Abstract

Objective

This study aimed to develop and validate a midwifery core outcome set (M-COS) for Italian settings based on a salutogenic framework of maternity care.

Design

A multi-phase and multi-method study was performed. In phase one, we conducted a literature review to identify a preliminary set of outcomes sensitive to midwifery care. In phase two, the qualitative and quantitative content validity of the M-COS was tested. Finally, in the third phase, construct validity was explored through a cross-sectional study to assess the psychometric properties of the M-COS through exploratory and confirmative factor analysis. This study was conducted from December 2019 to April 2020 in Italy.

Participants

Three main groups of experts/midwives were involved. Group One (n=10) was involved in the content validity phase, while the other two groups (Group Two and Group Three) were involved in the construct validity phase (n=300).

Results

The M-COS includes six outcome domains and thirty-one core outcomes perceived as sensitive to midwifery care, namely: mortality and morbidity (n=6 outcomes), childbirth (n=3), postnatal period (n=6), maternal health (n=11), maternal-infant bonding (n=3), and maternal self-care (n=2). All domains showed good evidence of internal consistency.

Conclusion

The Italian M-COS is a novel tool that will facilitate the consistent measurement of core outcomes sensitive to midwifery care from the antenatal to the postnatal period in Italian settings. This initial work will be followed by further studies, including validation by service users.

Implications for practice

The use of the M-COS in clinical practice would facilitate evidence-based data collection and thus contribute to promoting high-quality maternity care.

Keywords: maternal and newborn health, midwifery, quality indicators, core outcome set.

1.0 Introduction

A core outcome set (COS) is an agreed, standardised set of outcomes that should be measured and reported, as a minimum, in research studies for specific areas of health or healthcare (Williamson et al., 2020, 2012). Thus, a COS ensures that care quality, patient safety, and workforce measures are reported in a standard way in healthcare systems. Awareness of the need for standardised and shared COSs specific to maternity care has increased over the past decade among researchers and clinicians (Duffy et al., 2017). A recent scoping review reported that the number of COSs developed for maternal and neonatal healthcare is currently growing worldwide (Slavin et al., 2021). However, the review authors identified low consistency in outcome definitions and low adherence to the minimum standards for COSs development as the main limits for implementing COSs in routine clinical settings (Slavin et al., 2021). These issues could undermine the ability to compare and synthesise research outcomes and require to be addressed to improve care for healthy childbearing women.

Although the majority of the available COSs in maternal and neonatal care is related to illness or adverse event prevention, researchers have increasingly drawn upon a salutogenic framework of maternity care, emphasizing the achievement of positive health and wellbeing outcomes (Lazzaretto et al., 2018; Murphy and Fullerton, 2006; Smith et al., 2014). This paradigm shift reflects a renewed emphasis on the full scope of midwifery care including midwives' role in promoting normal birth (Escuriet et al., 2015; Sandall et al., 2016). Therefore, in developing a new COS, it is important to consider midwifery-sensitive outcomes and include them as proxy quality indicators of midwifery care. It is also essential to take the characteristics of the national and clinical contexts into account to develop a relevant and implementable midwifery core outcome set (M-COS) (Williamson et al., 2017).

Thus far, the M-COSs using a salutogenic framework are still poorly developed worldwide (Slavin et al., 2021). Precisely, developing M-COSs with a salutogenic framework means to shift the attention from the care of at-risk women to the care that enhances wellbeing operationalized in terms of mental health, perceived health, and general quality of life (Lazzaretto et al., 2018; Murphy and Fullerton, 2006; Smith et al., 2014). In this regard, the International Consortium for Health Outcomes Measurement (ICHOM) has recently developed the first international COS for perinatal care using a salutogenic approach, which needs to be adapted and pilot tested in diverse health care settings (Nijagal et al., 2018). However, there is no locally-developed M-COS focused on the Italian context; this gap undermines the possibility of highlighting the contribution of midwives within the healthcare system. A context-specific M-COS is pivotal to address and measure the impact of midwifery care in supporting health and wellbeing for users of maternity services, enabling midwives, in the Italian context, to implement quality improvement strategies targeting poor practice (Nijagal et al., 2018).

1.1 Objective

The objectives of this study were: (a) to identify the core outcomes previously reported in the literature and develop a preliminary M-COS reflecting the outcomes that midwives, within the Italian context, perceive as sensitive to their clinical practice (i.e., midwifery-sensitive outcomes), and (b) to initially validate the psychometric structure of the preliminary M-COS among a representative group of midwives. Specifically, the M-COS scope covers healthy childbearing women with low-risk pregnancies and their newborns and is intended to be used for both clinical and research purposes in the Italian context.

2.0 Methods

We conducted a multi-method and multi-phase study, consistently with the recommendations of The COMET Handbook: version 1.0 (Williamson et al., 2017) and the COS-Standards for Development (COS-STAD) (Kirkham et al., 2017). The reporting of this study is in line with the COS-Standards for reporting (COS-STAR) (Kirkham et al., 2016) (as evidenced by **Supplementary File 1**). The study comprised two main phases conducted from December 2019 to April 2020. Phase one focused on the development of a preliminary evidence-based M-COS through a literature review. Phase two focused on the initial validation process of the M-COS, which encompassed two sub-phases: content validity and construct validity. **Figure 1** summarises the process of developing and validating the M-COS.

Please insert Figure 1

2.1 Developmental phase (phase one): Literature review

Based on the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) statement and flow chart, a literature review was performed to identify a preliminary set of midwifery sensitive outcomes (Page et al., 2021). We searched three main databases: PubMed, CINAHL, and Scopus, using the following keywords: *maternal and newborn health*, *midwifery*, *quality indicators*, *core outcome set*. Study selection was based on the following inclusion criteria: (a) papers in the English language with full-text availability without time limits; (b) all types of study design; (c) studies describing maternal and fetal/neonatal outcomes during pregnancy, childbirth, and postnatal period that are sensitive to midwifery care delivered in any care setting; (d) studies conceived within a physiological framework of maternity care. The studies were screened independently by two authors in the eligibility phase to identify and synthesize the outcomes sensitive to midwifery care. In the inclusion phase, any divergences among the authors were solved through a consensus discussion, or if necessary, a third reviewer was consulted. Finally, a parallel search was conducted on the COMET, CROWN, and ICHOM databases to assess existing COSs

relevant to maternity care and hand-searched the bibliography of relevant papers to identify any additional references. This phase was conducted from December 2019 to February 2020. The study selection process is shown in **Figure 2**.

Eleven papers were included, namely: two COSs for maternity care (Devane et al., 2007; Nijagal et al., 2018), five observational studies (Cheyney et al., 2014; Collins-Fulea et al., 2005; Greener, 1991; Ickovics et al., 2019; Murphy and Fullerton, 2001) and four systematic reviews (Escuriet et al., 2015; Lazzaretto et al., 2018; Moorhead et al., 2018; Smith et al., 2014). Data extraction was performed using a predefined data collection form: first author, year of publication, geographic area, study design, objective and outcomes sensitive to midwifery care, initially, by compiling a list of the relevant outcomes from the identified COSs and then adding pertinent outcomes from the other studies, deleting duplicates as appropriate. The selected studies and the outcomes they each included are listed in **Supplementary File 2**.

The preliminary evidence-based M-COS resulting from the literature review contained 37 outcomes sensitive to midwifery care, grouped into three outcome domains: (a) clinical core outcomes referring to the maternal and fetal/neonatal state of health; (b) functional core outcomes referring to maternal and neonatal physical and psychosocial functions; (c) self-reported core outcomes referring to women's experience of care.

Please insert Figure 2

2.2 Validation phase (phase two)

2.2.1 Content validity

The preliminary M-COS was shared among the author group through a peer debriefing approach (Nowell et al., 2017). Subsequently, the preliminary M-COS was shared with a larger group of experts (Group One) to assess each outcome's relevance or susceptibility to improvement through midwifery care (quantitative content validity) and the comprehensiveness of the

preliminary M-COS as a whole (qualitative or face content validity). This phase was conducted in February 2020.

2.2.1.1 Participants in Group One

The panel of experts was selected based on the following criteria: (a) relevant experience in midwifery care, (b) documented involvement in professional education, and (c) experience in COSs development. The minimum number of experts was determined as the minimum size needed to have sufficient control over the chance agreement (Polit and Beck, 2014). A purposive sampling procedure was used to enrol experts from two universities in northern Italy. The final panel included 10 experts: 7 midwives and 3 healthcare researchers with a scientific background in public health. Specifically, most experts involved were females (n=9; 90%) with a mean age of 36.8 years (±SD=12.96). The educational backgrounds of the involved experts ranged from bachelor degree or equivalent diploma (n=2; 20%) to postgraduate education (n=6; 60%) and doctoral studies (n=2; 20%). Those with work experience of more than 10 years represent 50% of the group.

2.2.1.2 Consensus process and content validity

To quantitatively assess the degree of agreement reached among the experts, the Content Validity Ratio (CVR) and Content Validity Index (CVI) were computed to assess the pertinence and relevance of each outcome to the construct of interest (i.e., the outcome domains), respectively (Ayre and Scally, 2014; Polit and Beck, 2014). As per the CVR, each expert rated each outcome's pertinence on a three-point Likert scale (1=not necessary; 2=useful, but not essential; 3=essential). The CVR's critical values were defined considering the exact binomial distribution applied to the panel size to determine the CVR score indicating that the level of agreement among rates was greater than 50% (Wilson Vanvoorhis and Morgan, 2007). The outcomes' relevance was assessed at the item-level (I-CVIs) on a four-point Likert scale (1=not relevant; 2=somewhat relevant; 3=quite relevant; 4=highly relevant), and it was computed as the number of experts giving the highest rating

(i.e., ≥ 3) divided by the total number of experts. After that, the content validity was also computed at the scale level (S-CVI) through the mean values of I-CVIs. The critical threshold values are equal to 0.70 for I-CVIs and 0.80 for S-CVI (Polit and Beck, 2014). Finally, to assess qualitative content validity, the following open-ended questions were asked of the group: «*Do the proposed outcomes exhaustively cover the scope of midwifery practice? Does the adopted wording clearly describe the meaning of each outcome? Should any additional core outcomes be included*?». We carried out the content analysis of the answers to assess each item's clarity and comprehensibility (Hsieh and Shannon, 2005). All data were collected through an online survey sent to an expert panel.

2.2.2 Construct validity

The second sub-phase of validating the M-COS was carried out on the results of the previous sub-phase (i.e., content validity). This sub-phase aimed to identify the most plausible latent structure (i.e., outcome domains) that could best explain the observed common variance between midwives' responses when asked to rate the specificity and sensitivity of each outcome to their practice. These data were collected by a cross-sectional survey from February 2020 to March 2020 and were used to assess the psychometric characteristics of the M-COS through Exploratory Factor Analysis (EFA), Confirmative Factor Analysis (CFA) and reliability testing. This approach enabled the development of a theory-grounded and bottom-up taxonomy of the outcomes into domains.

2.2.2.1 Participants in Group Two and Group Three

Data was collected from two convenience samples of experts for the factor analysis. The data from the first sample (Group Two of experts) was used to carry out EFA and from the second sample (Group Three of experts) to carry out CFA and reliability tests. The enrolled experts were all midwives practicing in Italy with at least six months' work experience and willing to participate in the study. All eligible midwives were informed about the study's aim, data management policy,

and publication policy before giving informed consent. The desired size for each sample (n=150) was established based on the rule-of-thumb suggesting that 50 participants are needed for each hypothetical outcome domain (Wilson Vanvoorhis and Morgan, 2007). Accordingly, the final desirable sample size was equal to 300 midwives [(Group2$_{50\ midwives}$ * 3 hypothetical outcome domains) + ([(Group3$_{50\ midwives}$ * 3 hypothetical outcome domains)]. The data collection included socio-demographic and professional data (i.e., sex, age, education, years of experience, clinical setting and workplace's geographical area). Data were collected through an online survey.

The overall sample (i.e., Group Two and Group Three) was composed of 300 midwives. Group Two (n=149) was mostly female (n=148; 99.3%) with a bachelor degree (n=71; 48.6%) and more than 10 years' work experience (n=76; 21.7%). Respondents mainly worked in hospital settings (n=105; 70.5%) and in northern Italy (n=122; 82.4%). Their median age was 34 years (IQR=15). Group Three (n=151) was mostly female (n=150; 99.3%) with a bachelor degree (n=90; 59.6%) and more than 10 years' work experience (n=61; 40.4%). Respondents mainly worked in hospital settings (n=104; 69.3%) and in northern Italy (n=122; 83.0%). Their median age was 32 years (IQR=12). **Table 1** presents the descriptive characteristics of the samples.

Please insert Table 1

2.2.2.2 Consensus process and construct validity analyses

The expert groups involved in this sub-phase were asked to assess the dependence of each outcome on midwifery care, using a 5-point Likert scale (1=not dependent; 5=totally dependent). The question was: *«How much do you think this outcome is sensitive to midwifery care?»*. Before conducting the EFA, the Kaiser-Meyer-Olkinindex and Bartlett's tests were performed to determine the adequacy of the sample and its factoriality. EFA was performed using the Weighted Least Squares Estimator (WLSM) and an Oblique Factor Rotation (Geomin Rotation), considering a moderate skewness in the answers' sample distribution. The analysis of the following criteria

guided the selection of the number of M-COS domains to be extracted: (a) eigenvalues, (b) the screen test interpretation, (c) the percentage of variance represented and (d) appropriate values of fit indices indicating a plausible factor solution to explain the sample statistics. The fit indices considered were: the Root Mean Square Error of Approximation (RMSEA<0.08 indicate a good fit), Comparative Fit Index (CFI>0.90 indicate a good fit), Tucker-Lewis Index (TLI>0.9 indicate a good fit); Standardized RMR (SRMR<0.06 indicate a good fit). Once a suitable solution from the EFA was obtained, a CFA was subsequently performed on the answers derived from another group to corroborate the most plausible factor structure using the same estimator adopted in the EFA (WLSM). In this phase, the authors had to achieve a consensus about theoretical interpretability, supported by testing alternative factor solutions. The latter was compared with the original using the Satorra-Bentler scaled chi-square (χ^2) difference test, where non-significant differences indicated that alternative structures explained the data equally well. Internal consistency was computed using Cronbach's α for each outcome domain. The analyses were conducted using SPSS software version 22 (SPSS Inc., Chicago, IL, USA) and Mplus version 8.1.

2.2.3 Ethics and consent

The Institutional board committee of the University of Pavia (Italy) approved the study protocol. In this study, only healthcare professionals were involved as experts and clear information about the study aims and methods were provided before obtaining informed consent. Data collection was conducted following good clinical practice (GPC) and the European regulation for data protection and privacy (GDPR 2016/679), using online survey software (SurveyMonkey®TM) that is compliant with GDPR data security standards, and data were stored in a protected and GDPR-compliant cloud (SOC 2 accredited data centre). Only two authors had access to the data, which were protected with a multi-layer strategy to ensure privacy: data protection mandatory

education for researchers who had access to data, policies to password strength, reuse and expiration, and account verification.

3.0 Results

3.1 Content validity and consensus

Consensus discussion among the authors led to several changes to the preliminary M-COS: dropping certain outcomes because they did not refer to evidence-based interventions within a physiological framework and further specification, linguistic reformulation or translation of others. Thus, 31 out of 37 outcomes identified by the literature were submitted to Group One for content validity assessment. Group One assessed all proposed outcomes as pertinent to midwifery care (CVR >0.80). Likewise, all outcomes reached adequate threshold values for relevance both at item-level (I-CVI>0.90) and at overall COS level (S-CVI >0.98) (CVR and CVI scores are reported in **Supplementary Files 3 and 4,** respectively). Finally, qualitative content analysis of the text responses did not reveal any critical issues regarding the clarity and comprehensibility of the outcomes and no further outcomes were added.

3.2 Construct Validity and consensus

The correlation matrix derived from Group Two's answers (n=149) determined that the responses were suitable for EFA as the Bartlett's test of sphericity was significant ($\chi^2_{(465)}$=2322.278; P<0.001) and KMO=0.867. Based on the hypothesized theoretical structure of M-COS derived from the previous phase, we tested a model with three outcome domains. However, this model reported poor fit in explaining the variance in midwives' perception of each outcome as sensible to their practice ($\chi^2_{(375)}$=1011.106, *P*<0.001; χ^2/df=2.7; RMSEA=0.107; 90%CI[0.099-0.115]; CFI=0.790; TLI=0.740; SRMR=0.055). Therefore, based on the authors' interpretation of a possible latent structure explaining the obtained answers, the interpretation of the screen test and the eigenvalues, a

structure of seven outcome domains was tested. The seven outcome domains model showed exhaustive fit indices ($\chi^2_{(269)}$=524.726, *P*<0.001; χ^2/df=1.9; RMSEA=0.08; 90%CI[0.070-0.091], p < 0.000; CFI=0.916; TLI=0.854; SRMR=0.035), thus appearing as the most plausible domain structure for M-COS, explaining 61.91% of the answers' total variance. As reported in **Table 2**, factor loadings indicated a theoretical structure that could be intuitively interpreted.

We labelled the outcome domains based on the included outcomes: maternal mortality and stillbirth ($N_{outcomes}$= 3); maternal and perinatal morbidity ($N_{outcomes}$= 3); childbirth ($N_{outcomes}$= 3); postnatal period ($N_{outcomes}$= 6); maternal health ($N_{outcomes}$= 11); maternal-infant bonding ($N_{outcomes}$= 3); maternal self-care ($N_{outcomes}$= 2). Internal consistency supported the identified factor structure, given that it was adequate for each outcome domain: maternal mortality and stillbirth had Cronbach's α=0.806, maternal and perinatal morbidity had Cronbach's α=0.875, childbirth had Cronbach's α=0.802, postnatal period had Cronbach's α=0.828, maternal health had Cronbach's α=0.895, maternal-infant bonding had Cronbach's α=0.866 and maternal self-care had Cronbach's α=0.913.

The confirmatory and unconstrained model based on Group Three's responses was adequate in indicating a seven outcome domains structure as suitable in explaining the variance in perceived sensitivity of the outcomes to midwifery care ($\chi^2_{(413)}$=797.059, *P*<0.001; χ^2/df=1.9; RMSEA=0.079; 90%CI[0.071-0.087, p< 0.000]; CFI=0.940; TLI=0.932; WRMR=1.199). As reported in **Table 2**, the tested underlying structure (hypothetical outcome domains) significantly explained the variance for each outcome and explained 58.4% of the total variance.

In interpreting the content for each hypothetical outcome domain, due to their conceptual overlap the authors proposed to merge the first two hypothetical outcome domains (maternal mortality and stillbirth; maternal and perinatal morbidity) into a single domain covering mortality and morbidity. This choice was corroborated because the correlation between these hypothetical

outcome domains was the highest between any two domains (r=0.875; *P*<0.001). Thus, a second-order factor labelled as "mortality and morbidity" was tested to explain the variances of the two original domains and found to explain the data equally well ($\chi^2_{(417)}$=795.059, *P*<0.001; χ^2/df=1.9; RMSEA=0.078; 90%CI[0.070-0.086, p< 0.000]; CFI=0.941; TLI=0.934). The authors agreed to include the second-order factor as an outcome domain. Internal consistency in the new version was adequate for each outcome domains: mortality and morbidity had Cronbach's α=0.838, childbirth had Cronbach's α=0.769, postnatal period had Cronbach's α=0.805, maternal health had Cronbach's α=0.845, maternal-infant bonding had Cronbach's α=0.845, and maternal self-care had Cronbach's α=0.924.

Please insert Table 2

Given the evidence derived from the factor analyses (EFA and CFA), all authors agreed in considering the identified six outcome domains as suitable (**Table 3**). The authors provide conceptual definitions for each outcome domain to ensure clarity in the interpretation and use of the M-COS.

Please insert Table 3

4.0 Discussion

4.1 Statement of principal findings

This study presents the results of the development and initial validation process of the first set of outcomes sensitive to midwifery care (M-COS) focused on the Italian maternity context. M-COS is a valid and reliable outcome set comprising 31 midwifery-sensitive core outcomes classified into six outcome domains: mortality and morbidity, childbirth, postnatal period, maternal health, maternal-infant bonding, and maternal self-care.

4.2 Strengths and limitations

In Italy, similarly to other countries, the outcomes reported in clinical documentation of maternity care are typically disease-oriented and lack an evidence-based approach (de Jonge et al., 2021; Renfrew et al., 2015). In this context, M-COS represents a first attempt to promote the standardised reporting of midwifery-sensitive outcomes within a salutogenic framework. Using a robust methodology and adhering to the international standards for COS development, this study provides a validated instrument that has the potential to enhance clinical practice, as it highlights the contribution of midwifery care within maternity care systems and facilitates further research on the quality of care with implications for maternal and neonatal health and wellbeing.

Several limitations require to be acknowledged and discussed. Firstly, while an adequate number of midwives and healthcare researchers were involved in the study, we did not consider women's perspectives at the initial validation stage covered by this paper. Therefore, caution is required when generalising the research results, and other studies involving the perspectives of women are needed to corroborate the M-COS. Secondly, it was beyond the scope of the study to provide instructions related to the timing and measurement instruments to be used for embedding the M-COS in routine data collection: however, further research on this aspect will be required. The shortage of valid and reliable patient-reported outcome measures (PROMs) available in Italian is a key contextual limitation, undermining the possibility of providing specific recommendations for the adoption of the M-COS in clinical practice. Specifically, most available PROMs are generic and not specific to maternity care (e.g., Patient Health Questionnaire-2) or focus on assessing the impact of disease or complications during pregnancy and childbirth (Dickinson et al., 2019). Moreover, there is no agreement at the international level on the most suitable PROMs to assess a number of core outcomes related to women's health (e.g., women's knowledge and levels of engagement) and self-care skills during pregnancy and postnatal period. Thirdly, we chose a multi-phase and multi-method study design favouring a quantitative approach. Although this approach satisfies the

minimum criteria for COS development, future studies should also integrate a qualitative approach, in particular when enquiring what aspects of quality of care are important to women/service users. Lastly, medical doctors have not been enrolled in the validation phases (e.g., content validity), and their point of view on midwifery-sensitive outcomes considering a salutogenic perspective might be relevant as they are close stakeholders of midwifery care.

4.3 Comparison with similar studies

The outcome domains and associated core outcomes included in the developed M-COS using a salutogenic framework are overall aligned with the COSs available in the literature for maternity care, considering both clinical and patient-reported outcomes from the antenatal to the postnatal period (Devane et al., 2007; Nijagal et al., 2018). This comparison means that despite the M-COS has being developed using a salutogenic framework to define the core outcomes, it also included additional outcomes from the salutogenesis perspective, including mortality and morbidity. In fact, mortality and morbidity are not embedded in Smith's salutogenic framework (Smith et al., 2014), and we agreed to keep this domain, following the recent ICHOM proposal suggesting that this could support the tracking and auditing of cases of maternal mortality (Nijagal et al., 2018). Likewise, the Apgar score is routinely recorded in clinical practice to assess the newborn infant's condition immediately after birth and is considered as a core outcome in several midwifery datasets and COSs (Collins-Fulea et al., 2005; Devane et al., 2007; Escuriet et al., 2015; Lazzaretto et al., 2018; Moorhead et al., 2018; Murphy and Fullerton, 2001).

4.4 Implications for clinical practice

Embedding the M-COS in routine data collection would enable comprehensive evaluations of the impact of midwifery care on maternal and fetal/neonatal health and wellbeing outcomes and sustain quality improvement initiatives within a salutogenic conceptualisation of midwifery care (Lazzaretto et al., 2018; Murphy and Fullerton, 2006; Smith et al., 2014). Given that a medical-led

model of care for all childbearing women is still prevalent across most of Italy, the integration of the M-COS in routine datasets may highlight the benefits of midwifery and midwife-led models of care, potentially resulting in encouraging more widespread adoption of such models (Ricchi et al., 2019; Svanera et al., 2017). Full involvement of key stakeholders such as service users and decision-makers will be undertaken in future studies to ensure the applicability and usability of M-COS, taking into consideration the resources available and health priorities identified at the national level. In addition, it is important to point out that the M-COS does not necessarily provide an exhaustive list of the outcomes that local stakeholders may deem important to collect.

4.5 Unanswered questions and future research

Three main areas require further study: a) exploring women's perceptions of what they perceive the goals of midwifery care to be in the Italian context; b) identifying the most suitable PROMs for use in maternity care; c) developing pilot studies to support the implementation of M-COS in clinical practice.

Further, to promote evidence-based care and boost the design, conduct and reporting of midwifery research, core outcome measures in midwifery should be embedded in an overall midwifery minimum dataset (MMDS) that includes both outcome indicators and intervention metrics. Thus far, most MMDSs have been developed in the United States by the leading national associations such as the American College of Nurse-Midwives, the Midwives Alliance of North America and the American Association of Birth Centers (Burlon et al., 2017). Next steps for improving the standardised reporting of midwifery care in Italy should be taken within the overarching aim of developing and validating an Italy-focused MMDS.

In this context, it is worth noting that many professional activities carried out by midwives are not recorded, undermining the possibility of providing evidence on the effectiveness and quality of care and avoiding unjustified or inappropriate interventions (de Jonge et al., 2021; Devane et al.,

2007; Lazzaretto et al., 2018). Future studies should thus be oriented to measure the midwifery care process (i.e., midwifery care interventions) underlying the achievement of midwifery core outcomes during maternity care.

5.0 Conclusion

This study provides initial evidence of the validity and reliability of an M-COS comprising key outcomes sensitive to midwifery care within a salutogenic maternity care framework in the Italian context. Further research is required to corroborate and complete the validation of the M-COS by exploring women's perspectives. The use of the M-COS in clinical practice could highlight the often under-valued midwifery contribution to the provision of quality and safe maternity care.

References

Ayre, C., Scally, A.J., 2014. Critical Values for Lawshe's Content Validity Ratio. Meas. Eval. Couns. Dev. 47, 79–86. https://doi.org/10.1177/0748175613513808

Burlon, B., D'Agostino, F., Colaceci, S., Cocchieri, A., Vellone, E., Zega, M., Alvaro, R., 2017. Standardization of data collection in midwifery care: a literature review. Ig. Sanita Pubbl. 73, 311–324.

Cheyney, M., Bovbjerg, M., Everson, C., Gordon, W., Hannibal, D., Vedam, S., 2014.

Development and validation of a national data registry for midwife-led births: The Midwives Alliance of North America Statistics Project 2.0 Dataset. J. Midwifery Women's Heal. 59, 8–16. https://doi.org/10.1111/jmwh.12165

Collins-Fulea, C., Mohr, J.J., Tillett, J., 2005. Improving midwifery practice: The American College of Nurse-Midwives' Benchmarking Project. J. Midwifery Women's Heal. 50, 461–471. https://doi.org/10.1016/j.jmwh.2005.06.007

de Jonge, A., Dahlen, H., Downe, S., 2021. 'Watchful attendance' during labour and birth. Sex. Reprod. Healthc. 28, 100617. https://doi.org/10.1016/j.srhc.2021.100617

Devane, D., Begley, C.M., Clarke, M., Horey, D., OBoyle, C., 2007. Evaluating Maternity Care: A Core Set of Outcome Measures. Birth 34, 164–172. https://doi.org/https://doi.org/10.1111/j.1523-536X.2006.00145.x

Dickinson, F., McCauley, M., Smith, H., Van Den Broek, N., 2019. Patient reported outcome measures for use in pregnancy and childbirth: A systematic review. BMC Pregnancy Childbirth 19, 1–8. https://doi.org/10.1186/s12884-019-2318-3

Duffy, J.M.N., Rolph, R., Gale, C., Hirsch, M., Khan, K.S., Ziebland, S., McManus, R.J., van't Hooft, J., Brown, M., Grobman, W., Fitzpatrick, R., Karumanchi, S., Lucas, N., Magee, L., Mol, B., Stark, M., Thangaratinam, S., Wilson, M., von Dadelszen, P., Williamson, P., 2017. Core outcome sets in women's and newborn health: a systematic review. BJOG An Int. J. Obstet. Gynaecol. https://doi.org/10.1111/1471-0528.14694

Escuriet, R., White, J., Beeckman, K., Frith, L., Leon-Larios, F., Loytved, C., Luyben, A., Sinclair, M., Van Teijlingen, E., 2015. Assessing the performance of maternity care in Europe: A critical exploration of tools and indicators. BMC Health Serv. Res. 15, 1–13. https://doi.org/10.1186/s12913-015-1151-2

Greener, D., 1991. Development and validation of the Nurse-Midwifery Clinical Data Set. J. Nurse.

Midwifery. 36, 174–180. https://doi.org/https://doi.org/10.1016/0091-2182(91)90004-9

Hsieh, H.F., Shannon, S.E., 2005. Three approaches to qualitative content analysis. Qual. Health Res. 15, 1277–1288. https://doi.org/10.1177/1049732305276687

Ickovics, J.R., Lewis, J.B., Cunningham, S.D., Thomas, J., Magriples, U., 2019. Transforming prenatal care: Multidisciplinary team science improves a broad range of maternal-child outcomes. Am. Psychol. 74, 343–355. https://doi.org/10.1037/amp0000435

Kirkham, J.J., Davis, K., Altman, D.G., Blazeby, J.M., Clarke, M., Tunis, S., Williamson, P.R., 2017. Core Outcome Set-STAndards for Development: The COS-STAD recommendations. PLoS Med. 14, 1–10. https://doi.org/10.1371/journal.pmed.1002447

Kirkham, J.J., Gorst, S., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E., Moher, D., Schmitt, J., Tugwell, P., Tunis, S., Williamson, P.R., 2016. Core Outcome Set–STAndards for Reporting: The COS-STAR Statement. PLoS Med. 13, 1–11. https://doi.org/10.1371/journal.pmed.1002148

Lazzaretto, E., Nespoli, A., Fumagalli, S., Colciago, E., Perego, S., Locatelli, A., 2018. Intrapartum care quality indicators: A literature review. Minerva Ginecol. 70, 346–356. https://doi.org/10.23736/S0026-4784.17.04177-6

Moorhead, S., Johnson, M., Maas, M.L., Swanson, E., 2018. Nursing Outcomes Classification (NOC). Measurement of health outcomes., 6th ed. Elsevier Inc.

Murphy, P.A., Fullerton, J.T., 2006. Development of the Optimality Index as a New Approach to Evaluating Outcomes of Maternity Care. JOGNN - J. Obstet. Gynecol. Neonatal Nurs. 35, 770–778. https://doi.org/10.1111/J.1552-6909.2006.00105.x

Murphy, P.A., Fullerton, J.T., 2001. Measuring outcomes of midwifery care: Development of an instrument to assess optimality. J. Midwifery Women's Heal. 46, 274–284. https://doi.org/10.1016/S1526-9523(01)00158-1

Nijagal, M.A., Wissig, S., Stowell, C., Olson, E., Amer-Wahlin, I., Bonsel, G., Brooks, A., Coleman, M., Karalasingam, S.D., Duffy, J.M.N., Flanagan, T., Gebhardt, S., Greene, M.E., Groenendaal, F., Ravichandran, J., Jeganathan, R., Kowaliw, T., Lamain-De-Ruiter, M., Main, E., Owens, M., Petersen, R., Reiss, I., Sakala, C., Speciale, A.M., Thompson, R., Okunade, O., Franx, A., 2018. Standardized outcome measures for pregnancy and childbirth, an ICHOM proposal. BMC Health Serv. Res. 18, 1–18. https://doi.org/10.1186/s12913-018-3732-3

Nowell, L.S., Norris, J.M., White, D.E., Moules, N.J., 2017. Thematic Analysis: Striving to Meet the Trustworthiness Criteria. Int. J. Qual. Methods 16, 1–13. https://doi.org/10.1177/1609406917733847

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, *372*. https://doi.org/10.1136/bmj.n71

Polit, D.F., Beck, C.T., 2014. Essentials of Nursing Research. Appraising Evidence for Nursing Practice, 8th ed. Wolters Kluwer Health | Lippincott Williams & Wilkins.

Renfrew, M., Burns, E., Gross, M., Symon, A., 2015. Pathways to strenghtening midwifery in Europe, EntreNous.

Ricchi, A., Rossi, F., Borgognoni, P., Bassi, M.C., Artioli, G., Foa, C., Neri, I., 2019. The midwifery-led care model: A continuity of care model in the birth path. Acta Biomed. 90, 41–52. https://doi.org/10.23750/abm.v90i6-S.8621

Sandall, J., Soltani, H., Gates, S., Shennan, A., Devane, D., 2016. Midwife-led continuity models versus other models of care for childbearing women, Cochrane Database of Systematic Reviews. John Wiley and Sons Ltd. https://doi.org/10.1002/14651858.CD004667.pub5

Slavin, V., Creedy, D.K., Gamble, J., 2021. Core Outcome Sets Relevant to Maternity Service Users: A Scoping Review. J. Midwifery Women's Heal. 1–18.

https://doi.org/10.1111/jmwh.13195

Smith, V., Daly, D., Lundgren, I., Eri, T., Benstoem, C., Devane, D., 2014. Salutogenically focused outcomes in systematic reviews of intrapartum interventions: A systematic review of systematic reviews. Midwifery 30, e151–e156. https://doi.org/10.1016/j.midw.2013.11.002

Svanera, C., Nelli, E., Colciago, E., Fumagalli, S., 2017. Midwifery-led care for normal childbirth pathway in Lombardy Region. Minerva Ginecol. 69, 115–117. https://doi.org/https://doi.org/10.23736/s0026-4784.16.03941-1

Williamson, P.R., Altman, D.G., Bagley, H., Barnes, K.L., Blazeby, J.M., Brookes, S.T., Clarke, M., Gargon, E., Gorst, S., Harman, N., Kirkham, J.J., McNair, A., Prinsen, C.A.C., Schmitt, J., Terwee, C.B., Young, B., 2017. The COMET Handbook: Version 1.0. Trials 18, 1–50. https://doi.org/10.1186/s13063-017-1978-4

Williamson, P.R., Altman, D.G., Blazeby, J.M., Clarke, M., Devane, D., Gargon, E., Tugwell, P., 2012. Developing core outcome sets for clinical trials: Issues to consider. Trials 13, 1–8. https://doi.org/10.1186/1745-6215-13-132

Williamson, P.R., De Ávila Oliveira, R., Clarke, M., Gorst, S.L., Hughes, K., Kirkham, J.J., Li, T., Saldanha, I.J., Schmitt, J., 2020. Assessing the relevance and uptake of core outcome sets (an agreed minimum collection of outcomes to measure in research studies) in Cochrane systematic reviews: A review. BMJ Open 10, 1–7. https://doi.org/10.1136/bmjopen-2019-036562

Wilson Vanvoorhis, C.R., Morgan, B.L., 2007. Understanding Power and Rules of Thumb for Determining Sample Sizes. Tutor. Quant. Methods Psychol. 3, 43–50. https://doi.org/https://doi.org/10.20982/tqmp.03.2.p043