in a Physiological Cohort. Birth. 2026 May 2. doi: 10.1111/birt.70072. Epub ahead of print. PMID: 42068157.

- Garzon, S., Campetti, E., ***Panzeri, M.****, Fumagalli, S., Nespoli, A., Locatelli, A., Fruscio, R., Serafini, M., Antolini, L., Valletta, S., Adami, A., Bosco, M., Franchi, M. P., & Uccella, S. (2026). Cross-Cultural Adaptation and Validation of KidSIM Attitude Towards Teamwork in Training Undergoing Designed Educational Simulation (ATTITUDES) in Undergraduate Healthcare Professionals. Nursing open, 13(3), e70499. https://doi.org/10.1002/nop2.70499

- Borrelli, S., Fumagalli, S., ***Panzeri, M.***, Terry, R., Mink, M., Sturkenboom, S., & Spiby P. (2026). Online information available to women and families about antenatal education. *MIDIRS Midwifery Digest*, *36*(1), 58–69.

- Masè, C., ***Panzeri, M.****, Rovelli, N., Fumagalli, S., & Vaccari, S. (n.d.). *Work-Related Wellbeing, Organizational Support and Professional Recognition among Italian Midwives: Results from a National FNOPO Survey*. https://doi.org/10.82083/LUCINA_20263_295

- Ornaghi, S., Fumagalli, S., Antolini, L., ***Panzeri, M.****, Spandrio, R., Ferrini, S., Nespoli, A., Maini, M., & Locatelli, A. (2025). Rebozo and maternal postures to prevent persistent occiput posterior position of the fetal head: protocol for a randomised clinical trial 'the ReMaP-POPP RCT'. *BMJ open*, *15*(10), e103520. https://doi.org/10.1136/bmjopen-2025-103520

- Fumagalli, S., Nespoli, A., ***Panzeri, M.****, Antolini, L., Colciago, E., Adami, A., Canepa, M. M., Del Bo, E., Ferrara, R., Mauri, P. A., Cagnacci, A., Ceccaroni, M., Dattolo, C., Esposito, G., Franchi, M. P., Gorlero, F., Grisolia, G., Grosso, F., Lecis, A., Mazzeo Melchionda, M., … Locatelli, A. (2025). Effect of Maternity Units' Organizational Levels on Maternal Birth Satisfaction: A Multicentric Cohort Study. *Birth (Berkeley, Calif.)*, *52*(4), 641–651. https://doi.org/10.1111/birt.12909

- Cosmai, G., Biondini, M., ***Panzeri, M.****, Serafini, M., Lambicchi, L., Locatelli, A., Nespoli, A., & Fumagalli, S. (2025). The decision-making process in the choice of VBAC: Facilitators and barriers from women's perception, a systematic review. *European Journal of Midwifery*, *9*. https://doi.org/10.18332/ejm/205874

- Fumagalli, S., Torrenzano, G., Adami, A., ***Panzeri, M.****, De Lucia, A., Antolini, L., Caliari, E., Bassetti, S., Danza, M., Lopresti, E., Michelerio, V., Pellegrini, E., Nespoli, A., & Donisi, V. (2025). Validation of the Italian version of the Secondary Traumatic Stress Scale (STSS-I) within midwifery students. *Nurse education in practice*, *85*, 104342. https://doi.org/10.1016/j.nepr.2025.104342

- Fumagalli, S., ***Panzeri, M.****, Borrelli, S., Ornaghi, S., Antolini, L., Sacco, G. G. A., Formenti, M., Sala, S., & Nespoli, A. (2024). Midwives' stress and burnout during the Omicron wave in Italy: An observational survey. *Journal of advanced nursing*, *80*(11), 4510–4522. https://doi.org/10.1111/jan.16065

- Fumagalli, S., Antolini, L., Nespoli, A., ***Panzeri, M.****, Terenghi, T., Ferrini, S., Spandrio, R., Maini, I. M., Locatelli, A., & Ornaghi, S. (2024). Rebozo and advanced maternal postures: A promising set of intrapartum interventions to reduce persistent occiput posterior position of the fetal head. *European Journal of Midwifery*, *8*. https://doi.org/10.18332/ejm/191511

- Fumagalli, S., Nespoli, A., ***Panzeri, M.****, Pellegrini, E., Ercolanoni, M., Vrabie, P. S., Leoni, O., & Locatelli, A. (2024). Intrapartum quality of care among healthy women: A population-based cohort study in an Italian region. *International Journal of Environmental Research and Public Health*, *21*(5), 629.

- Nespoli, A., Sacco, G. G. A., Bouhachem, F. Z., Motta, F., Paredi, S., Antolini, L., Panzeri, M., Pellegrini, E., & Fumagalli, S. (2024). Assessment of the psychometric properties of the italian version of the midwifery student evaluation of practice (MIDSTEP-IT): A validity and reliability study. *Midwifery*, *133*, 103991. https://doi.org/10.1016/j.midw.2024.103991

*Corresponding author
[#] Co-first author

6.2. Research under review

- Socio-demographic, obstetric, and midwifery care impacts on perinatal outcomes: evidence from a multicentric prospective study focusing on maternal satisfaction. Submitted to *Birth* on 14/05/2025.

- Epidural analgesia during labour in healthy childbearing women: intrapartum care and maternal–neonatal outcomes. A retrospective cohort study. Submitted to *BMJOpen* on 27/01/2026.

- Assessing Pain Coping in Labour: Inter-Rater Reliability of the Pain Coping Scale Among Midwives. Submitted to *Birth* on 10/10/2025.

- Understanding clinical teaching in midwifery education: validation of the MidPaACT tool in the Italian context. Submitted to *Interdisciplinary Journal of Nursing and Health* on 18/04/2026.

- Healthcare students as research assistants: a phenomenological study. Submitted to *Nurse Education Today* on 05/06/2026

7. Appendix 1

STATA script used to generate the parity-stratified randomization lists for the ReMaP-POPP Randomized Controlled Trial.

```
**************************************************
clear all
set seed 123456
**************************************************
*RANDOMISATION LIST (Parity = 0)
**************************************************
tempfile list0 list1
*----------------------------
* Create empty dataset
*----------------------------
clear
set obs 500
gen parity = 0                    // Nulliparous women
gen block = .
gen trt = ""
*----------------------------
* Block generation procedure
* Blocks of variable size (4, 6, 8)
* ensuring total sample = 500
*----------------------------
local total = 0
local blocknum = 0
while `total' < 500 {
      local b = 4 + 2*runiformint(0,2)   // block size: 4, 6 or 8
    if (`total' + `b') <= 500 {
        local ++blocknum
        local blocksize`blocknum' = `b'
        local total = `total' + `b'
    }
}
*----------------------------
* Allocation within blocks
* 1:1 allocation (Standard care vs Intervention)
* Random ordering within each block
*----------------------------
local start = 1
forvalues i = 1/`blocknum' {
```

```
    local b = `blocksize`i''
    local half = `b'/2
    local end = `start' + `b' - 1
    replace block = `i' if _n >= `start' & _n <= `end'
    replace trt = "Standard care" if _n >= `start' & _n < (`start' +
`half')
    replace trt = "Intervention"  if _n >= (`start' + `half') & _n <= `end'
    gen u = runiform()
    sort block u
    drop u
    local start = `end' + 1
}
* Participant identifier
gen idprog = _n
save `list0', replace
************************************************
* RANDOMISATION LIST (Parity = 1)
************************************************
clear
set seed 12345610
set obs 500
gen parity = 1                      // Multiparous women
gen block = .
gen trt = ""
local total = 0
local blocknum = 0
while `total' < 500 {
    local b = 4 + 2*runiformint(0,2)
    if (`total' + `b') <= 500 {
        local ++blocknum
        local blocksize`blocknum' = `b'
        local total = `total' + `b'
    }
}
local start = 1
forvalues i = 1/`blocknum' {
    local b = `blocksize`i''
    local half = `b'/2
    local end = `start' + `b' - 1
    replace block = `i' if _n >= `start' & _n <= `end'
    replace trt = "Standard care" if _n >= `start' & _n < (`start' +
`half')
```

```
    replace trt = "Intervention"  if _n >= (`start' + `half') & _n <= `end'
    gen u = runiform()
    sort block u
    drop u
    local start = `end' + 1
}
gen idprog = _n
save `list1', replace
***********************************************
* EXPORT FINAL RANDOMISATION LIST
***********************************************
use `list0', clear
export excel using "`c(pwd)'/ReMaP_POP_randomisation_list.xlsx", ///
    sheet("Parity_0") firstrow(variables) replace
use `list1', clear
export excel using "`c(pwd)'/ReMaP_POP_randomisation_list.xlsx", ///
    sheet("Parity_1") firstrow(variables) sheetreplace
```

8. References

Barrowclough, J. A., Lin, L., Kool, B., Hofmeyr, G. J., & Crowther, C. A. (2022). Maternal postures for fetal malposition in labour for improving the health of mothers and their infants. *Cochrane Database of Systematic Reviews*, *8*(8), CD014615. https://doi.org/10.1002/14651858.CD014615

Caughey, A. B., Sharshiner, R., & Cheng, Y. W. (2015). Fetal malposition: impact and management. *Clinical Obstetrics and Gynecology*, *58*(2), 241–245. https://doi.org/10.1097/GRF.0000000000000106

Cheng, Y. W., Shaffer, B. L., & Caughey, A. B. (2006). Associated factors and outcomes of persistent occiput posterior position: A retrospective cohort study from 1976 to 2001. *The Journal of Maternal-Fetal & Neonatal Medicine: The Official Journal of the European Association of Perinatal Medicine, the Federation of Asia and Oceania Perinatal Societies, the International Society of Perinatal Obstetricians*, *19*(9), 563–568. https://doi.org/10.1080/14767050600682487

Cohen, S. R., & Thomas, C. R. (2015). Rebozo technique for fetal malposition in labor. *Journal of Midwifery & Women's Health*, *60*(4), 445–451. https://doi.org/10.1111/jmwh.12352

Digitale, J. C., Martin, J. N., & Glymour, M. M. (2022). Tutorial on directed acyclic graphs. *Journal of Clinical Epidemiology*, *142*, 264–267. https://doi.org/10.1016/j.jclinepi.2021.08.001

Fumagalli, S., Antolini, L., Nespoli, A., Panzeri, M., Terenghi, T., Ferrini, S., … Ornaghi, S. (2024). Rebozo and advanced maternal postures: A promising set of intrapartum interventions to reduce persistent occiput posterior position of the fetal head. *European Journal of Midwifery*, *8*. https://doi.org/10.18332/ejm/191511

Fumagalli, S., Colciago, E., Antolini, L., Riva, A., Nespoli, A., & Locatelli, A. (2021). Variables related to maternal satisfaction with intrapartum care in Northern Italy. *Women and Birth: Journal of the Australian College of Midwives*, *34*(2), 154–161. https://doi.org/10.1016/j.wombi.2020.01.012

Gardberg, M., Laakkonen, E., & Sälevaara, M. (1998). Intrapartum sonography and persistent occiput posterior position: a study of 408 deliveries. *Obstetrics and Gynecology*, *91*(5 Pt 1), 746–749. https://doi.org/10.1016/s0029-7844(98)00074-x

Ghi, T., Youssef, A., Martelli, F., Bellussi, F., Aiello, E., Pilu, G., … Rizzo, G. (2016). Narrow subpubic arch angle is associated with higher risk of persistent occiput posterior position at delivery. *Ultrasound in Obstetrics & Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, *48*(4), 511–515. https://doi.org/10.1002/uog.15808

Hernán, M. A., & Robins, J. M. (2016). Using big data to emulate a target trial when a randomized trial is not available. *American Journal of Epidemiology*, *183*(8), 758–764. https://doi.org/10.1093/aje/kwv254

Hernán, M. A., & Robins, J. M. (2020). *Causal inference: what if*. Boca Raton: Chapman & Hall/CRC. Retrieved from https://static1.squarespace.com/static/675db8b0dd37046447128f5f/t/677676888e31cc50c2c33877/1735816881944/hernanrobins_WhatIf_2jan25.pdf

Hochman, N., Galper, A., Stanger, V., Levin, G., Herzog, K., Cahan, T., … Meyer, R. (2023). Risk factors for a negative birth experience using the Birth Satisfaction Scale-Revised. *International Journal of Gynaecology and Obstetrics: The Official Organ of the International Federation of Gynaecology and Obstetrics*, *163*(3), 904–910. https://doi.org/10.1002/ijgo.14884

Iversen, M. L., Midtgaard, J., Ekelin, M., & Hegaard, H. K. (2017). Danish women's experiences of the rebozo technique during labour: A qualitative explorative study. *Sexual & Reproductive Healthcare: Official Journal of the Swedish Association of Midwives*, *11*, 79–85. https://doi.org/10.1016/j.srhc.2016.10.005

Joensuu, J. M., Saarijärvi, H., Rouhe, H., Gissler, M., Ulander, V.-M., Heinonen, S., … Mikkola, T. (2023). Effect of the maternal childbirth experience on a subsequent birth: a retrospective 7-year cohort study of primiparas in Finland. *BMJ Open*, *13*(3), e069918. https://doi.org/10.1136/bmjopen-2022-069918

Johansson, C., & Finnbogadóttir, H. (2019). First-time mothers' satisfaction with their birth experience - a cross-sectional study. *Midwifery*, *79*(102540), 102540. https://doi.org/10.1016/j.midw.2019.102540

Lewis, L., Hauck, Y. L., Ronchi, F., Crichton, C., & Waller, L. (2016). Gaining insight into how women conceptualize satisfaction: Western Australian women's perception of their maternity care experiences. *BMC Pregnancy and Childbirth*, *16*(1), 29. https://doi.org/10.1186/s12884-015-0759-x

Ministero della salute. (2025). *Certificato di assistenza al parto, per la rilevazione dei dati di sanita' pubblica e statistici di base relativi agli eventi di nascita, di nati-mortalita' ed ai nati affetti da patologie congenite*.

Ministero della Sanità. (2010). *Accordo stato-regioni*.

Nilvér, H., Begley, C., & Berg, M. (2017). Measuring women's childbirth experiences: a systematic review for identification and analysis of validated instruments. *BMC Pregnancy and Childbirth*, *17*(1), 203. https://doi.org/10.1186/s12884-017-1356-y

Ornaghi, S., Fumagalli, S., Antolini, L., Panzeri, M., Spandrio, R., Ferrini, S., … Locatelli, A. (2025). Rebozo and maternal postures to prevent persistent occiput posterior position of the fetal head: protocol for a randomised clinical trial "the ReMaP-POPP RCT." *BMJ Open*, *15*(10), e103520. https://doi.org/10.1136/bmjopen-2025-103520

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350

Srivastava, A., Avan, B. I., Rajbangshi, P., & Bhattacharyya, S. (2015). Determinants of women's satisfaction with maternal health care: a review of literature from developing countries. *BMC Pregnancy and Childbirth*, *15*(1), 97. https://doi.org/10.1186/s12884-015-0525-0

VanderWeele, T. (2015). *Explanation in causal inference: Methods for mediation and interaction*. *706*. Retrieved from https://psycnet.apa.org/record/2015-10736-000

World Health Organization. (2016). *Standards for improving quality of maternal and newborn care in health facilities*. Geneva: World Health Organization. Retrieved from https://cdn.who.int/media/docs/default-source/mca-documents/advisory-groups/quality-of-care/standards-for-improving-quality-of-maternal-and-newborn-care-in-health-facilities.pdf

World Health Organization. (2018). *WHO recommendations on intrapartum care for a positive childbirth experience*. Genève, Switzerland: World Health Organization. Retrieved from https://iris.who.int/bitstream/handle/10665/260178/9789241550215-eng.pdf

World Medical Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *JAMA: The Journal of the American Medical Association*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.2013.281053

La borsa di dottorato cofinanziata con risorse dell'Unione europea-*NextGeneration EU*
Piano Nazionale di Ripresa e Resilienza Missione 4 – Componente 1 – Riforma 4.1 Riforma dei Dottorati – Inv. 4.1
Borse PNRR patrimonio Culturale –*CUP: H43C22000830006*