Master Thesis for the Health Economics, Policy and Management

University of Oslo,

The Faculty of Medicine,

Department of Health Management and Health Economics

The Cost-Effectiveness of Screening Strategies for Perinatal Depression in Nepal

Oslo, 15th November 2023

Submitted by: Anup Neupane

Student Number: 651346

Thesis Supervisor: Dr Lumbwe Chola

Acknowledgments

I would like to express my sincere gratitude to all the extraordinary individuals who have supported me throughout the completion of this work.

First and foremost, I would like to thank Lumbwe for being a constant source of support and discussions to lead my thesis toward the right path.

I would also like to thank my friends and family and my wife, Manjila, who supported me through this extensive journey.

Table of contents

Tables

Figures

Abbreviations

ANC	Antenatal Care
ANM	Auxiliary Nurse Midwives
CBT	Cognitive Behavioural Therapy
CEAC	Cost-effectiveness Acceptability Curves
CEAF	Cost-effectiveness Acceptability Frontier
CHE	Current Health Expenditure
CIDI	Composite International Diagnostic Interview
CUA	Cost Utility Analysis
DALY	Disability Adjusted Life Year
DSM 5	Diagnostic and Statistical Manual of Mental Disorders 5th edition
EPDS	Edinburgh Postnatal Depression Scale
EVPI	Expected Value of Perfect Information
EVPPI	Expected Value of Partial Perfect Information
FCHVs	Female Community Health Volunteers
GAD	Generalized Anxiety Disorder
GDP	Gross Domestic Product
HYE	Healthy Years Equivalent
HTA	Health Technology Assessment
ICD-10	International Classification of Disease Tenth Revision
ICER	Incremental Cost Effectiveness Ratio
IHME	Institute for Health Metrics and Evaluation
IPT	Interpersonal Therapy
MDD	Major Depressive Disorder
mhGAP	mental health Gap Action Program
NMB	Net Monetary Benefit
NICE	National Institute for Health and Care Excellence
EVPI	Expected Value of Perfect Information
EVPPI	Expected Value of Partial Perfect Information
OCD	Obsessive Compulsive Disorder
OECD	Organization for Economic Cooperation and Development

PCA	Person Centred Approach
pEVPI	Population Expected Value of Perfect Information
PHQ	Patient Health Questionnaire
PHQ-9	Patient Health Questionnaire 9 items
PMADs	Perinatal Mood and Anxiety Disorders
PND	Perinatal depression
PSA	Probabilistic Sensitivity Analysis
PTSD	Post Traumatic Stress Disorder
QALY	Quality Adjusted Life Year
RCT	Randomized Controlled Trial
THHP	Thinking Healthy Program peer-delivered
THP	Thinking Healthy Program
USD	United States Dollar
VOI	Value of Information
WHO	World Health Organization
WTP	Willingness to Pay

Abstract

Objective: The prevalence of antenatal depression is estimated to be around 24.8% (95% CI 19.2% - 30.7%) in the Nepalese context. Perinatal depression leads to poor maternal and infant outcomes. Screening and treatment of perinatal depression offers the possibility of preventing adverse outcomes. This study explored the cost-effectiveness of screening alternatives during pregnancy compared to no screening.

Methods: A decision tree was developed to model the identification and treatment pathways of depression from the 4th to 8th antenatal care visit to 6 months postpartum using the Edinburgh Postnatal Depression Scale (EPDS) and the Patient Health Questionnaire 9 items (PHQ-9), compared to no screening. The economic evaluation took a healthcare perspective. Model parameters were taken from published literature. The outcome was assessed in terms of the incremental cost to detect, the incremental cost to treat, and the incremental cost per quality-adjusted life years (QALYs). Cost-effectiveness planes, cost-effectiveness acceptability curves (CEAC), cost-effectiveness acceptability frontier (CEAF), and population Expected Value of Perfect Information (pEVPI) were produced using a net-benefit approach based on Monte Carlo simulations of cost-outcome data for outcome with QALYs whereas for the outcome with treated cost-effectiveness plane were produced based on Monte Carlo simulations of cost-outcome data.

Results: The incremental cost to detect the case of depression was less than one USD for both screening tools. The incremental cost per treated for EPDS and PHQ-9 relative to no screening was USD 100 and USD 111, respectively. In terms of QALYs, the incremental cost per QALYs for EPDS and PHQ-9 relative to no screening was USD 649 and USD 716, respectively. Two-way CEAC and CEAF, EPDS vs. no screening, and PHQ-9 versus no screening showed that the screening alternatives were optimal choices compared to no screening at 1*Gross domestic product per capita (GDP) Willingness to Pay (WTP) threshold. There was value in collecting further information; at 1*GDP per capita WTP threshold for two-way model EPDS vs no screening and PHQ-9 versus no screening, the pEVPI was 2.99 million USD and 2.86 million USD.

Conclusion: The low incremental cost to detect and the incremental cost per treated ensured that the screening alternatives were cost-effective. Both screening alternatives were cost-effective compared to no screening for incremental cost per QALYs. Limitations include data availability and a short time horizon; thus, further research is needed.

Background

Context

Perinatal mood and anxiety disorders (PMADs) are mood and anxiety disorders that occur during pregnancy (Meltzer-Brody & Rubinow, 2021). PMADs are serious global health issues that negatively impact women's lives and additionally place a burden on their families and communities. A United States-based longitudinal cross-sectional study of national inpatient sample data showed women with PMAD experienced a higher incidence of severe maternal morbidity and mortality along with higher delivery costs associated with increased length of stay, hospital transfer, and other delivery-related costs (McKee et al., 2020). Perinatal depression (PND) is the most common type of PMAD, which encompasses major and minor depressive episode that occurs during pregnancy or within the first 12 months postpartum. A systematic review of maternal and infant consequences of postpartum depression showed that postpartum depression is associated with more negative maternal physical and psychological health with worse quality of life (Slomian et al., 2019). Bauer et al. reported that the offspring bear most of the social cost of PMAD (Bauer et al., 2016). More than 10% of pregnant women and women who have just given birth experience depression globally (World Health Organization, 2023a). The prevalence of PND differs by country and is higher in low and middle-income countries than in high-income countries. A systematic review with meta-analysis on the prevalence of PND in low and middle-income countries had a pooled prevalence of 24.7% (95% CI 23.7% - 25.6%) (Roddy Mitchell et al., 2023). PND in the literature is often differentiated into antenatal and postpartum depression. The PND prevalence range in Nepal is between 18% and 50% (Chalise et al., 2022; Joshi et al., 2019; Shakya et al., 2008; Stuge et al., 2022).

National guidance

At present, Nepal lacks the screening and treatment guidelines for perinatal depression. Perinatal depression is treated as a case of general population depression by healthcare professionals. Addressing the lack of guidelines for the treatment of mental health issues, the World Health Organization (WHO) provides the framework for the treatment of mental health issues in the mental health Gap Action Programme (mhGAP) intervention guide for mental, neurological, and substance use disorders in non-specialized health settings version 2.0 (World Health Organization, 2016). The WHO guidance focuses on pregnant and breastfeeding women as a special population and comments on any deviation and precautions that need to be considered while treating this special population, such as avoiding the use of

long-lasting medicines like fluoxetine. WHO recommendations on maternal and newborn care for a positive postnatal experience guide postnatal care in case of maternal depression (World Health Organization, 2022). The WHO recommends screening for postpartum depression and anxiety (recommendation 18) and prevention of postpartum depression and anxiety (recommendation 19). The guideline recommends the Edinburgh Postnatal Depression Scale (EPDS) or Patient Health Questionnaire-9 items (PHQ-9) to screen for common mental health conditions. Women with clinically significant symptoms or risk factors should be offered psychological interventions (e.g., cognitive behavioral therapy (CBT) or interpersonal therapy (IPT)).

Evidence on the cost-effectiveness of screening for depression in the perinatal period

In a systematic review of the screening programs coupled with treatment intervention for common maternal mental health disorders among perinatal women, the meta-analysis revealed a reduction in perinatal depression and anxiety among perinatal women undergoing screening programs (Waqas et al., 2022). Screening methods such as EPDS and PHQ 9 were used in the different studies within inclusion criteria, and the treatment interventions varied a lot among studies such as CBT, non-directive counseling, care plans, IPT, education plans were used in combination of screening techniques above in the systematic review above (Waqas et al., 2022).

In a cost-effectiveness analysis for screening and treating postpartum depression and psychosis in the USA, the incremental cost per quality-adjusted life years (QALY) gained and cost per remission were calculated; the incremental cost-effectiveness ratio (ICER) per QALYs was below the willingness to pay (WTP) threshold (A. Wilkinson et al., 2017). Similarly, the cost-effectiveness of screening for depression with EPDS was cost-effective compared to standard care in Canada (Premji et al., 2021). Mixed results are reported for the cost-effectiveness of screening in the UK setting in a cost-effectiveness study to assess routine screening in primary care, which showed it was not cost-effective to screen for postnatal depression (Paulden et al., 2009). In contrast, a recent cost-effectiveness analysis from the UK assessing the cost-effectiveness of screening tools for identifying depression in early pregnancy demonstrated that screening was cost-effective (Heslin et al., 2022).

In the context of Nepal, a parallel pilot randomized controlled trial (RCT) was conducted to assess the impact of psychosocial intervention to enhance mental health in women experiencing domestic and family violence in Nepal showed lower depressive symptoms in the intervention group with higher quality of life scores (Diksha et al., 2021). Instead, in a

general population setting, RCT on psychological intervention within services for depression delivered by primary care workers in Nepal was cost-effective under the 1* Gross domestic product (GDP) per capita WTP threshold (Aldridge et al., 2022). Despite some primary studies establishing the effect of mental health treatment intervention in the general population as well as a subgroup of pregnant women, no studies assess the impact of screening and treatment for PND.

Study objective

The objective of this study was to explore the cost-effectiveness of screening alternatives for perinatal depression compared to no screening in Nepal. The screening alternatives include EPDS and PHQ-9, and the outcome is explored in terms of the number of women treated and QALYs.

Methods

Target population and setting

The target population was pregnant women aged 16+ attending their monthly visits with Female Community Health Volunteers (FCHVs) in Nepal. FCHVs identify pregnant women in their catchment area as early as possible, and they help to distribute monthly supplies of iron and folic acid supplements (Paudyal et al., 2022), in addition to providing support in-home delivery cases (Government of Nepal, 2023a). The Government of Nepal has adopted the eight Antenatal care (ANC) contact protocol recommended by WHO, with the first visit occurring at 12 weeks, the second at 16 weeks, the third from 20–24 weeks, the fourth within 28 weeks, the fifth in 32 weeks, the sixth in 34 weeks, the seventh in 36 weeks, and the eighth between 38–40 weeks (Ministry of Health and Population [Nepal] et al., 2023). The pregnant population that is depressed was calculated using the population projection from the Institute for Health Metrics and Evaluation (IHME) for 2023 (Institute for Health Metrics and Evaluation (IHME), 2020) and the crude birth rate projection from WHO estimates for 2023 (World Health Organization, 2023b) with the prevalence of depression in pregnant women.

Screening strategies

The screening strategies were chosen based on the WHO recommendation on postnatal care (World Health Organization, 2022). The following screening strategies were included:

EPDS: The EPDS is a ten-item self-administered tool developed to assist in identifying possible symptoms of depression in the postnatal period (Cox et al., 1987). It has shown adequate sensitivity and specificity to identify depressive symptoms in antenatal and postnatal periods. EPDS was validated in a Nepalese setting. A score of 12/13 was used to indicate a positive screen (Bhusal et al., 2016).

PHQ-9: PHQ-9 is a depression module with a nine-item self-administered tool developed to assist in identifying possible symptoms of depression (Spitzer et al., 1999). PHQ-9 was a validated measure in the Nepalese population; a score of 10 was used to indicate a positive screen (Kohrt et al., 2016).

Treatment strategy:

Usual care

With no current guidelines on PND treatment in Nepal, the usual care treatment of PND follows the mhGAP intervention guide developed by WHO (World Health Organization, 2016). The Intervention Guide outlines a comprehensive approach to managing priority Mental, Neurological, and Substance (MNS) conditions using algorithms designed for

clinical decision-making, which was developed for non-specialized health settings. The intervention guide treats depressed pregnant women as a distinct and special population, offering specific considerations to be taken into account in the decision-making process (World Health Organization, 2016).

Thinking Healthy Program peer delivered (THHP)

Thinking Healthy Program (THP) is a program adopted by the WHO for perinatal depression that encompasses low-intensity psychological intervention (World Health Organization, 2023a). THP peer delivered (THPP) is a modified THP delivery by trained peer counselors (Fuhr et al., 2019). The intervention delivered in the Indian setting was incorporated with FCHVs as the intervention delivery agent for the Nepalese setting for this study used in scenario analysis.

Time Horizon

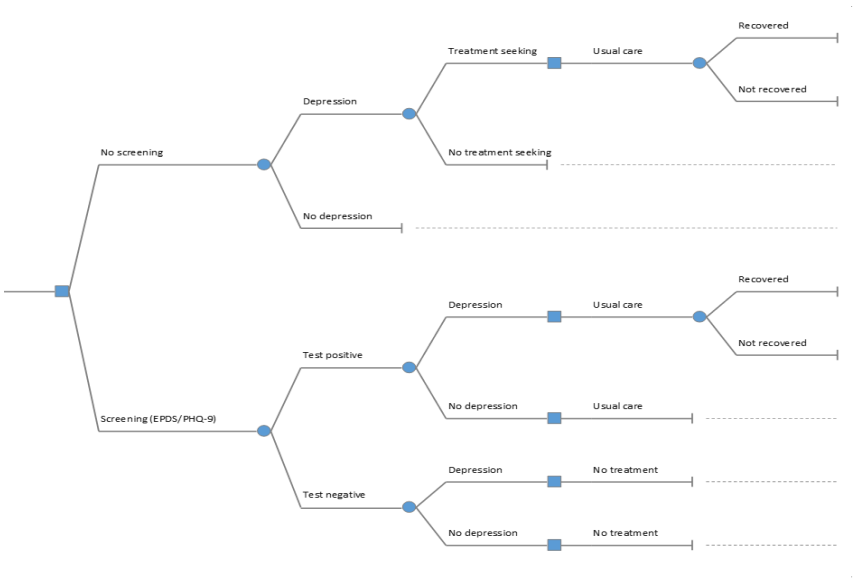
A systematic review of screening programs for common maternal mental health disorders among perinatal women showed varied timepoints for screening 23 to 23 to 32weeks of gestation and four to six weeks after birth was reported in different studies, with the end points 3months postpartum to 12 months postpartum (Waqas et al., 2022). More on screening time and intervals are discussed in part II (extended theory, guidelines for screening and treatment of PND). For this study, the time horizon, form the second or third trimester of pregnancy to six months post-childbirth, was chosen, a total of approximately nine months. A similar time horizon was used by (Heslin et al., 2022) to assess the cost-effectiveness of screening tools for early pregnancy, and nine months represents the duration of the THHP study outcome that is analyzed as a scenario analysis.

Model structure

We developed a decision tree model in Microsoft Excel to evaluate the cost-effectiveness of the screening strategies. This model incorporates the screening along with usual care **Error! Reference source not found.**. At the start of the model, women are either screened using EPDS or PHQ-9 or not screened. All women who screen positive get the usual care, and they respond to the treatment. All women who screen negative do not receive any treatment. We assumed that the number of cases detected in the do-nothing arm as a proportion of women seeking treatment for general mental health conditions get usual care.

The model is analyzed at three levels. Part I was the impact of screening tools (EPDS, PHQ-9) in terms of cases detected in the system. Part II is where the outcome is the cost per case detected and treated—and part III is where the outcome is the cost per QALY.

Figure 1 *Detection and treatment pathway*



Model parameters

Clinical input parameters

Probabilities associated with the sensitivity and specificity of the screening tools and treatment pathway modeled are reported in ***Table 1***. The data for diagnostic properties of screening tools were taken from primary studies in a Nepalese setting. A literature review was performed to check for the available literature in the context of PND in Nepal, and the search query is provided in Appendix F. Data on sensitivity and specificity for the Nepalese version of EPDS were taken from a hospital-based cross-sectional study in Kathmandu, Nepal, which aimed to validate the EPDS as a screening tool (see paper for full details) (Bhusal et al., 2016). Data on sensitivity and specificity for the Nepalese version of PHQ-9 were taken from a cross-sectional study, randomly selecting participants in Chitwan, Nepal, which aimed to validate the PHQ-9 (Kohrt et al., 2016). The International Classification of Disease tenth revision (ICD-10) and validated Nepali depression module of the Composite International Diagnostic Interview (CIDI) was used as the gold standard for validating EPDS and PHQ-9 respectively (Bhusal et al., 2016; Kohrt et al., 2016). Different diagnostic parameters from a meta-analysis that was a part of the National Institute for Health and Care Excellence (NICE) guidelines (National Collaborating Centre for Mental Health (UK), 2014) for the same diagnostic threshold in Nepalese settings were explored in a scenario analysis.

Leverton et al. demonstrated the health visitor's ability to detect depression during the postnatal period, which is reported to have a sensitivity of 8% and specificity of 98%

(Leverton & Elliott, 2000). However, this was deemed irrelevant in our setting as the FCHVs had no formal training to assess and detect depression. In the absence of a screening tool, women had to choose to seek care for their depression in order to receive treatment. Treatment-seeking behavior on mental health problems in the female population in the context of Nepal was used as a proxy for the treatment pathway in no screening arm, which was 7% (Ministry of Health and Population [Nepal] et al., 2023). This treatment-seeking behavior was deemed appropriate for our model as no precise mechanism or training is provided to FCHVs to detect perinatal depression.

The lack of primary studies assessing the impact of usual care in managing PND led to the use of secondary sources. A brief literature review on the cost-effectiveness studies of PND globally was conducted as the data from Nepal was not enough to feed the model, and the studies were assessed with title and abstract for further reading. The search strategy is provided in Appendix F. In terms of treatment, the absolute risk for no improvement with standard care was 0.61 (National Collaborating Centre for Mental Health (UK), 2014), which was similar to the absolute risk for no improvement with standard care 0.59 as in the THHP intervention study in Indian setting (Fuhr et al., 2019). The NICE estimate on treatment response was used in the base model. We assumed that anyone who screened positive (whether true positive or false positive) received treatment. For scenario analysis, THHP intervention was provided in screening alternatives in addition to usual care (Fuhr et al., 2019). Data on response to treatment for THHP was taken from the same THHP study (Fuhr et al., 2019). The probability of recovery in THHP plus usual care was 0.53, and the probability of recovery with usual care alone was 0.41. More information on the THHP intervention is described in the extended background section. For scenario analysis with spontaneous recovery, the absolute risk for no improvement without any care, 0.67, was taken from (Dennis et al., 2009), and it was in line with the NICE recommendation of 0.67 absolute risk for no improvement without any care (National Collaborating Centre for Mental Health (UK), 2014). The beta distribution was chosen for all the model parameters in Table 1. The distribution parameters were directly provided or estimated from the primary source reported, except for the spontaneous recovery, where model parameters were calculated from the 95% CI (0.242 – 0.425) taken from (Dennis et al., 2009). The model pathway for the spontaneous recovery model is presented in Appendix A.

Table 1 *Model parameters for screening accuracy and treatment pathway*

Parameters	Base case probabilities	Source	Distribution	Distribution parameters	
				Alpha (α)	Beta (β)
Prevalence	24.8	(Chalise et al., 2022)	Beta	62	188
Screening Pathway					
EPDS					
Sensitivity	0.92	(Bhusal et al., 2016)	Beta	46	4
Specificity	0.96	(Bhusal et al., 2016)	Beta	283	13
PHQ-9					
Sensitivity	0.94	(Kohrt et al., 2016)	Beta	16	1
Specificity	0.8	(Kohrt et al., 2016)	Beta	86	22
Treatment seeking	0.07	(Nepal Health Research Council, 2020)	Beta	420	5572
Treatment pathway					
Response to treatment					
Usual care	0.39	(National Collaborating Centre for Mental Health (UK), 2014)	Beta	793	808
THHP intervention					
THHP plus usual care	0.53	(Fuhr et al., 2019)	Beta	61	55
Usual care only	0.41	(Fuhr et al., 2019)	Beta	49	71
Spontaneous recovery	0.33	(Dennis et al., 2009)	Beta	33	66.9
NICE estimates					
EPDS					
Sensitivity	0.61	(National Collaborating Centre for Mental Health (UK), 2014)	Beta	45	29
Specificity	0.94	(National Collaborating Centre for Mental Health (UK), 2014)	Beta	494	31
PHQ-9					
Sensitivity	0.75	(National Collaborating Centre for Mental Health (UK), 2014)	Beta	53	18
Specificity	0.88	(National Collaborating Centre for Mental Health (UK), 2014)	Beta	582	82

Note. EPDS = Edinburgh Postnatal Depression Score; PHQ-9 = Patient Health Questionnaire 9 items; THHP = Peer-delivered Thinking Healthy Program; NICE = National Institute for Health and Care Excellence.

Outcomes

Three parts of the model have different outcomes: part I (cost per case detected), part II (cost per case detected and treated), and part III (cost per QALYs). Outcomes for part III are described in Table 2. Quality-adjusted life years (QALYs) measure health outcomes depending on period and weight, ranging from 0 to 1. The weight is health-related quality of life during that period, where one corresponds to optimal health, and zero corresponds to being dead. The QALYs used in this study are taken from a cost-effectiveness analysis of screening tools for identifying depression in early pregnancy in the United Kingdom (Heslin et al., 2022) due to a lack of primary sources and lower-middle-income countries sources on the antenatal utility. Heslin et al. QALYs are based on the utility values from a prospective diagnostic accuracy cohort study (BaBY PaNDA) that estimates the utility values of

depression for the prenatal period and postnatal period where the treatment pathway is the standard treatment pathway based on the (NICE) guidelines (Littlewood et al., 2018). The QALYs are from the antenatal to the postnatal period (9 months) in terms of depressed versus not, i.e., depressed in the antenatal period to depressed in the postnatal period, depressed in antenatal period to not depressed postnatal period, and not depressed in antenatal period to not depressed postnatal period (Heslin et al., 2022). The study only provided utility in the case treatment was offered; the QALYs for depressed to depressed in case of not receiving the treatment were estimated by keeping the utility at the baseline and utility at the end the same. For the spontaneous recovery model, additional QALY for the case of self-recovery was needed; the QALY for depressed to not depressed in case of no treatment was assumed to have a 5% reduction in QALYs relative to depressed to not depressed with treatment. For the scenario analysis where the baseline adjusted utility method for QALY estimation was used, the utility gained (start and end point of utility values) accounted for the duration stayed in a particular utility state was used to estimate the QALYs. The QALY values are reported in ***Table 2***. The beta distribution was deemed appropriate for propagating uncertainty. For the beta distribution, standard error was assumed to be 30% of mean and the distribution parameters were estimated using 95% CI.

Table 2 *Model parameters for the outcome - QALYs*

Parameter	Values	Source	Distribution	Standard Error	95% CI
QALYs (9months, treatment)					
Depressed to non-depressed	0.6553	(Heslin et al., 2022)	Beta	0.1966	0.270-1.00
Depressed to depressed	0.5991	(Heslin et al., 2022)	Beta	0.1797	0.247-0.951
Non-depressed to non-depressed	0.7422	(Heslin et al., 2022)	Beta	0.2227	0.306-1.179
QALYs (9months, no treatment)					
Depressed to depressed[a]	0.5606	(Heslin et al., 2022)	Beta	0.1681	0.231-0.889
Depressed to non-depressed[b]	0.6225	(Heslin et al., 2022)	Beta	0.1867	0.256-0.988
QALYs (9 months, treatment)[c]					
Depressed to non-depressed	0.0946	(Heslin et al., 2022)	Beta	0.0284	0.039-0.150
Depressed to depressed	0.0384	(Heslin et al., 2022)	Beta	0.0115	0.015-0.060

Note. QALYs = Quality Adjusted Life Years; CI = Confidence Interval

[a] estimated from the utility values from (Heslin et al., 2022), keeping utility at the baseline and utility at the end the same. [b] estimated from QALYs from (Heslin et al., 2022), the QALY for depressed to not depressed in case of no treatment was assumed to have a 5% reduction in QALYs relative to depressed to not depressed with treatment. [c] estimated from baseline adjusted values of utility (The baseline adjustment of utility gain method was used to estimate QALYs).

Resource use and unit costs

The economic evaluation used a healthcare perspective. The costs associated with each screening and treatment are presented in ***Table 3***. Data on the resource use for screening was estimated using the time to screen and per minute tariff of the FCHVs. Different screening time for EPDS was reported in different studies: 3.54 minutes (Heslin et al., 2022), 5 minutes (Paulden et al., 2009), and 15 minutes (A. Wilkinson et al., 2017). WHO recommendation on time for screening was, for EPDS, 10 minutes for face-to-face, 5 minutes for self-administered, and 3-10 minutes for PHQ, depending on the version of PHQ (*WHO recommendations on maternal and newborn care for a positive postnatal experience*, 2022). As EPDS is a 10-item questionnaire and PHQ-9 has nine, the same time to screen was used for both of the screening techniques, 10 minutes. The per-minute tariff of FCHVs was estimated using the minimum salary in Nepal (Government of Nepal, 2023b), as the FCHVs are not paid a salary per the government rule but are incentivized with other perks. In a scenario analysis where auxiliary nurse midwives (ANM) are considered the agents to screen, the cost of screening was estimated using the per-minute tariff of ANM (Loksewajob, 2023).

The cost of usual care was taken from a study estimating the annual cost of the World Health Organization's mhGAP treatment component per average case of disorder (Chisholm et al., 2016). This was taken as a proxy for the usual care due to a lack of studies estimating the usual care cost for the management of PND in pregnant women.

For the THHP intervention model, the cost estimate for the THHP delivery was based on the duration of intervention from the THHP India study. The intervention delivery tariff for Sakhis in the Indian setting for the RCT (Fuhr et al., 2019) was even higher than the general practioner salary in the Nepalese context (Loksewajob, 2023), so the cost was estimated using per minute tariff for FCHVs (Government of Nepal, 2023b).

For true positives, the total treatment cost was assigned. For false positives, according to NICE guidelines, they would receive the same treatment as true positives, but they would stop treatment as their false positive status is recognized and would consume 20% of resources (National Collaborating Centre for Mental Health (UK), 2014), which was also adapted in (Heslin et al., 2022); a similar approach was adopted in this study.

The cost for usual care was from 2016, which was adjusted to July 2023 using the consumer price index (Trading Economics, 2023). For the currency conversion, 1 USD = 132.76 Nrs date: 13th September was used for this study. The gamma distribution was used to propagate

all the cost parameters, and the standard error was assumed to be 30% of the mean for all the parameters.

Table 3 *Model parameters for the cost of screening and treatment*

Parameter	Cost (USD)	Distribution	Standard error	Source	Note
Screening (FCHVs)					
EPDS	0.147	Gamma	0.0441	(Government of Nepal, 2023b)	Based on 10 minutes to screen, FCHV costs 0.0104 USD per minute, based on the minimum salary.
PHQ	0.147	Gamma	0.0441	(Government of Nepal, 2023b)	Based on 10 minutes to screen, FCHV costs 0.0104 USD per minute, based on the minimum salary.
Screening (ANM)	0.215	Gamma	0.0647	(Loksewajob, 2023)	Based on 10 minutes to screen for EPDS and PHQ-9, ANM costs 0.0215 USD per minute, based on the ANM salary.
No screening					
GP Diagnosis	0.954	Gamma	0.2862	(Loksewajob, 2023)	Based on a 30-minute examination by G.P., 0.0318 USD per minute, based on salary.
Treatment					
Usual Care	37.6	Gamma	11.28	(Chisholm et al., 2016)	33.81 USD from treatment cost per average case of disorder per annum from 2016 adjusted to inflation with CPI index to 2023 adjusted to 9 month period
THHP	5.17	Gamma	1.55	(Fuhr et al., 2019)	Based on 14 sessions *(37.5 minutes) at 0.3693 USD per session for FCHVs (Fuhr et al., 2019),.

Note. USD = US Dollar; EPDS = Edinburgh Postnatal Depression Score; PHQ-9 = Patient Health Questionnaire 9 items; FCHVs = Female Community Health Volunteer; ANM = Auxiliary Nurse Midwives; GP = General Practioner; THHP = Peer-delivered Thinking Healthy Program; CPI = Consumer Price Index.

Assumptions

The following assumptions were made:

- All screening tools are used with all women at the antenatal visits by FCHVs;
- Women screened positive by the FCHVs are not already receiving the treatment, and therefore, all the women screened positive will be referred to THHP plus usual care;
- Those who screen negative and are true negative will not be depressed in the later visits;
- In the no-screening arm, those with depression getting the treatment depend on the treatment-seeking behavior and get the usual care;
- No spontaneous recovery in no screening arm as well as screening arm;
- In the false negative arm, there is no later identification.

Model outputs

Results are presented in different ways for different parts of the model. For part I of the model, the result is presented in terms of incremental cost per case detected in the system for screening tools relative to no screening. For part II of the model, the result is presented in terms of incremental cost per case detected and treated in the system. Results in part III are presented in four ways: ICERs, scatterplots, cost-effectiveness acceptability curves (CEAC), cost-effectiveness acceptability frontier (CEAF), and population Expected Value of Perfect Information (pEVPI). The cost-effectiveness of the model outcome with QALYs is established with a WTP threshold per QALY. 1* GDP per capita was chosen as the WTP threshold based on the WHO recommendation of 1*GDP per capita to 3* GDP per capita WTP threshold, where the WTP threshold is not established. Nepal's GDP per capita for 2023 was 1371.971 USD (CEIC, 2023). ICERs are calculated as the incremental cost (difference between cost in two groups) divided by the incremental effect (difference in outcome between two groups), providing incremental cost per health outcome unit.

The probabilistic model was only implemented in part II and part III. Probabilistic models give us distribution over incremental cost, incremental effect, and joint cost-effect distribution. CEAC and CEAF were produced using a net-benefit approach based on the simulations of cost-outcome data from probabilistic sensitivity analysis PSA. CEAC provides the probability that an intervention is cost-effective according to increasing threshold values. CEAF provides the probability that optimal intervention is cost-effective according to increasing threshold values. The highest expected net benefit determines the optimal option, whereas CEAC represents the proportion of iterations in which each option had the highest net benefit. Two-way CEAC and CEAF were analyzed for EPDS versus no screening and PHQ-9 versus no screening, respectively. In pEVPI was calculated for the base model. The expected Value of Perfect Information (EVPI) provides the maximum sum worthwhile for conducting further research to eliminate all uncertainties in the model. pEVPI is estimated from EVPI, considering the beneficiary population's size and the time horizon where the information generated will be useful.

As no WTP threshold was available for part II with model outcome incremental cost per incremental treated, the model is not further analyzed with CEAC and CEAF; only a scatterplot is discussed. For part III, with QALY as an outcome measure, the incremental cost per QALY, scatterplot, CEAC, CEAF, and pEVPI are presented.

Sensitivity analysis

In decision models for cost-effectiveness analysis, the integrity of the economic model depends on assumptions made and the validity of model input parameters. Sensitivity analysis helps analyze the impact of changes in model parameters and assumptions. Sensitivity analysis can be broadly classified into deterministic sensitivity analysis (to assess the impact of uncertainty around the value of individual parameters or uncertainty around model structure) and probabilistic sensitivity analysis (to assess the impact of joint uncertainty of multiple parameters simultaneously). The probabilistic method best estimates mean costs and outcomes in a non-linear decision model. PSA assigns the distribution around model parameters to represent uncertainty and propagate the uncertainty using simulations. In the current study, 5,000 simulations were made, resulting in a joint distribution of cost and health outcomes. The high number of simulation was chosen as the test model convergence for the ICERs were not performed; relatively high simulations, such as 5000, are deemed sufficient.

A range of scenario analyses with probabilistic sensitivity analyses were conducted for parts II and III of the model:

- Scenario 1: THHP intervention provided to the screening alternatives,
- Scenario 2: ANM is used for screening,
- Scenario 3: Use of QALYs estimated from controlling for baseline utility,
- Scenario 4: Sensitivity and specificity for screening alternatives from NICE guidelines,
- Scenario 5: Spontaneous recovery model.

Results

The results of base case analysis are presented in three parts following the model structure.

Part I:

The model outcome in terms of cases detected is presented in ***Table 4***. 2478 cases are detected in the health system if there is no screening in place, with a cost of USD 2364. With EPDS, 32568 cases were detected, and 32276 cases were detected with PHQ-9 with the same cost of USD 14845. The cost per incremental case detected for EPDS and PHQ-9 relative to no screening was USD 0.414 and USD 0.405, respectively.

Table 4 *Total costs and number of cases detected for each screening approach*

	Total detected	Total Costs (USD)	Incremental detected	Incremental costs	ICER
No screening	2478	2364			
EPDS	32568	14845	30090	12481	0.414[a]
PHQ-9	33276	14845	30798	12481	0.405[a]

Note. EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio.

[a] ICER relative to no screening.

Part II:

Model outcomes for cases detected and treated for screening with usual care are presented in ***Table 5***. The total number of cases treated in the no screening arm was 966, with a total cost of 95532 USD, whereas in the EPDS arm, the total treated cases were 12701, and in the PHQ-9 arm, 12977, with a mean cost of 1.27 million USD and 1.42 million USD to the system respectively. This resulted in an incremental cost per treated of USD 100.49 for EPDS screening and USD 110.88 for PHQ-9, relative to no screening.

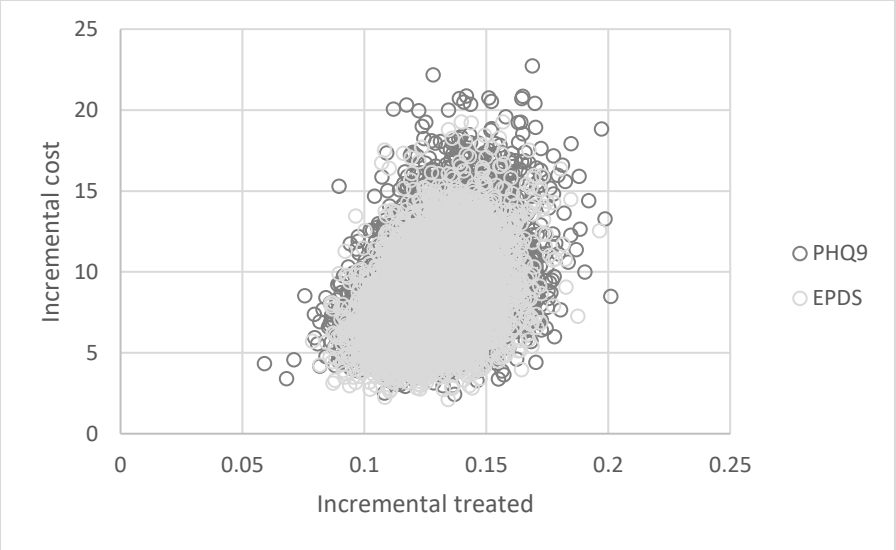
Table 5 *Mean costs and treated for each screening approach*

	Total treated	Total Costs (USD)	Incremental treated	Incremental costs	ICER
No screening	966	95532			
EPDS	12701	1274864	11735	1179331	100.49
PHQ-9	12977	1427403	12011	1331870	110.88

Note. EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio.

[a] ICER relative to no screening.

Figure 2 *Scatterplot for incremental treated and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*



The scatter plot of the incremental cost and incremental cases treated for the EPDS screening and PHQ-9 screening relative to no screening is plotted in ***Figure 2***. The scatter plot shows that the EPDS and PHQ-9 simulations overlap, with PHQ-9 with some simulations on the right side with higher incremental costs and higher incremental treated.

Scenario analysis

All scenario analysis results are presented in ***Table 6***The model with THHP intervention delivered along with usual care in the screening alternatives showed the highest reduction in the ICER per case treated for EPDS and PHQ-9, USD 83.26 per case treated and 91.82 per case treated relative to no screening, respectively. In other models, using ANM to screen and the sensitivity and specificity values from NICE guidelines, the ICERs were close to the baseline model, USD 101.85 per case treated for EPDS and USD 112.21 per case treated for PHQ-9, and USD 104.58 per case treated for EPDS and USD 108.05 per case treated for PHQ-9, relative to no screening respectively. The model with baseline-adjusted utilities had no difference from the baseline model as the model impacts the utility values. But in the scenario where spontaneous recovery was allowed, ICER per case treated for EPDS and PHQ-9 was USD 653.22 per case treated and USD 720.75 per case treated relative to no screening, respectively. The scatterplots for the scenario analysis are reported in Appendix C.

Table 6 *Mean costs and treated for each screening approach for different scenarios*

	Mean treated	Mean cost (USD)	Incremental treated	Incremental cost	ICER
Scenario 1					
No screening	0.0071	0.67			
EPDS	0.1209	10.15	0.1138	9.47	83.26
PHQ-9	0.1235	11.36	0.1164	10.69	91.82
Scenario 2					
No screening	0.0067	0.67			
EPDS	0.0889	9.04	0.0822	8.37	101.85
PHQ-9	0.0909	10.11	0.0841	9.44	112.21
Scenario 3					
No screening	0.0067	0.67			
EPDS	0.0889	8.93	0.0822	8.26	100.49
PHQ-9	0.0909	9.99	0.0841	9.33	110.88
Scenario 4					
No screening	0.0067	0.67			
EPDS	0.0589	6.13	0.0522	5.46	104.58
PHQ-9	0.0725	7.77	0.658	7.11	108.05
Scenario 5					
No screening	0.0828	0.67			
EPDS	0.0955	8.93	0.0126	8.26	653.22
PHQ-9	0.0958	9.99	0.0129	9.33	720.75

Note. EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio; Scenario 1 = Peer delivered Thinking Healthy Program delivered to screening arm in addition to usual care; Scenario 2 = Auxiliary Nurse Midwives are used as the screening agent; Scenario 3 = Baseline utility adjustment method used to estimate QALYs and utility gains are provided only to the treatment arm; Scenario 4 = Sensitivity and Specificity parameters taken from meta-analysis from National Institute for Health and Care Excellence guidelines; Scenario 5 = No spontaneous recovery relaxed and allowed on the in the no-treatment-seeking population of the no-screening arm and the false negative arm of screening approaches.

Part III:

The results for part III are presented in. In the model with usual care in all alternatives, the mean QALY per person was highest for PHQ-9 at 0.7112, followed by EPDS at 0.7109 and no screening at 0.698. Total cost per person was highest for PHQ-9, USD 9.99, followed by EPDS, USD 8.93, and do nothing, USD 0.67. The ICER was USD 648.7 per QALY for EPDS compared to no screening and USD 715.76 per QALY for PHQ-9 compared to no screening. Compared to a threshold of 1*GDP per capita (US$1371), both EPDS and PHQ-9 are potentially highly cost-effective.

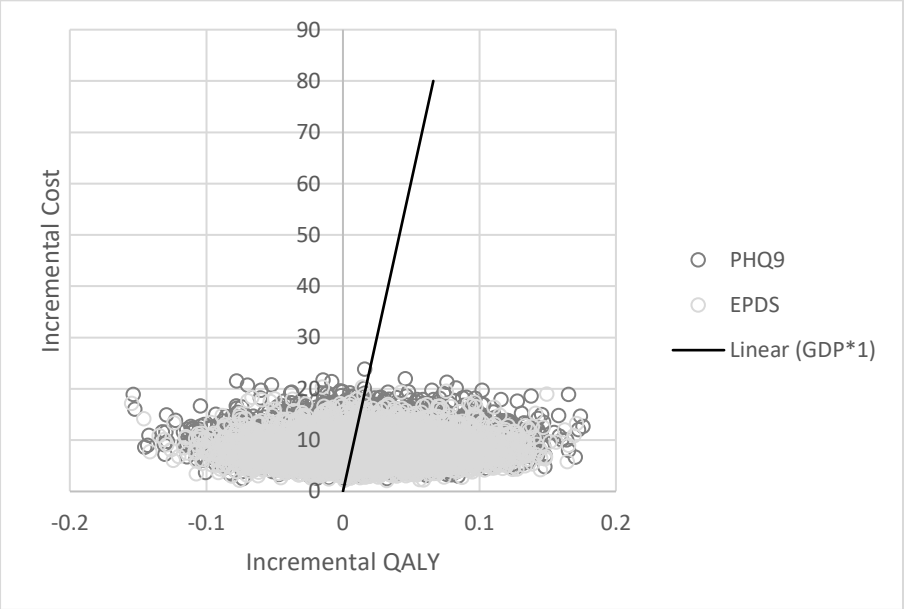
Table 7 Mean costs and QALYs for each screening approach

	Mean QALY	Mean cost	Incremental QALYs	Incremental costs	ICER
No screening	0.698	0.67			
EPDS	0.7109	8.93	0.0127	8.26	648.7
PHQ-9	0.7112	9.99	0.0130	9.33	715.76

Note. QALY = Quality Adjusted Life Year; EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio.

The scatter plot of the incremental cost and incremental QALYs for the EPDS screening and PHQ-9 screening relative to no screening is plotted in ***Figure 3***. The scatter plot shows that the EPDS and PHQ-9 simulations overlap, PHQ-9 with some simulations on the outer side with higher incremental costs and higher incremental treated. As the scatter plot has 5000 simulations, the inference of having more or less simulated ICERs on the right side to the 1*GDP per capita WTP threshold line could not be made as it was not visibly distinguishable.

Figure 3 *Scatterplot for incremental Quality Adjusted Life Years (QALY) and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*



The probabilistic ICERs for EPDS and PHQ-9 relative to no screening were 552.15 and 612.29, which is reported in Appendix B. The CEAC for EPDS versus no-screening and PHQ-9 versus no-screening were evaluated and are presented in ***Figure 4*** and ***Figure 5***. The trajectory of EPDS and PHQ-9 probability to be cost-effective along different WTP thresholds was similar. The CEAC indicates that at a willingness to pay at USD 0 per QALY, no screening has the highest probability of being cost-effective. With the increase in willingness to pay, the probability of screening alternatives being cost-effective increases. At

the 1*GDP per capita WTP threshold, the probability of being cost-effective was 0.57 and 0.43 for EPDS and no screening, respectively, when EPDS versus no screening is considered. When PHQ-9 was considered, CEAC for 1*GDP per capita WTP threshold was 0.56 and 0.44 for PHQ-9 and no screening, respectively. The results from the CEAF frontier are shown in ***Figure 6*** and ***Figure 7***; in the models EPDS and PHQ-9 relative to no screening, the decision switch is around WTP USD 750, which is close to the probabilistic ICER for EPDS and PHQ-9 relative to no screening 552.15 and 612.29 respectively. At the 1*GDP per capita WTP threshold, EPDS and PHQ-9 were the optimal choice of intervention over no screening. The pEVPI for the two-way model is presented in ***Figure 6*** and ***Figure 7***, EPDS versus no screening and PHQ-9 versus no screening, was 2.99 million USD and 2.86 million USD, respectively, at 1*GDP per capita WTP threshold presented in.

Figure 4 *Cost-effectiveness acceptability curve for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

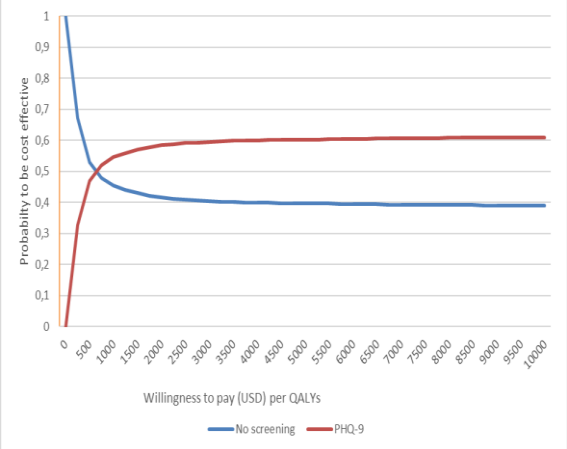


Figure 5 *Cost-effectiveness acceptability curve for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*

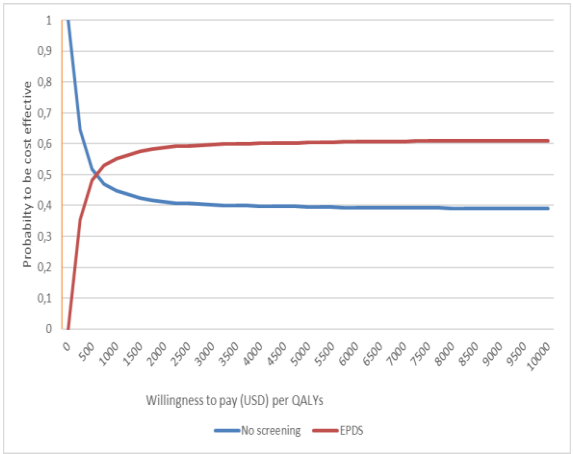


Figure 6 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening and population expected value of perfect information (pEVPI)*

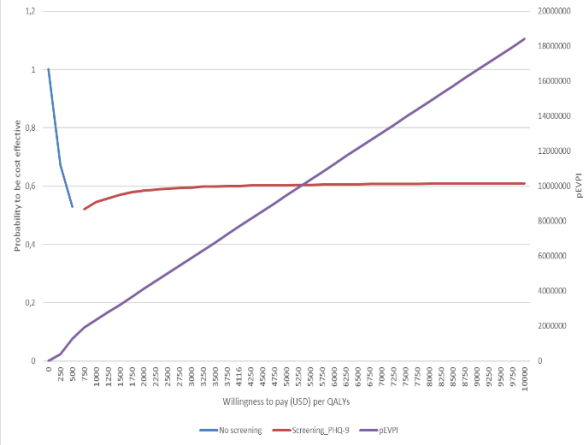


Figure 7 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening and population expected value of perfect information (pEVPI)*

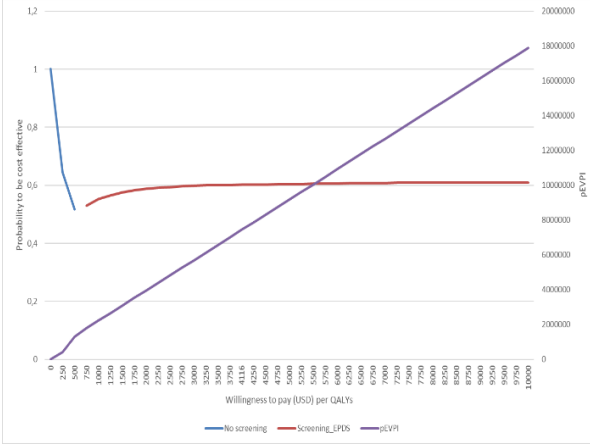


Table 8 *Mean costs and QALYs for each screening approach for different scenarios*

	Mean QALY	Mean Costs (USD)	Incremental QALYs	Incremental costs	ICER
Scenario 1					
No screening	0.6982	0.67			
EPDS	0.7127	10.15	0.0145	9.48	652.97
PHQ-9	0.7130	11.36	0.0148	10.69	719.93
Scenario 2					
No screening	0.698	0.67			
EPDS	0.7109	9.04	0.0127	8.37	657.46
PHQ-9	0.7112	10.11	0.0130	9.44	724.32
Scenario 3					
No screening	0.001	0.67			
EPDS	0.0137	8.93	0.0127	8.26	649.77
PHQ-9	0.014	9.99	0.0130	9.33	716.95
Scenario 4					
No screening	0.698	0.67			
EPDS	0.7063	6.21	0.008	5.46	675.03
PHQ-9	0.7084	7.78	0.0101	7.11	697.48
Scenario 5					
No screening	0.703	0.67			
EPDS	0.7113	8.93	0.0084	8.26	980.33
PHQ-9	0.7115	9.99	0.0086	9.33	1081.68

Note. QALY = Quality Adjusted Life Years; EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio; Scenario 1 = Peer delivered Thinking Healthy Program delivered to screening arm in addition to usual care; Scenario 2 = Auxiliary Nurse Midwives are used as the screening agent; Scenario 3 = Baseline utility adjustment method used to estimate QALYs and utility gains are provided only to the treatment arm; Scenario 4 = Sensitivity and Specificity parameters taken from meta-analysis from National Institute for Health and Care Excellence guidelines; Scenario 5 = No spontaneous recovery relaxed and allowed on the in the no-treatment-seeking population of the no-screening arm and the false negative arm of screening approaches.

Scenario analysis

The ICERs per QALYs for the scenario are reported in ***Table 8***. The model with THHP intervention delivered along with usual care in the screening alternatives showed a slight increase in the ICER per QLAYs for EPDS and PHQ-9, USD 652.97 per QALYs, and USD 719.93 per QALYs, relative to no screening, respectively. In another model, using ANM to screen the ICERs was also higher than the baseline model, USD 657.46 per QALYs for EPDS and USD 724.32 per QALYs for PHQ-9 relative to no screening. In the model where the sensitivity and specificity values from NICE guidelines were taken, the ICERs were USD 675.03 per QALYs for EPDS and USD 697.48 per QALYs for PHQ-9 relative to no screening. In the model with baseline adjusted utilities, QALY gains were almost identical to the baseline model, resulting in ICERs, USD 649.77 per QALYs for EPDS and USD 716.95 per QALYs for PHQ-9 relative to no screening. The probabilistic ICERs for all scenarios are presented in Appendix B, and all the probabilistic ICERs were in the same conclusion as the deterministic ICERs reported in ***Table 8***; in all scenarios, both screening approaches were cost-effective under 1*GDP per capita threshold.

Scatterplots, CEAC, and CEAF for all different scenario models are presented in Appendix C. In the model with THHP intervention delivered along with usual care in the screening alternatives, the optimal choice was EPDS screening with a 0.55 probability of being cost-effective at 1* GDP per capita WTP threshold for the model with no screening and EPDS and the optimal choice was PHQ-9 screening with a 0.56 probability of being cost-effective at 1* GDP per capita WTP threshold for the model with no screening and PHQ-9. In the models with ANM as a screening agent and baseline adjusted utilities, EPDS and PHQ-9 was the optimal choice with 0.57, 0.56, and 0.99, 0.98 probability of being cost-effective at 1*GDP per capita threshold for the model with no screening and EPDS and with no screening and PHQ-9, respectively. In the model with NICE guidelines test parameters for screening tests, both EPDS and PHQ-9 were the optimal choice of intervention with a 0.57 and 0.56 probability of being cost-effective 1* GDP per capita WTP threshold for the model with no screening and EPDS and with no screening and PHQ-9, respectively. In the model with spontaneous recovery, EPDS and PHQ-9 were the optimal choice of intervention with a 0.54 and 0.53 probability of being cost-effective 1* GDP per capita WTP threshold for the model with no screening and EPDS and with no screening and PHQ-9, respectively.

Discussion

Main findings

This study compared two screening approaches for the detection and treatment of depression against a no-screening alternative in pregnant women at their antenatal visit with FCHVs. The study presented three outcomes: cost per case detected, cost per case treated, and cost per QALY gained.

The incremental cost per case detected for screening (EPDS, PHQ-9) compared to no screening was substantially low, even less than one USD, which indicates that implementing a screening tool to identify PND cases is potentially beneficial. Screening alone and detecting cases of PND may not be useful if no treatment option is available. Therefore, the incremental cost per case detected and treated was estimated. Both screening options (EPDS, PHQ-9) led to higher cases detected and treated compared to no screening. Furthermore, we estimated the cost per QALY gained, and both screening alternatives were cost-effective compared to no screening, falling below the willingness to pay threshold of 1*GDP per capita.

Both the treated (part II) and QALY (part III) models were explored under different scenarios. The implementation of THHP intervention along with the screening had slightly higher ICER than the baseline model but well within 1*GDP per capita WTP threshold in part III, suggesting that the introduction of such CBT-based intervention that could be delivered by FCHVs can be implemented for PND. When we used ANM as health personnel delivering the intervention, the intervention costs were higher, resulting higher ICER but the ICER was still below the WTP threshold. In a scenario where the values for screening alternatives were based on the meta-analysis from NICE, the ICER for EPDS increased but was within the 1*GDP per capita WTP threshold. In contrast, the ICER for PHQ-9 decreased and was within the 1*GDP per capita WTP threshold. In our base model, the sensitivity and specificity of EPDS and PHQ-9 were taken from each primary study, respectively. The test parameters were relatively better for the primary studies for EPDS (both higher sensitivity and specificity) than the estimates in the meta-analysis published in the NICE guidelines, so that explains the increase in the EPDS ICER, whereas for PHQ-9 baseline model had higher sensitivity and lower specificity compared to NICE guideline, which resulted in slightly lower ICER than the baseline model.

Manca et al. provide the framework for the utility estimation techniques and controlling for the baseline utility (Manca et al., 2005). The utility gain from an intervention is commonly calculated in two ways, one taking the average utility before and after intervention adjusted for the duration stayed in a certain utility state, and another taking only the utility gains before and after intervention adjusted for the duration stayed in a certain utility state. The baseline model uses the former average method for utility estimation as taken from (Heslin et al., 2022); as a scenario analysis, the latter method controlling for the baseline utility was explored. It resulted in very similar ICER values for the QALY model; as for the treated model, that does not change the outcome of the model. In the scenario where spontaneous recovery was allowed in the model for those who do not receive treatment for depression, the ICER values of both screening alternatives were increased relative to no screening but were within the 1*GDP per capita WTP threshold. Probabilistic sensitivity analysis was implemented for all the QALY models, with CEAC and CEAF for the different scenarios discussed above. The optimal choice of intervention was still screening tools (EPDS/PHQ-9), as in the baseline model at WTP 1*GDP per capita. The nature of the pEVPI curve is explained in Appendix E.

Comparison of results

This is the first study evaluating the cost-effectiveness of screening and treatment of PND in the Nepalese context. Most studies of screening and treatment for PND in the literature are conducted in higher-income countries and show mixed results on cost-effectiveness. U.K.-based studies have shown contrasting evidence on the screening for PND. A study to assess the cost-effectiveness of screening tools for identifying depression in early pregnancy found the Whooley, EPDS, and EPDS followed by Whooley had a similar probability of being cost-effective around 30% than no-screen option 20% at WTP thresholds £20,000 - £30,000 per QALY gained (Heslin et al., 2022). The analysis lacked CEAF, which would have helped to detect the optimal choice among the screening options. Whereas in the postnatal setting, Plauden 2009, evaluates the cost-effectiveness of routine screening for postnatal depression in primary care. A decision model was implemented to evaluate the cost-effectiveness of routine application of either postnatal or general depression questionnaires; routine screening was not cost-effective compared with routine care only. The ICER was highly impacted by the cost of managing the incorrectly identified depression, as shown in the sensitivity analysis (Paulden et al., 2009). A cost-effectiveness study in the Canadian context, taking a public payer perspective with a two-year horizon, showed that screening for depression with EPDS

was cost-effective compared to the standard of care(Premji et al., 2021). Similarly, a modeled physician screening for and treating postpartum depression and psychosis in partnership with a psychiatrist in a U.S. Medicaid payer perspective for the time horizon of two years found the ICER for screening and treatment was a cost-effective intervention (Andra Wilkinson et al., 2017). More studies were in the postpartum period rather than the pre-pregnancy period.

The studies above had only usual care or current practice settings to assess the impact of screening alone, whereas there is an abundant amount of research on the intervention strategies to treat PND, which leads us to explore the combination of these intervention strategies with the screening tools. In our study, as in scenario 1, we couple the THHP intervention with the usual care setting to assess the impact of having a fairly simple intervention that could be delivered by the health personnel in place. There are several RCTs on the intervention alternative that show the treatment intervention is cost-effective (Henderson et al., 2019; Lee et al., 2016a; Ride et al., 2016), but most of them do not use QALYs as outcome measures (Dukhovny et al., 2013; Grote et al., 2017; Petrou et al., 2006). A web-based CBT measure as an intervention in an RCT to treat PND with QALYs as an outcome measure found that it was a cost-effective alternative (Lee et al., 2016b), similar in another cluster RCT model from PONDER trial had intervention as cost-effective (Henderson et al., 2019). In both of these studies, the intervention was cost-saving. In what were we thinking of cluster RCT, the intervention arm had a higher cost, but the ICER for intervention was below the national guideline (Ride et al., 2016). Even with the QALY measure, the approach of cost-effectiveness analysis is often challenging in these studies to compare with our study, as these studies are RCT-based cost-effectiveness, and our study has a model-based cost-effectiveness that is coupled with screening tools. Scenario 1 could be viewed as an extension of the RCT-based THHP intervention that was a cost-saving and cost-effective intervention in RCT based setting in the Indian context when the RCT-based intervention translated into a population setting in the Nepalese context with the incorporation of screening tools to identify the depression cases leads to a different result due the impact of screening test properties. The results were not cost-saving, but cost-effective in both cost per cases treated and cost per QALY setting with 1*GDP per capita WTP. RCT assessing the cost-effectiveness of psychological intervention within services for depression delivered by primary care workers in Nepal showed that the psychological intervention was cost-effective under the 1*GDP per capita threshold (Aldridge et al., 2022). The study had lower cost estimates for standard care and standard care plus intervention for service delivery. However,

in terms of implementation cost, the cost of standard care and standard care plus intervention was higher. The study had total cost as the implementation cost, which led to higher values of the total cost for the RCT, leading to higher ICER values for the psychological intervention even though it was cost-effective under the 1*GDP per capita threshold (Aldridge et al., 2022). As it was conducted in a general population setting, the QALY's gains could not be directly compared with our pregnant women population, and the study only provided the overall QALY gains for the whole study group.

Strength and limitations

This is the first study to assess the impact of screening and treatment of PND in the Nepalese context. This study was modeled with screening and treatment pathways, with validated screening approaches to the Nepalese context. This study incorporated several scenarios to inform better the questions that might be relevant. This study explored the use of the Thinking Healthy program treatment developed by WHO, with some adaptation (THHP) to be delivered by the non-specialist health worker. The clinical outcome of the THHP can be viewed as a limitation as well as a strength as there is no primary study in the Nepalese setting, but a study setting in India, a close neighbor with a similar setting, can be of more value than the estimates from the developed countries. The criteria to define the recovery had been chosen rather conservatively, defined as a PHQ-9 score of less than five at three months and six months by the authors in the India study (Fuhr et al., 2019). This study also highlighted the methodological issue with the QALY estimation. Despite having similar ICERs for baseline controlled utility and average utility models, it might be of more relevance when moving further with CEAC and EVPI models and the questions that stakeholders may raise regarding resource allocation for new research.

Several limitations which could have influenced the results should be considered. The model was based on the assumption that (all the women screened positive would be referred to usual care; those who screen negative and are true negative will not be depressed in the later visits; in the no screening arm, those with depression getting the treatment depend on the treatment-seeking behavior and get usual care; in the false negative arm, there is no later identification)—additional assumption of no spontaneous recovery in the baseline model, which was relaxed in scenario analysis. The assumptions were deemed necessary to simplify the model. Another limitation of the study would be on-the-ground depression progression, as the model did not relapse due to a shorter time horizon.

The choice of WTP threshold for this study was based on the WHO CHOICE recommendation of 1*GDP per capita to 3*GDP per capita, where 1*GDP per capita was chosen as the WTP threshold. No studies are available exploring the WTP threshold per QALY in Nepal, which makes it challenging to inform policy decisions. With a method to estimate cost-effectiveness thresholds based on health expenditures per capita and life expectancy at birth, Pichon-Riviere et al. estimate the WTP threshold per QALY for Nepal as 0.35 proportion of GDP per capita per QALY, resulting in USD 413 per QALY for the year 2019 (Pichon-Riviere et al., 2023). The choice of such a threshold can hugely impact the policy decision.

As there are no rigid guidelines on the screening time and interval around the screening of PND, our model only uses one-point screening antenatally. Is that sufficient? Multiple-point screening was not implemented to see the differences. A recent systematic review of recommendations on perinatal depression screening from the Organization for Economic Co-operation and Development (OECD) member countries, on the screening intervals, recommendations started from identifying as early as possible, first contact up to 32 weeks, third trimester antenatally, and four weeks to six months postnatally where some publications did not state the screening frequency and timing (El-Den et al., 2022). The question of either single-point screening or multiple points stays relevant to inform the policymakers.

The ongoing RCT ("Thinking Healthy Programme" for Perinatal Depression in Nepal, NCT05393479) in Nepal could help answer the question of the cost of care and provide a proper estimate for the intervention delivery cost. In addition, cost analysis of depression case management would be beneficial to check the usual care cost from the ongoing clinical trial as the cost plays a major role in developing countries. The use of the QALYs gained was based on a secondary study in a UK setting. So, given the uncertainty around the parameter itself and a different population setting, it calls for estimating the utility gains in Nepalese settings or LIMC settings to have representative utility values. The primary assumption of all screened positive getting the treatment and also all the 100% population being screened should be considered given that the health service utilization in mental health issues is low in Nepal.

Implication for policy

This study showed that the screening for PND was cost-effective, but we need to be careful with the limitations. Defining cost-effectiveness is a part of informing a policy, and the question of feasibility in terms of implementation should be considered. As out-of-pocket

payments make up a major share of current health expenditures in Nepal, affordability and accessibility should be considered.

Implication for further research

An integral part of any cost-effectiveness study lies with the WTP threshold. WTP stays an integral part of the decision-making process to have a complete value of information analysis. As no studies establish that, studies estimating the local values of the WTP threshold can be helpful for all policy decision-making. Health utility values from the local population help strengthen the cost-effectiveness's validity, rather than taking it from a secondary population when we want to inform the policy.

Conclusion

In the Nepalese setting, the cost per case detection via screening tools (EPDS/PHQ-9) was less than one USD, which paved the way for assessing the cost-effectiveness of screening tools. The cost per case treated was around 100 USD for screening alternatives, which suggested the screening alternatives be cost-effective as no WTP threshold could be applied. EPDS and PHQ-9 screening approaches were cost-effective under the 1*GDP per capita WTP threshold per QALY. However, multiple considerations must be taken when making decisions in resource-constrained settings; coverage and equity considerations should be considered due to limitations of data availability and short time horizons. The results should be viewed as provisional, with the need for additional research.

References

Aldridge, L. R., Luitel, N. P., Jordans, M. J. D., Bass, J. K., & Patenaude, B. (2022). Cost-effectiveness of psychological intervention within services for depression delivered by primary care workers in Nepal: economic evaluation of a randomized control trial. *Global Mental Health*, *9*, 499-507. https://doi.org/10.1017/gmh.2022.54

Bauer, A., Knapp, M., & Parsonage, M. (2016). Lifetime costs of perinatal anxiety and depression. *Journal of Affective Disorders*, *192*, 83-90. https://doi.org/https://doi.org/10.1016/j.jad.2015.12.005

Bhusal, B. R., Bhandari, N., Chapagai, M., & Gavidia, T. (2016). Validating the Edinburgh Postnatal Depression Scale as a screening tool for postpartum depression in Kathmandu, Nepal. *International Journal of Mental Health Systems*, *10*, Article 71. https://doi.org/10.1186/s13033-016-0102-6

CEIC. (2023). *Nepal GDP per Capita*. Retrieved 10/05/2023 from https://www.ceicdata.com/en/indicator/nepal/gdp-per-capita

Chalise, A., Shrestha, G., Paudel, S., & Poudyal, A. K. (2022). Antenatal depression and its associated factors among women of Godawari Municipality, Lalitpur, Nepal: a cross-sectional study. *BMJ Open*, *12*(11), e063513. https://doi.org/10.1136/bmjopen-2022-063513

Chisholm, D., Burman-Roy, S., Fekadu, A., Kathree, T., Kizza, D., Luitel, N. P., Petersen, I., Shidhaye, R., De Silva, M., & Lund, C. (2016). Estimating the cost of implementing district mental healthcare plans in five low- and middle-income countries: The PRIME study. *Br J Psychiatry*, *208*(s56), s71-s78. https://doi.org/10.1192/bjp.bp.114.153866

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry*, *150*(6), 782-786. https://doi.org/10.1192/bjp.150.6.782

Dennis, C. L., Hodnett, E., Kenton, L., Weston, J., Zupancic, J., Stewart, D. E., & Kiss, A. (2009). Effect of peer support on prevention of postnatal depression among high risk women: multisite randomised controlled trial. *BMJ*, *338*(7689), 280-284. https://doi.org/10.1136/bmj.a3064

Diksha, S., Debra, A., Amornrat, S., & Kathleen, B. (2021). Domestic and Family Violence and its Association with Mental Health Among Pregnant Women Attending Antenatal Care in a Tertiary Hospital of Eastern Nepal. *Journal of Nepal Health Research Council*, *19*(1). https://doi.org/10.33314/jnhrc.v19i1.2508

Dukhovny, D., Dennis, C.-L., Hodnett, E., Weston, J., Stewart, D. E., Mao, W., & Zupancic, J. A. F. (2013). Prospective Economic Evaluation of a Peer Support Intervention for Prevention of Postpartum Depression among High-Risk Women in Ontario, Canada. *Amer J Perinatol*, *30*(8), 631-642. https://doi.org/10.1055/s-0032-1331029

El-Den, S., Pham, L., Anderson, I., Yang, S., Moles, R. J., O'Reilly, C. L., Boyce, P., Raine, K. H., & Raynes-Greenow, C. (2022). *Perinatal depression screening: a systematic review of recommendations from member countries of the Organisation for Economic Co-operation and Development (OECD)* [871-893]. Wien ; New York :.

Fuhr, D. C., Weobong, B., Lazarus, A., Vanobberghen, F., Weiss, H. A., Singla, D. R., Tabana, H., Afonso, E., De Sa, A., D'Souza, E., Joshi, A., Korgaonkar, P., Krishna, R., Price, L. N., Rahman, A., & Patel, V. (2019). Delivering the Thinking Healthy Programme for perinatal depression through peers: an individually randomised controlled trial in India. *Lancet Psychiatry*, *6*(2), 115-127. https://doi.org/10.1016/S2215-0366(18)30466-8

Government of Nepal. (2023a). *Female Community Health Programme*. Ministry of Health and Population. Retrieved 25/10/2023 from https://mohp.gov.np/program/female-community-health-programme/en

Government of Nepal. (2023b). *Minimum Wages with effect from 2080-81*. Ministry of Labour, Employment and Social Security Retrieved from https://moless.gov.np/en/post/show/548

Grote, N. K., Simon, G. E., Russo, J., Lohr, M. J., Carson, K., & Katon, W. (2017). Incremental Benefit-Cost of MOMCare: Collaborative Care for Perinatal Depression Among Economically

Disadvantaged Women. *Psychiatr Serv*, *68*(11), 1164-1171. https://doi.org/10.1176/appi.ps.201600411

Henderson, C., Dixon, S., Bauer, A., Knapp, M., Morrell, C. J., Slade, P., Walters, S. J., & Brugha, T. (2019). Cost-effectiveness of PoNDER health visitor training for mothers at lower risk of depression: findings on prevention of postnatal depression from a cluster-randomised controlled trial. *Psychol Med*, *49*(8), 1324-1334. https://doi.org/10.1017/s0033291718001940

Heslin, M., Jin, H., Trevillion, K., Ling, X., Nath, S., Barrett, B., Demilew, J., Ryan, E. G., O'Connor, S., Sands, P., Milgrom, J., Bick, D., Stanley, N., Hunter, M. S., Howard, L. M., & Byford, S. (2022). *Cost-effectiveness of screening tools for identifying depression in early pregnancy: a decision tree model*. London :.

Institute for Health Metrics and Evaluation (IHME). (2020). *Population Forecasting* https://vizhub.healthdata.org/population-forecast/

Joshi, D., Shrestha, S., & Shrestha, N. (2019). Understanding the antepartum depressive symptoms and its risk factors among the pregnant women visiting public health facilities of Nepal. *PLoS One*, *14*(4), Article e0214992. https://doi.org/10.1371/journal.pone.0214992

Kohrt, B. A., Luitel, N. P., Acharya, P., & Jordans, M. J. D. (2016). Detection of depression in low resource settings: validation of the Patient Health Questionnaire (PHQ-9) and cultural concepts of distress in Nepal. *BMC Psychiatry*, *16*(1), 58. https://doi.org/10.1186/s12888-016-0768-y

Lee, E. W., Denison, F. C., Hor, K., & Reynolds, R. M. (2016a). *Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review*. London.

Lee, E. W., Denison, F. C., Hor, K., & Reynolds, R. M. (2016b). Web-based interventions for prevention and treatment of perinatal mood disorders: a systematic review. *BMC Pregnancy Childbirth*, *16*(1), 38-38. https://doi.org/10.1186/s12884-016-0831-1

Leverton, T. J., & Elliott, S. A. (2000). Is the EPDS a magic wand?: 1. A comparison of the Edinburgh Postnatal Depression Scale and health visitor report as predictors of diagnosis on the Present State Examination. *Journal of reproductive and infant psychology*, *18*(4), 279-296. https://doi.org/10.1080/713683048

Littlewood, E., Ali, S., Dyson, L., Keding, A., Ansell, P., Bailey, D., Bates, D., Baxter, C., Beresford-Dent, J., Clarke, A., Gascoyne, S., Gray, C., Hackney, L., Hewitt, C., Hutchinson, D., Jefferson, L., Mann, R., Marshall, D., McMillan, D., . . . Gilbody, S. (2018). Identifying perinatal depression with case-finding instruments: a mixed-methods study (BaBY PaNDA – Born and Bred in Yorkshire PeriNatal Depression Diagnostic Accuracy). *Health services and delivery research*, *6*(6), 1-210. https://doi.org/10.3310/hsdr06060

Loksewajob. (2023). *New Salary Scale of Health Workers in Nepal with Position 2079*. https://loksewajob.com/salary-scale-of-health-workers-in-nepal/

Manca, A., Hawkins, N., & Sculpher, M. J. (2005). Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ*, *14*(5), 487-496. https://doi.org/10.1002/hec.944

McKee, K., Admon, L. K., Winkelman, T. N. A., Muzik, M., Hall, S., Dalton, V. K., & Zivin, K. (2020). Perinatal mood and anxiety disorders, serious mental illness, and delivery-related health outcomes, United States, 2006-2015. *BMC Womens Health*, *20*(1), 150-150. https://doi.org/10.1186/s12905-020-00996-6

Meltzer-Brody, S., & Rubinow, D. (2021). An Overview of Perinatal Mood and Anxiety Disorders: Epidemiology and Etiology. In E. Cox (Ed.), *Women's Mood Disorders: A Clinician's Guide to Perinatal Psychiatry* (pp. 5-16). Springer International Publishing. https://doi.org/10.1007/978-3-030-71497-0_2

Ministry of Health and Population [Nepal], New ERA, & ICF. (2023). *Nepal Demographic and Health Survey 2022*. Ministry of Health and Population [Nepal],. https://dhsprogram.com/pubs/pdf/FR379/FR379.pdf

National Collaborating Centre for Mental Health (UK). (2014). *Antenatal and postnatal mental health : clinical management and service guidance* (updated edition. ed., Vol. Volume 192). The British Psychological Society ; The Royal College of Psychiatrists. https://www.ncbi.nlm.nih.gov/books/NBK305023/

Nepal Health Research Council. (2020). *National Mental Health Survey, Nepal-2020 Factsheets (Adults)*. Retrieved from https://nhrc.gov.np/wp-content/uploads/2020/09/Factsheet-Adults-1.pdf

Paudyal, N., Parajuli, K. R., Garcia Larsen, V., Adhikari, R. K., Devkota, M. D., Rijal, S., Chitekwe, S., & Torlesse, H. (2022). A review of the maternal iron and folic acid supplementation programme in Nepal: Achievements and challenges. *Matern Child Nutr*, *18 Suppl 1*(Suppl 1), e13173. https://doi.org/10.1111/mcn.13173

Paulden, M., Palmer, S., Hewitt, C., & Gilbody, S. (2009). *Screening for postnatal depression in primary care: cost effectiveness analysis* [b5203-b5203]. London :.

Petrou, S., Cooper, P., Murray, L., & Davidson, L. L. (2006). Cost-effectiveness of a preventive counseling and support package for postnatal depression. *Int J Technol Assess Health Care*, *22*(4), 443-453. https://doi.org/10.1017/S0266462306051361

Pichon-Riviere, A., Drummond, M., Palacios, A., Garcia-Marti, S., & Augustovski, F. (2023). Determining the efficiency path to universal health coverage: cost-effectiveness thresholds for 174 countries based on growth in life expectancy and health expenditures. *Lancet Glob Health*, *11*(6), e833-e842. https://doi.org/10.1016/S2214-109X(23)00162-6

Premji, S., McDonald, S. W., McNeil, D. A., & Spackman, E. (2021). Maximizing maternal health and value for money in postpartum depression screening: a cost-effectiveness analysis using the All Our Families cohort and administrative data in Alberta, Canada. *J Affect Disord*, *281*, 839-846. https://doi.org/10.1016/j.jad.2020.11.051

Ride, J., Lorgelly, P., Tran, T., Wynter, K., Rowe, H., & Fisher, J. (2016). *Preventing postnatal maternal mental health problems using a psychoeducational intervention: the cost-effectiveness of What Were We Thinking* [e012086]. [London] :.

Roddy Mitchell, A., Gordon, H., Lindquist, A., Walker, S. P., Homer, C. S. E., Middleton, A., Cluver, C. A., Tong, S., & Hastie, R. (2023). Prevalence of Perinatal Depression in Low- and Middle-Income Countries: A Systematic Review and Meta-analysis. *JAMA Psychiatry*, *80*(5), 425-431. https://doi.org/10.1001/jamapsychiatry.2023.0069

Shakya, R., Situala, S., & Shyangwa, P. M. (2008). Depression during Pregnancy in a Tertiary Care Center of Eastern Nepal. *Journal of Nepal Medical Association*, *47*(3), 128-131. <Go to ISI>://WOS:000264499700009

Slomian, J., Honvo, G., Emonts, P., Reginster, J.-Y., & Bruyère, O. (2019). Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. *Womens Health (Lond)*, *15*, 1745506519844044-1745506519844044. https://doi.org/10.1177/1745506519844044

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & and the Patient Health Questionnaire Primary Care Study, G. (1999). Validation and Utility of a Self-report Version of PRIME-MD: The PHQ Primary Care Study. *JAMA : the journal of the American Medical Association*, *282*(18), 1737-1744. https://doi.org/10.1001/jama.282.18.1737

Stuge, B., Sorensen, H., Acharya, R. S., & Tveter, A. T. (2022). Prevalence of Antenatal Depressive Symptoms and its Associated Factors among Pregnant Nepalese Women with and Without Low Back- and/or Pelvic Girdle Pain. *J Nepal Health Res Counc*, *19*(4), 784-791. https://doi.org/10.33314/jnhrc.v19i04.3849

Trading Economics. (2023). *Nepal Consumer Price Index Cpi*. Retrieved 11/10/2023 from https://tradingeconomics.com/nepal/consumer-price-index-cpi

Waqas, A., Koukab, A., Meraj, H., Dua, T., Chowdhary, N., Fatima, B., & Rahman, A. (2022). Screening programs for common maternal mental health disorders among perinatal women: report of

the systematic review of evidence. *BMC Psychiatry*, *22*(1), 54. https://doi.org/10.1186/s12888-022-03694-9

WHO recommendations on maternal and newborn care for a positive postnatal experience. (2022). World Health Organization.

Wilkinson, A., Anderson, S., & Wheeler, S. B. (2017). Screening for and Treating Postpartum Depression and Psychosis: A Cost-Effectiveness Analysis. *Matern Child Health J*, *21*(4), 903-914. https://doi.org/10.1007/s10995-016-2192-9

Wilkinson, A., Anderson, S., & Wheeler, S. B. (2017). *Screening for and Treating Postpartum Depression and Psychosis: A Cost-Effectiveness Analysis* [903-914]. [New York, N.Y.] :.

World Health Organization. (2016). *mhGAP intervention guide for mental, neurological and substance use disorders in non-specialized health settings: mental health Gap Action Programme (mhGAP)* (version 2.0 ed.). World Health Organization. https://apps.who.int/iris/handle/10665/250239

World Health Organization. (2022). *WHO recommendations on maternal and newborn care for a positive postnatal experience*. World Health Organization.

World Health Organization. (2023a). *Depressive disorder (depression)*. World Health Organiation. Retrieved 26/10/2023 from https://www.who.int/news-room/fact-sheets/detail/depression/?gclid=CjwKCAjws9ipBhB1EiwAccEi1C2dYipmIX9Siri4-YYbapOM6nTzONiAFQbe5LRcwkqEVEpE3FfVVRoCcJAQAvD_BwE

World Health Organization. (2023b). *Indicators* Maternal, newborn, child and adolescent health and ageing. https://www.who.int/data/maternal-newborn-child-adolescent-ageing/indicator-explorer-new/mca/crude-birth-rate-(births-per-1000-population)

Appendix

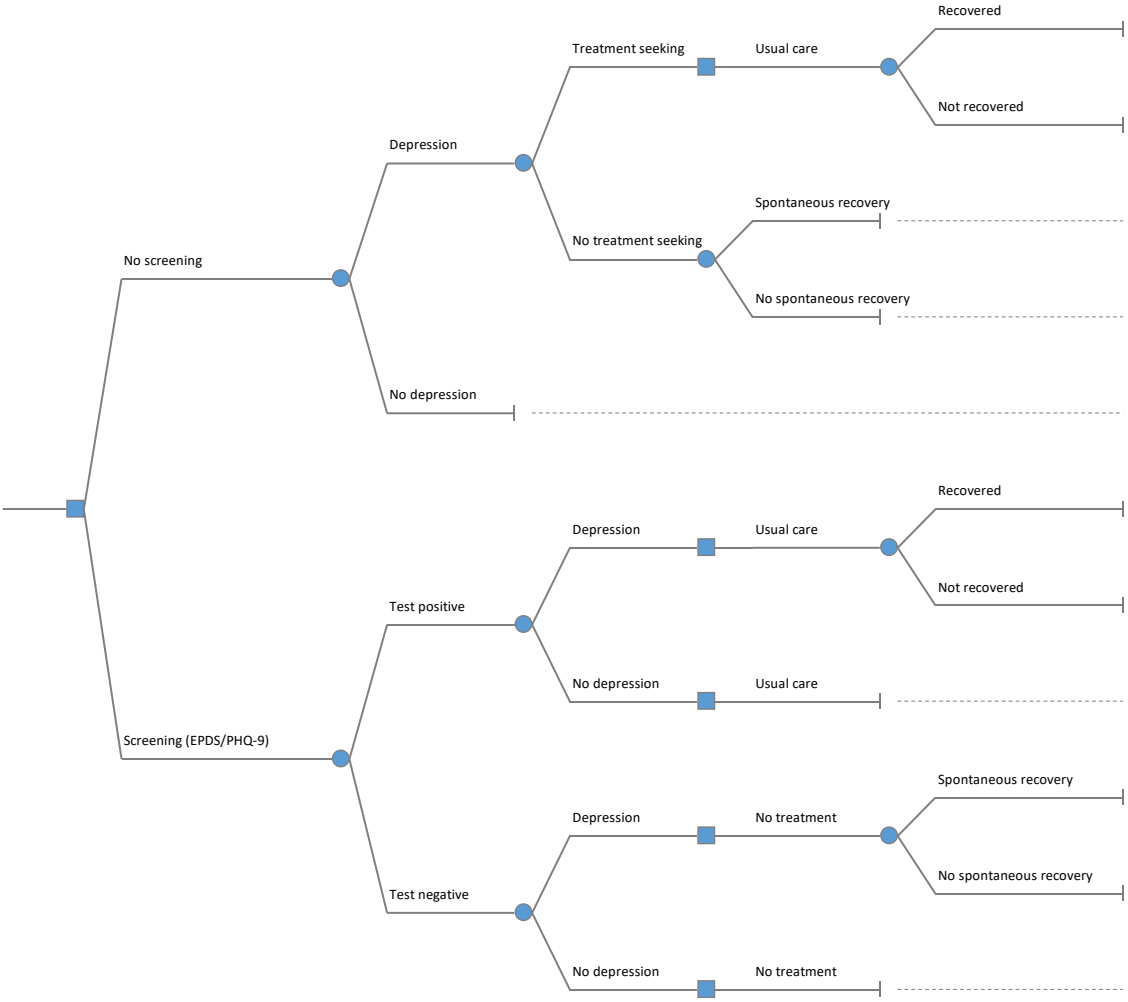
Table of Contents

Tables

Figures

Appendix A: Model pathway for spontaneous recovery model

Figure 8 *Detection and treatment pathway for scenario with spontaneous recovery*



Appendix B: Probabilistic ICERs

Outcome as number of treated model

Table 9 *Probabilistic ICERs per treated for EPDS and PHQ-9 compared to no screening for the base model and different scenarios*

	Incremental treated	Incremental costs (USD)	ICER
Base case			
EPDS	0.128	8.24	64.16
PHQ-9	0.132	9.35	70.93
Scenario 1			
EPDS	0.097	10.05	103.78
PHQ-9	0.1	11.65	116.85
Scenario 2			
EPDS	0.129	11.52	86.71
PHQ-9	0.132	12.63	95.84
Scenario 3			
EPDS	0.128	10.96	85.27
PHQ-9	0.131	12.41	94.31
Scenario 4			
EPDS	0.07	9.02	128.22
PHQ-9	0.073	10.52	144.72
Scenario 5			
EPDS	0.059	8.13	138.3
PHQ-9	0.06	9.22	152.84

Note. EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio, Scenario 1 = Peer delivered Thinking Healthy Program delivered to screening arm in addition to usual care; Scenario 2 = Auxiliary Nurse Midwives are used as the screening agent; Scenario 3 = Baseline utility adjustment method used to estimate QALYs and utility gains are provided only to the treatment arm; Scenario 4 = Sensitivity and Specificity parameters taken from meta-analysis from National Institute for Health and Care Excellence guidelines; Scenario 5 = No spontaneous recovery relaxed and allowed on the in the no-treatment-seeking population of the no-screening arm and the false negative arm of screening approaches.

Outcome as QALY model

Table 10 *Probabilistic ICERs per QALY for EPDS and PHQ-9 compared to no screening for the base model and different scenarios*

	Incremental QALYs	Incremental costs	ICER
Base case			
EPDS	0.0149	8.23	552.15
PHQ-9	0.0152	9.34	612.29
Scenario 1			
EPDS	0.0128	10.16	796.57
PHQ-9	0.0131	11.76	894.47
Scenario 2			
EPDS	0.015	8.42	555.44
PHQ-9	0.016	9.52	611.79
Scenario 3			
EPDS	0.015	8.21	539.09
PHQ-9	0.016	9.3	596.29
Scenario 4			
EPDS	0.009	5.43	602.83
PHQ-9	0.011	7.04	626.31
Scenario 5			
EPDS	0.009	8.3	846.19
PHQ-9	0.01	9.43	933.22

Note. QALY = Quality Adjusted Life Years; EPDS = Edinburgh Postnatal Depression; PHQ-9 = Patient Health Questionnaire 9 items; USD = US Dollar, ICER = Incremental Cost Effectiveness Ratio, Scenario 1 = Peer delivered Thinking Healthy Program delivered to screening arm in addition to usual care; Scenario 2 = Auxiliary Nurse Midwives are used as the screening agent; Scenario 3 = Baseline utility adjustment method used to estimate QALYs and utility gains are provided only to the treatment arm; Scenario 4 = Sensitivity and Specificity parameters taken from meta-analysis from National Institute for Health and Care Excellence guidelines; Scenario 5 = No spontaneous recovery relaxed and allowed on the in the no-treatment-seeking population of the no-screening arm and the false negative arm of screening approaches.

Appendix C: Scatterplots, CEAC, and CEAF for different scenarios

Scenario 1: Peer-delivered Thinking Healthy Program (THHP) intervention provided to the screening alternatives

Figure 9 *Scatterplot for incremental treated and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

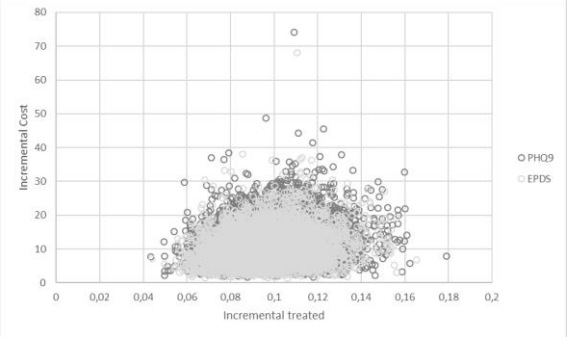


Figure 10 *Scatterplot for incremental Quality Adjusted Life Years (QALY) and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

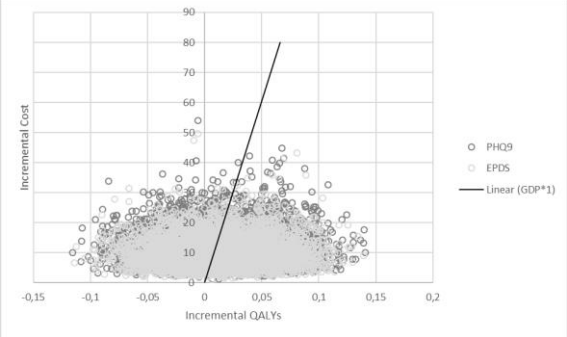


Figure 11 *Cost-effectiveness acceptability curve for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

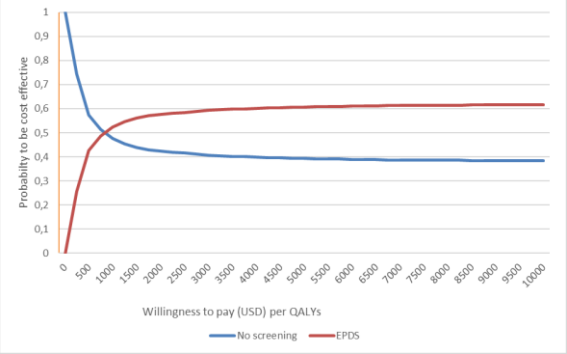


Figure 12 *Cost-effectiveness acceptability curve for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*

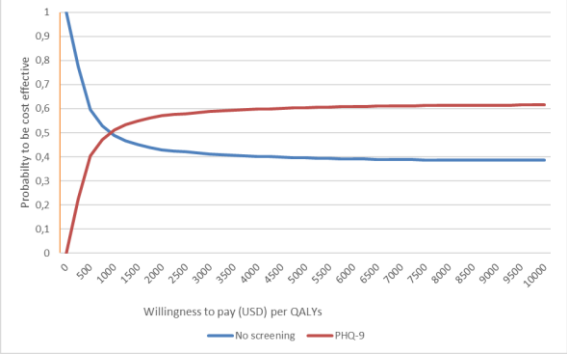


Figure 13 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

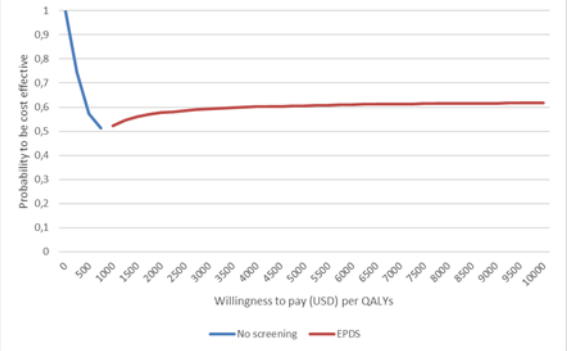
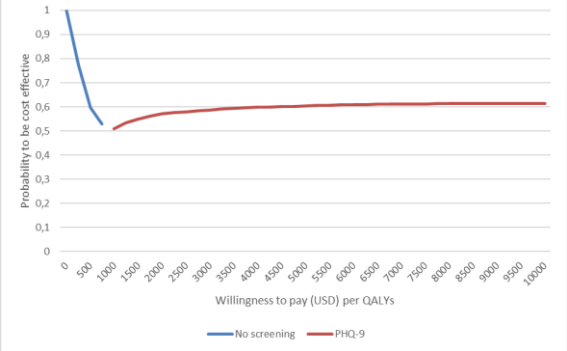


Figure 14 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*



Scenario 2: Auxiliary Nurse Midwives (ANM) is used for screening

Figure 15 *Scatterplot for incremental treated and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

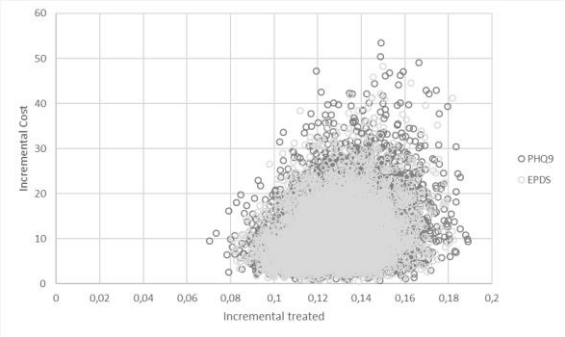


Figure 16 *Scatterplot for incremental Quality Adjusted Life Years (QALY) and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

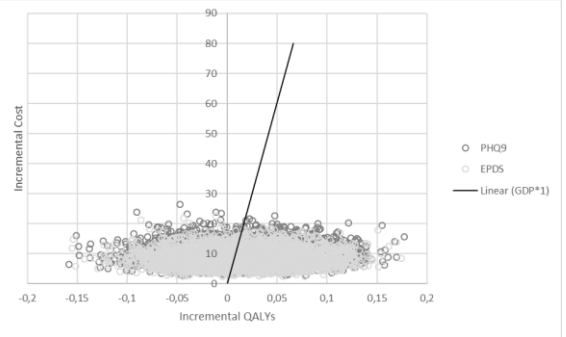


Figure 17 *Cost-effectiveness acceptability curve for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

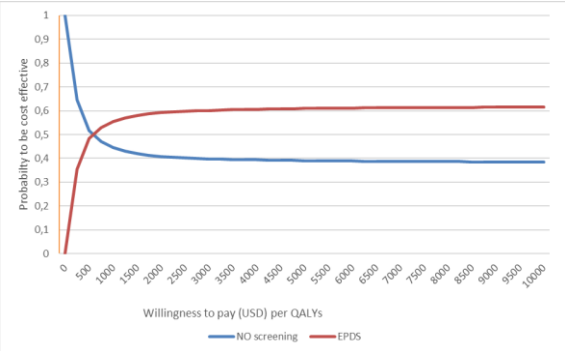


Figure 18 *Cost-effectiveness acceptability curve for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*

Figure 19 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

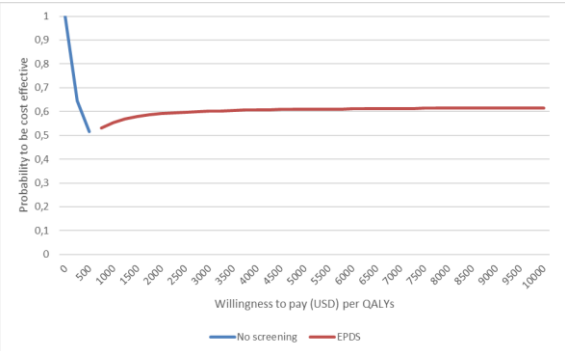
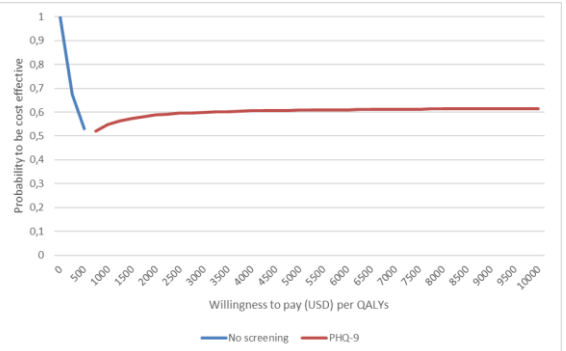


Figure 20 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*



Scenario 3: Use of Quality Adjusted Life Years (QALYs) estimated from controlling for baseline utility

Figure 21 *Scatterplot for incremental treated and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

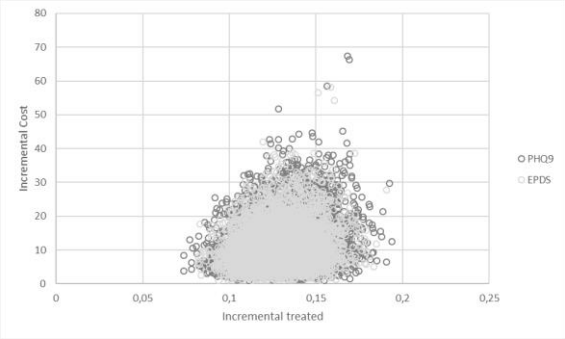


Figure 22 *Scatterplot for incremental Quality Adjusted Life Years (QALY) and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

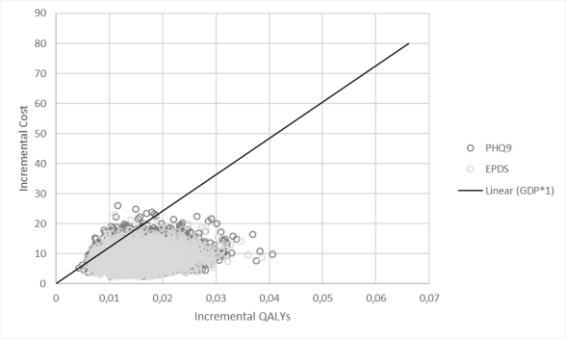


Figure 23 *Cost-effectiveness acceptability curve for Edinburgh Postnatal Depression Score (EPDS) relative to no screening*

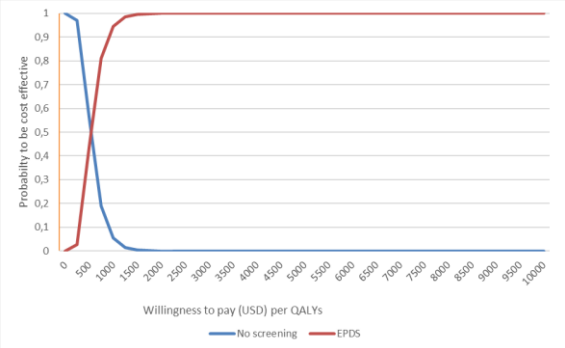


Figure 24 *Cost-effectiveness acceptability curve for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*

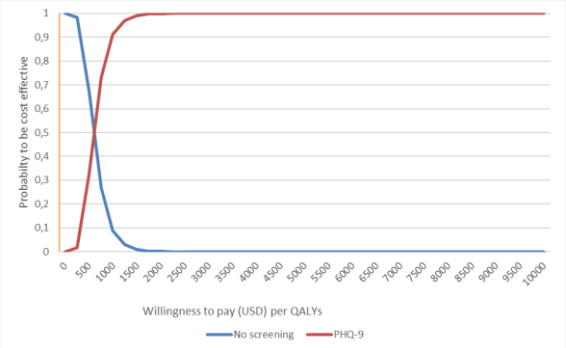


Figure 25 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

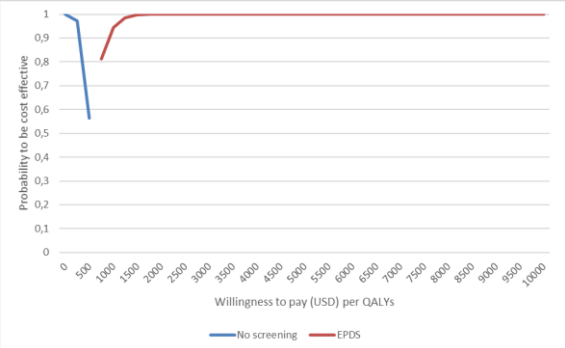
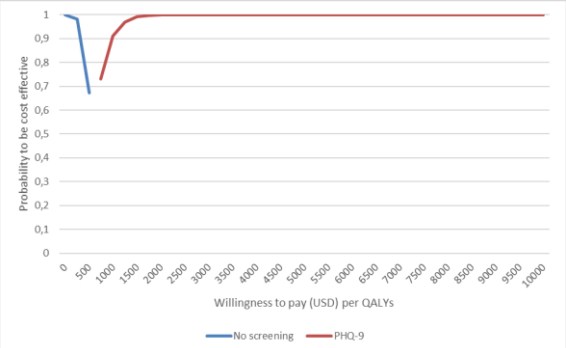


Figure 26 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*



Scenario 4: Sensitivity and specificity for screening alternatives from National Institute for Health and Care Excellence (NICE) guidelines

Figure 27 *Scatterplot for incremental treated and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

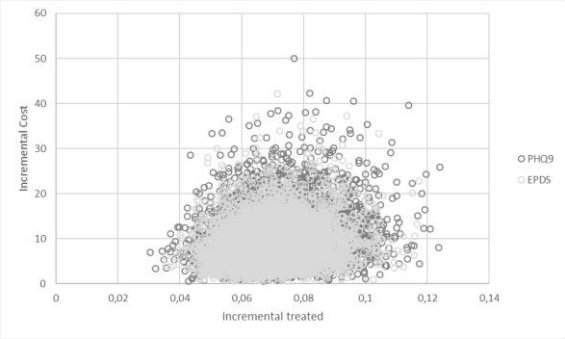


Figure 28 *Scatterplot for incremental Quality Adjusted Life Years (QALY) and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

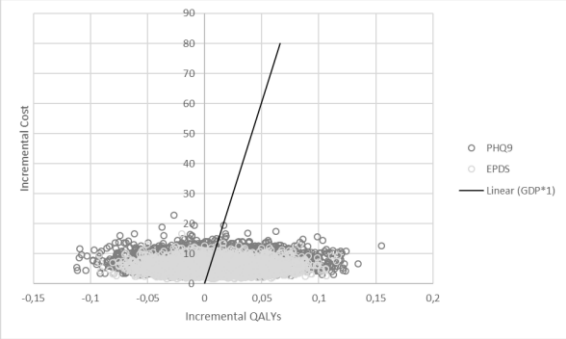


Figure 29 *Cost-effectiveness acceptability curve for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

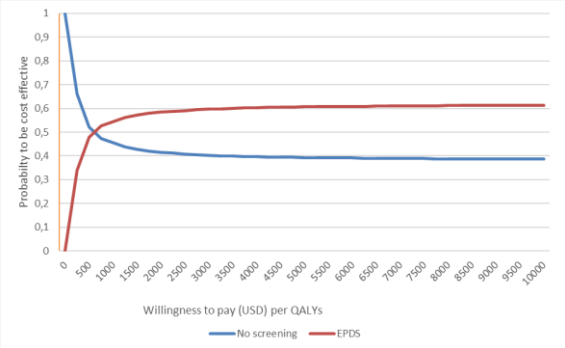


Figure 30 *Cost-effectiveness acceptability curve for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*

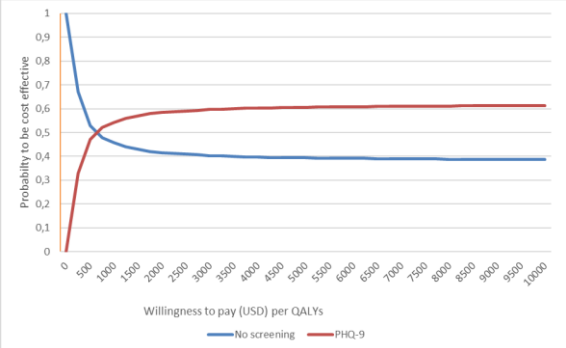


Figure 31 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

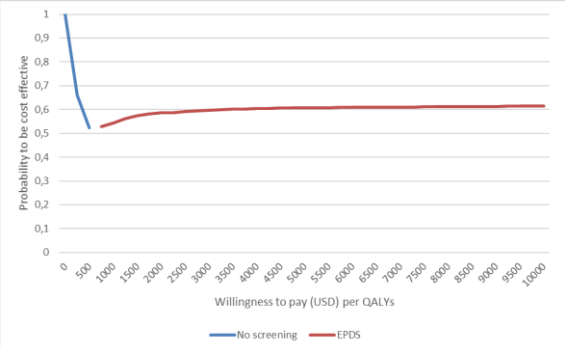
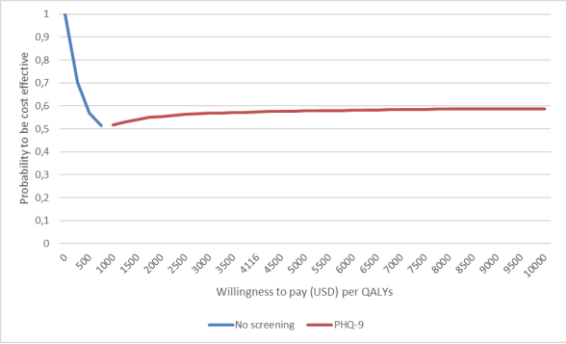


Figure 32 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*



Scenario 5: Spontaneous recovery model

Figure 33 *Scatterplot for incremental treated and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

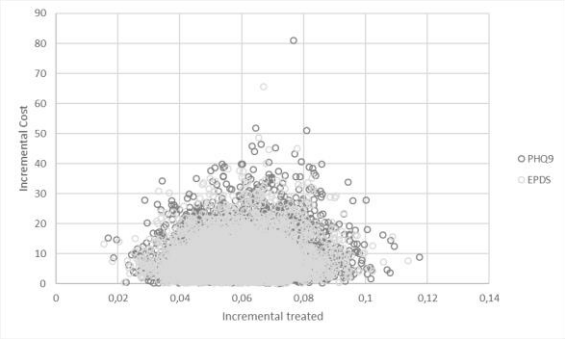


Figure 34 *Scatterplot for incremental Quality Adjusted Life Years (QALY) and incremental cost for Edinburgh Postnatal Depression Score (EPDS) and Patient Health Questionnaire 9 items (PHQ-9) relative to no screening*

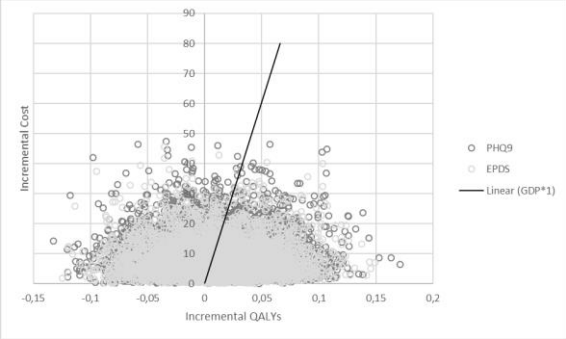


Figure 35 *Cost-effectiveness acceptability curve for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

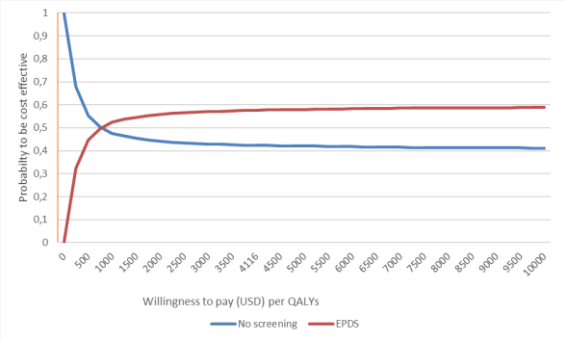


Figure 36 *Cost-effectiveness acceptability curve for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*

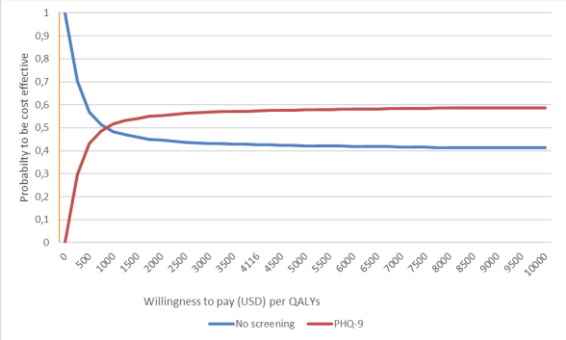


Figure 37 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening*

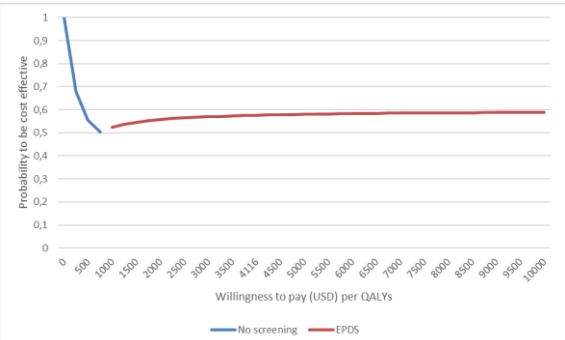
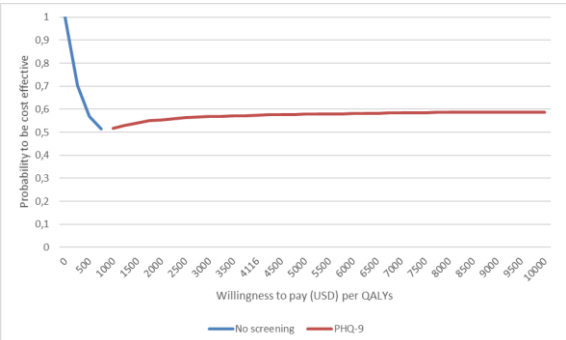


Figure 38 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening*



Appendix D: CHEERS checklist

CHEERS checklist

Table 11 *CHEERS checklist*

	Item	Guidance for Reporting	Reported in section
Title			
Title	1	Identify the study as an economic evaluation and specify the interventions being compared.	Page I
Abstract			
Abstract	2	Provide a structured summary that highlights context, key methods, results, and alternative analyses.	Page VIII
Introduction			
Background and objectives	3	Give the context for the study, the study question and its practical relevance for decision making in policy or practice.	Page 1-3
Methods			
Health economic analysis plan	4	Indicate whether a health economic analysis plan was developed and where available.	NA
Study population	5	Describe characteristics of the study population (such as age range, demographics, socioeconomic, or clinical characteristics).	Methods: Target population and setting, clinical input parameters
Setting and location	6	Provide relevant contextual information that may influence findings.	Methods: Target population and setting
Comparators	7	Describe the interventions or strategies being compared and why chosen.	Methods: Screening strategies and treatment strategies
Perspective	8	State the perspective(s) adopted by the study and why chosen.	Methods: Resource use and unit costs
Time horizon	9	State the time horizon for the study and why appropriate.	Methods: Time horizon
Discount rate	10	Report the discount rate(s) and reason chosen.	NA
Selection of outcomes	11	Describe what outcomes were used as the measure(s) of benefit(s) and harm(s).	Methods: Outcomes
Measurement of outcomes	12	Describe how outcomes used to capture benefit(s) and harm(s) were measured.	Methods: Outcomes
Valuation of outcomes	13	Describe the population and methods used to measure and value outcomes.	Methods: Outcomes
Measurement and valuation of resources and costs	14	Describe how costs were valued.	Methods: Resource use and unit costs

Currency, price date, and conversion	15	Report the dates of the estimated resource quantities and unit costs, plus the currency and year of conversion.	Methods: Page 10
Rationale and description of model	16	If modelling is used, describe in detail, and why used. Report if the model is publicly available and where it can be accessed.	Methods: Model structure
Analytics and assumptions	17	Describe any methods for analysing or statistically transforming data, any extrapolation methods, and approaches for validating any model used.	Methods: Model assumptions
Characterizing Heterogeneity	18	Describe any methods used for estimating how the results of the study vary for sub-groups.	NA
Characterizing distributional effects	19	Describe how impacts are distributed across different individuals or adjustments made to reflect priority populations.	NA
Characterizing uncertainty	20	Describe methods to characterize any sources of uncertainty in the analysis.	Methods: Sensitivity analysis
Approach to engagement with patients and others affected by the study	21	Describe any approaches to engage patients or service recipients, the general public, communities, or stakeholders (e.g., clinicians or payers) in the design of the study.	NA
Results			
Study parameters	22	Report all analytic inputs (e.g., values, ranges, references) including uncertainty or distributional assumptions.	Methods: Table1, table2 and table 3
Summary of main results	23	Report the mean values for the main categories of costs and outcomes of interest and summarise them in the most appropriate overall measure.	Results: Results, table 4, table 5, table 6, table 7, table 8, appendix
Effect of uncertainty	24	Describe how uncertainty about analytic judgments, inputs, or projections affect findings. Report the effect of choice of discount rate and time horizon, if applicable.	Results: Results, figure 2, figure 3, figure 4, figure 5, figure6, figure7, and appendix (partially applicable)
Effect of engagement with patients and others affected by the study	25	Report on any difference patient/service recipient, general public, community, or stakeholder involvement made to the approach or findings of the study	NA
Discussion			
Study findings, limitations, generalizability, and current knowledge	26	Report key findings, limitations, ethical or equity considerations not captured, and how these could impact patients, policy, or practice.	Discussion
Other relevant information			

Source of funding	27	Describe how the study was funded and any role of the funder in the identification, design, conduct, and reporting of the analysis	NA
Conflicts of interests	28	Report authors conflicts of interest according to journal or International Committee of Medical Journal Editors requirements.	NA

Husereau D, Drummond M, Augustovski F, de Bekker-Grob E, Briggs AH, Carswell C, Caulley L, Chaiyakunapruk N, Greenberg D, Loder E, Mauskopf J, Mullins CD, Petrou S, Pwu RF, Staniszewska S; CHEERS 2022 ISPOR Good Research Practices Task Force. Consolidated Health Economic Evaluation Reporting Standards 2022 (CHEERS 2022) Statement: Updated Reporting Guidance for Health Economic Evaluations. BMJ. 2022;376:e067975.

Appendix E: Understanding pEVPI curves

Figure 39 *Cost-effectiveness acceptability frontier for Edinburgh Postnatal Depression Score (EPDS) screening relative to no screening and population expected value of perfect information (pEVPI)*

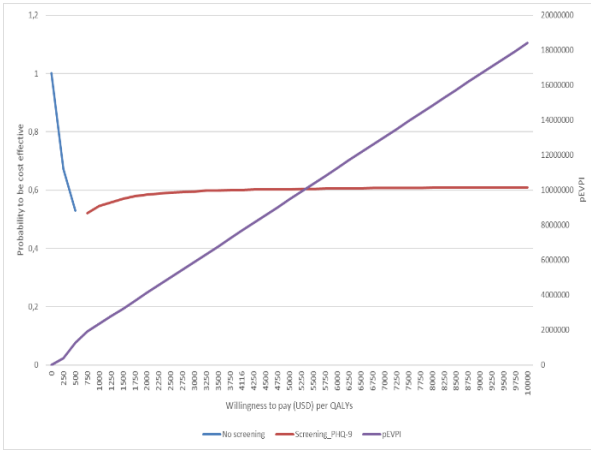
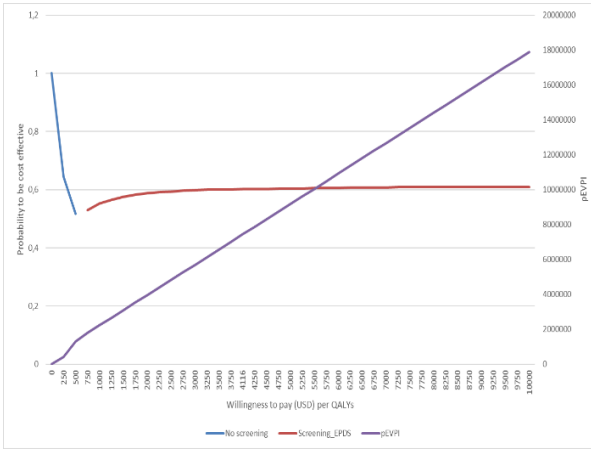


Figure 40 *Cost-effectiveness acceptability frontier for Patient Health Questionnaire 9 items (PHQ-9) screening relative to no screening and population expected value of perfect information (pEVPI)*



pEVPI increases up to a local maximum at the point where the threshold value is equal to the value of the ICER for the adoption of the screening approaches relative to no screening approach. Up to this value, EVPI is increasing. This is because the uncertainty surrounding the adoption decision is increasing (error probability increasing), as is the value applied to the consequences of making an incorrect decision. After this point, the uncertainty surrounding the adoption decision begins to fall, and the consequences associated with making an incorrect decision continue to rise. The overall EVPI depends on the interaction between these terms. For the threshold values slightly higher than the ICER for adopting the screening approaches relative to no screening approach, the EVPI falls, implying that the probability of an incorrect decision is reducing at a rate sufficient to offset the increasing cost of making an incorrect decision. However, this fall is not noticeable as the WTP threshold is plotted for

higher intervals, but in higher thresholds than the decision switch probabilistic ICER, the pEVPI curve rises. As in the later part, the value applied to the consequences of making incorrect decisions rises. As in the model, screening does not completely dominate the no screening, so there is always a substantial amount of error probability present, which then, with the higher value applied to the consequences of making decisions at higher WTP thresholds, the curve is only increasing.

Appendix F: Search strategy in Scopus

Context of PND in Nepal

Search string: (TITLE-ABS-KEY (depression* OR depressive OR (mood PRE/1 disorder*) OR anxiet* OR anxious*) AND TITLE (pregnant OR pregnanc* OR gravidit* OR perinatal OR peri-natal OR prenatal OR pre-natal OR antenatal OR antepartum OR postnatal OR post-natal OR postpartum OR post-partum OR parturit* OR birth* OR childbirth OR maternal OR midwi* OR neonatal* OR neo-natal* OR obstetric*)) AND (TITLE-ABS-KEY (nepal))

47 studies, 08/07/2023

Cost effectiveness analysis in PND

Search string: (TITLE-ABS-KEY (depression* OR depressive OR (mood PRE/1 disorder*) OR anxiet* OR anxious*) AND TITLE (pregnant OR pregnanc* OR gravidit* OR perinatal OR peri-natal OR prenatal OR pre-natal OR antenatal OR antepartum OR postnatal OR post-natal OR postpartum OR post-partum OR parturit* OR birth* OR childbirth OR maternal OR midwi* OR neonatal* OR neo-natal* OR obstetric*)) AND (TITLE-ABS-KEY (cost* OR economy*))

943 studies, 09/07/2023

Part II

Preface

This section comprises two main components: an in-depth exploration of perinatal depression, encompassing an extended background and theory, and a concise overview of economic evaluation. The initial part will delve into perinatal depression, elucidating pertinent screening methods and treatments specifically focusing on the intervention used in the study above. Additionally, a brief insight into Nepal's health system will be presented. The subsequent section will offer a brief theory for economic evaluation.

Table of Contents

Figures

Extended Background

Perinatal depression

Pregnancy is when the fetus develops inside a woman's uterus, accompanied by emotional and physical changes for mothers. Mood and anxiety disorders are common in the process of psychological change. Perinatal mood and anxiety disorders (PMADs) are mood or anxiety disorders that occur during pregnancy or postpartum (Meltzer-Brody & Rubinow, 2021). The most common type of PMAD is perinatal depression (PND). Anxiety disorder also occurs in the perinatal period, often co-occurring with perinatal mood disorders. The spectrum of a perinatal anxiety disorder includes generalized anxiety disorder (GAD), obsessive-compulsive disorder (OCD), and post-traumatic stress disorder (PTSD) (Cox, 2021). About 25-40% of mothers suffer mood liability and mild depression during the first weeks after parturitions called postpartum blues, which is self-limiting in nature, about 10-15% have a depressive disorder, and rare cases of psychosis during the infant's first year (Riecher-Rössler & Steiner, 2005, p. 6).

PND encompass major and minor depressive episode that occurs either during pregnancy or within the first 12 months following delivery. PND is common with major depressive disorder (MDD) outside of the perinatal period. Looking at the Diagnostic and Statistical Manual of Mental Disorders, 5th edition (DSM 5), the criteria for depression during the perinatal period mirrors the diagnostic criteria for MDD. The DSM 5 uses the "peripartum" specifier to represent the women who meet the criteria for depression during pregnancy or within the first four weeks of postpartum. The diagnosis specifically requires depressed mood or loss of interest, in addition to depressed mood or loss of interest with at least four additional symptoms such as significant weight loss or weight gain, insomnia or hypersomnia nearly every day, agitation, fatigue, inappropriate guilt, concentration problems, suicidal thoughts (Cox, 2021).

Risk factors

Risk factors for perinatal depression, in general, can be adolescent pregnancy, difficult birth experience, poverty, gender discrimination, poor nutrition, low education opportunities, physical health conditions, little or no support, natural disasters, gender-based violence, unwanted pregnancy, fertility difficulties, and substance abuse (World Health Organization, 2022a). These risk factors can play different roles in different communities and country settings. Low educational attainment, low socio-economic status, history of mental illness,

delivery of a preterm baby, interpersonal violence, and lack of social support are major risk factors for developing PMAD (Joshi et al., 2019; Shrestha et al., 2014).

Impact of perinatal depression

PMADs are a challenging global health issue as the negative impact affects women, families, and communities. The offspring bear most of the social costs of PMAD, as highlighted in the study to assess the lifetime cost of perinatal anxiety and depression (Bauer et al., 2016). A systematic review conducted to assess the maternal and infant consequences of postpartum depression found maternal postpartum depression was associated with more negative maternal physical and psychological health with worse quality of life (Slomian et al., 2019). Depressed mothers seemed to experience more difficulties in their social relationships. Many studies showed significant and negative associations between maternal postpartum depressive symptoms and infant cognitive development, language development, infant behaviors, and quality of sleep (Slomian et al., 2019). In terms of mother-infant interaction, studies showed postpartum depression associated with poor maternal care along with negative effects on breastfeeding (discontinued breastfeeding, less healthy feeding practices, breastfeeding problems, lower satisfaction) (Slomian et al., 2019). A population-based retrospective birth cohort study of women in South Carolina found perinatal mental health was associated with a higher risk of severe maternal morbidity, higher risk of preterm birth and low birth weight, hypertensive disorders, and cesarean section (Runkle et al., 2023).

Guidelines for screening and treatment of perinatal depression

A recent systematic review of recommendations on perinatal depression screening from the Organization for Economic Co-operation and Development (OECD) member countries demonstrated different screening tools with different timing and frequency. EPDS was the most endorsed screening tool, followed by the PHQ-9. On the screening intervals, recommendations started from identifying as early as possible, first contact up to 32 weeks, third trimester antenatally, and four weeks to six months postnatally. Some publications did not state the screening frequency and timing (El-Den et al., 2022). Similar nature screening was also demonstrated in a systematic review focusing on the European nations; clinical practice guidelines with recommendations for peripartum depression, where Norway (Guideline for birth assistance. Mental health in pregnancy), Netherlands (Guideline for SSRI use during pregnancy and lactation) Serbia (Treatment of depression- TD- SNGn national guidelines of good clinical practice) and United Kingdom (Consensus guidance on the use of psychotropic medication preconception, in pregnancy and postpartum 2017) did not

recommend the screening. In contrast, clinical practice guidelines (CGPs) from Belgium, Denmark, Finland, Germany, Italy, and Malta, both CGPs from Spain, and antenatal and postnatal mental health from NICE and management of perinatal mood disorders from SIGN from the United Kingdom recommended screening (Motrico et al., 2022).

On the treatment referral front, a wide range of referrals was made from mental health care providers to primary care clinicians (El-Den et al., 2022). With CGPs, psychological treatment (most recommended CBT and IPT) was recommended as a first-line intervention and pharmacological initiation based on psychotherapy non-response or depression severity (Motrico et al., 2022).

WHO recommendation on maternal and newborn care for positive postnatal experience

WHO developed a consolidated guideline of new and existing recommendations on routine postnatal care, which provides recommendations for care during the postnatal period. Under the mental health interventions category, WHO recommended two recommendations: recommendation 18 (Screening for postpartum depression and anxiety) and recommendation 19 (Prevention of postpartum depression and anxiety) (World Health Organization, 2022b).

The guideline recommends the Edinburgh Postnatal Depression Scale (EPDS) or Patient Health Questionnaire-9 (PHQ-9) to screen for common mental health conditions. The Guideline Development Group observed that studies demonstrating a decrease in postpartum depression and anxiety involved widespread screening for mental health issues by well-trained healthcare professionals, along with the subsequent confirmation of diagnoses and implementation of treatment strategies. In cases where women exhibit clinically significant symptoms or have identifiable risk factors, they should be presented with psychological interventions, such as cognitive-behavioral therapy or interpersonal therapy. The decision to provide these interventions should be collaborative, considering the woman's preferences and the care provider's capacity to deliver the intervention, considering factors such as training, expertise, and experience(World Health Organization, 2022b).

Screening of perinatal depression

In a systematic review of screening programs for common maternal mental health disorders among perinatal women, the majority of the studies used EPDS for the assessment of postpartum depression, which was followed by PHQ-9 and non-directive counseling, psychoeducation, and pharmacological therapy, the most used treatment strategies. Meta-analysis of postpartum depression, assessing rates of depressive disorder among pregnant

women or postpartum women undergoing screening for perinatal depression, had the pooled result indicating a positive impact in favor of the intervention group (Ahmed Waqas et al., 2022). The screening tools recommended by WHO (EPDS/PHQ-9) (World Health Organization, 2022b), deemed relevant in our study, are further discussed.

EPDS

It is a tool that helps the health professional in screening for depressive symptoms among pregnant and postpartum women. It was developed in the UK by Cox et al. EPDS is a ten-item questionnaire that has a Likert scale response ranging from 0 (not at all) to 3 (all the time). The symptoms assessed by the EPDS tool are mood reactivity, anhedonia, self-blame, anxiety, feelings of panic, coping ability, difficulty in sleeping, feelings of sadness, crying episodes, and self-harm (Cox et al., 1987).

PHQ-9

It is a tool for assessing major depressive disorders in primary care settings. It is based on DSM-IV for diagnosis of major depressive disorder. It is a nine-item questionnaire, with responses on a Likert scale ranging from 0 (not at all) to 3 (nearly every day). The symptoms assessed by PHQ-9 are anhedonia, low mood and hopelessness, insomnia or hypersomnia, fatigue, poor self-esteem, lack of concentration, psychomotor retardation or agitation, and suicidal ideation (Spitzer et al., 1999).

Treatment for perinatal depression

The treatment of perinatal depression lies in two domains: pharmacological interventions and non-pharmacological interventions. In a health technology assessment (HTA) showing the clinical effectiveness, cost-effectiveness, safety, and acceptability of preventing postnatal depression in the UK, Morrel et al. assessed outcomes in three groups of women: 1) all pregnant women (universal group), 2) pregnant women at risk of PND because of social factors (selective group), and 3) pregnant women at risk of developing PND because of psychological risk factors (indicated group). In terms of clinical effectiveness, midwifery-redesigned postnatal care, person-centered approach (PCA), and cognitive behavioral therapy (CBT) interventions were most beneficial to the universal group. In selective groups, interpersonal psychotherapy (IPT) based intervention and education on preparing for parenting were shown to be beneficial. Promoting parent-infant interaction, peer support, and IPT-based, CBT-based, and PCA-based interventions was beneficial for indicated groups (Morrell et al., 2016). In the same HTA, midwifery-redesigned postnatal care was shown to

be cost-effective for the universal groups, education on preparing for parenting for the selective groups, and PCA-based interventions for the indicated groups (Morrell et al., 2016).

A systematic review of the prevention of common mental disorders among women in the perinatal period found that the majority of interventions on theoretical orientations were cognitive behavioral and psychoeducational, followed by mindfulness, social support, and interpersonal therapy (A. Waqas et al., 2022). WHO Guide for Integration of Perinatal Mental Health in Maternal Health in Maternal and Child Health Services lists the major evidence-based treatment techniques that could be used to treat PND (World Health Organization, 2022a). These are as follows:

- Behavioral activation is a psychological treatment to improve mood by re-engaging in the activities that used to be enjoyed; it is also part of the CBT.
- Relaxation training is related to training in techniques such as breathing exercises to bring a relaxation response.
- Problem-solving treatment is a psychological treatment that involves systematic problem identification and problem-solving techniques.
- Interpersonal therapy is a psychological treatment linking depressive symptoms and interpersonal problems.
- Cognitive behavioral therapy is a psychological treatment comprising a cognitive component (thinking differently) and a behavioral component (doing things differently).
- Parenting skills training is a group treatment that changes caregiving behavior and strengthens caregiving strategies.

WHO has developed brief psychological intervention manuals for depression that lay therapists may deliver to individuals and groups. The Thinking Healthy manual covers the use of CBT recommended by the mhGAP program, which community health workers could deliver for perinatal depression. The Thinking Healthy approach offers a well-defined structure and emphasis on addressing practical health and psychosocial issues commonly encountered by mothers during the perinatal period (World Health Organization, 2015). The intervention that is used in the study as a scenario analysis is discussed further.

Tailored Thinking Healthy/(THHP)

Compared to the original THP, adaptions in the content and delivery mechanism were made to make it deliverable by the peers (THHP). Fuhr et al. narrowed the focus from CBT to behavioral activation and reduced the number of sessions. THPP was delivered in four phases with six to 14 individual sessions. The first phase is the prenatal phase, in which intervention is delivered during the second or third trimester of pregnancy in one to six sessions; second phase early infancy, in which intervention was delivered during two months after childbirth in one to four sessions; third phase middle infancy, in which intervention was delivered 3-4 months after childbirth in two sessions; fourth phase late infancy in which intervention was delivered 5-6 months after childbirth in two sessions. In THHP India, treatment completion was defined as receiving at least six sessions, with one in each phase and recovery (defined as a PHQ-9 score less than five at three months and six months) (Fuhr et al., 2019).

Peers(Sakhis) in the THHP were laywomen (i.e., without mental health training) who had shown interest in helping and supporting other women within their communities. In this study, we choose FCHVs as the intervention delivery agent; there might be a slight advantage with choosing FCHVs as they have basic training of 18 days as per government protocol while recruiting and revision/update training focusing on advocating healthy behaviors of mothers and community people to promote safe motherhood, child health, family planning, and other community-based health promotion and service delivery (Ministry of Health and Population [Nepal], 2023). Despite the advantage of basic training, there is a lack of training in identifying and helping women with mental health issues within the training module. This makes the Sakhis and FCHVs similar in the mental health training context.

Health system of Nepal

The Ministry of Health and Population manages Nepal's healthcare system; the structure can be visualized at three levels: federal, provincial, and local (Adhikari, 2023). Nepal has lower health allocations in terms of gross domestic product (5.75% of GDP for 2022/23) than the global average of 10% GDP (NHRC, 2022; UNICEF et al., 2021), which was a decrease from the previous year's 7.46% in 2021/22. Regarding budget, the central government has the highest health expenditure, followed by the local and provincial levels (73.9%, 20.9%, and 5.1%, respectively) (UNICEF et al., 2021). Prior to COVID, the share of donors in financing public health expenditure was in decline over the years, with 20.6% in 2019/20, but with COVID, the reliance on the donor rose significantly in 2021/22 to 54.6%, which shows the reliance over donors on the events of emergency such as COVID.

Out-of-pocket spending is the source of health expenditure, and out-of-pocket spending has been a crucial part of the proportion of current health expenditure. In 2020, out-of-pocket spending as a percentage of Current Health Expenditure (CHE) was 54.2%, and government health spending as a proportion of CHE was 30.1% (World Health Organization, 2023).

Theoretical framework

This section outlines the theoretical framework for economic evaluation in healthcare, encompassing the concepts of economic assessment, health-economic modeling, and decision-making in uncertain conditions.

Economic evaluation

Economic evaluation in healthcare compares the alternative options in terms of their costs and consequences. It is a technique utilized to distribute limited healthcare resources, acknowledging the constraints on healthcare budgets and the necessity to decide on resource allocation. It is grounded in two economic principles: the Extra-Welfarist approach, which aims to maximize results with minimal resources, and the notion of Pareto improvements, which suggests that decisions should improve the well-being of at least one person without harming anyone else. Cost-benefit analysis is a form of economic evaluation that is based on the concept of potential Pareto improvements. In cost-benefit analysis, the full range of health and other consequences of a policy change is compared with resource costs as a form of compensation test. Cost-effectiveness analysis is a form of economic evaluation that is based on the non-welfarist perspective. These non-welfarist perspectives use an exogenously defined societal objective and budget constraint for health care (Briggs et al., 2006).

The input parameters for economic evaluations are costs and consequences, compared in a strict incremental way among the alternatives, which is a prerequisite for economic evaluation. In applied research in health, the measures of health outcome are viewed in a wide range, such as true positive case detected, disease condition averted, percentage reduction of risk factors, and cases treated; however, a more generic measure is predominant in the field, quality-adjusted life years (QALYs) is the most frequently used measure of health outcome (Briggs et al., 2006; Drummond, 2015).

Cost-utility analysis

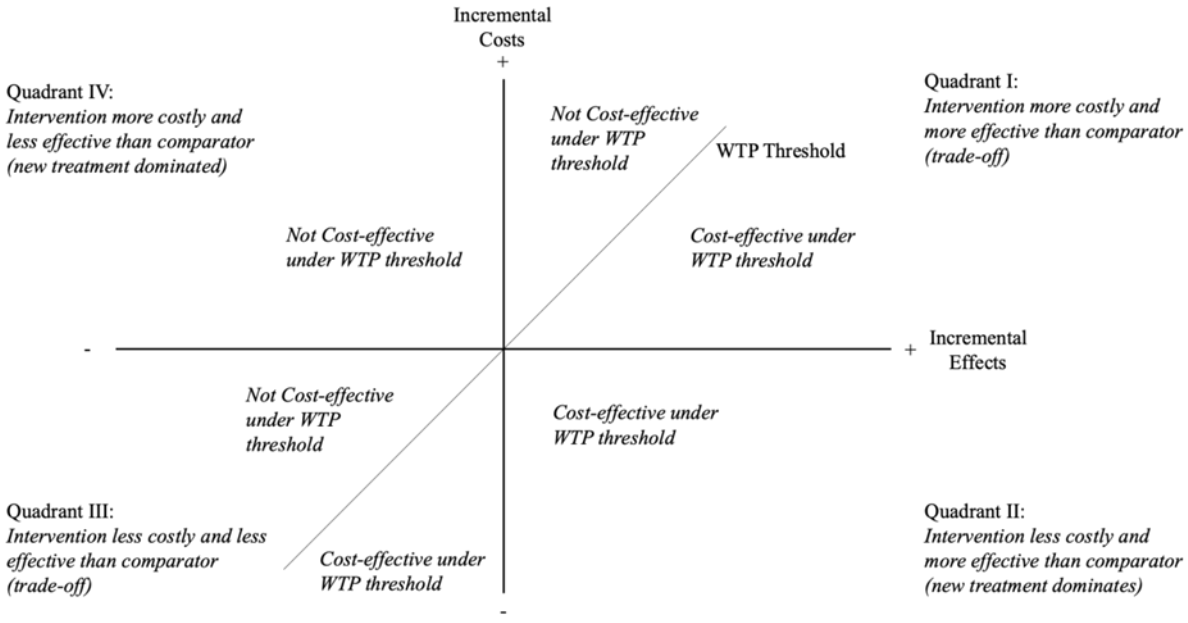
Cost-utility analysis (CUA) is a specific type of economic evaluation used to compare the outcomes of different interventions in generic terms. In economic evaluation, different generic measurement units exist through which the “utility” or the “health gain” is quantified. For this purpose, cost-utility analyses frequently make use of DALYs (Disability-adjusted life years), HYEs (Healthy-Years Equivalent), or QALYs (Quality-Adjusted Life Years). QALYs are the most frequently used measure in the literature. QALYs reflect two aspects of healthcare intervention: the impact on individuals’ length of life and health-related quality of life (Briggs et al., 2006).

In the CUA approach, the incremental cost-effectiveness ratio (ICER) is calculated to assess the cost-effectiveness of the intervention. ICER is defined as incremental cost (one intervention relative to another) to the incremental effect (in the same order as the cost). The ICER quantifies the cost per health outcome gained when comparing two treatment options and aids in identifying the most economically efficient choice. The ICER is compared with the willingness to pay(WTP) threshold per health outcome. The Willingness-to-Pay Threshold (WTP Threshold) represents the maximum amount that society is willing to pay for an additional health outcome (Briggs et al., 2006; Drummond, 2015).

$$\text{ICER} = \frac{Incremental\ cost\ (costA - costB)}{Incremental\ outcome\ (effectA - effectB)} = \frac{\Delta C}{\Delta E}$$

The results from the ICER can be viewed in a cost-effectiveness plane.

Figure 41 *Cost-effectiveness plane*



Note. Adapted from: Briggs, A., Claxton, K., & Sculpher, M. J. (2006). *Decision modelling for health economic evaluation* (Reprint 2011. ed.). Oxford University Press.
WTP = Willingness to Pay

If the ICER lies in quadrant II, the intervention is less costly and produces more health outcomes compared to the comparator, so it is cost-effective given any threshold. If the ICER lies within quadrant IV, the intervention is more costly and produces fewer health outcomes compared to the comparator, so it is not cost-effective given any WTP threshold. For the ICER lying in quadrant I and quadrant IV, there is an increment in the cost and a health outcome and a decrement in the cost and health outcome, respectively. In these quadrants,

WTP plays a role in defining what is cost-effective; if the ICER is below the WTP threshold, it is cost-effective. One of the challenges of viewing cost-effective in terms of ICER arises from these quadrants, as the ICER value is calculated by dividing positive health gain by positive health outcome, gives a positive numerical ICER as well as in the quadrant III negative cost gain divided by the negative health outcome gained gives positive numerical ICER. So, only with the ICER value it is not possible to distinguish whether that was a gain or reduction in both cost and health outcomes.

The concept of net monetary benefit(NMB) comes into play to tackle the issue with the ICER value alone, as discussed above, which provides us with a monetary value that can be presented to stakeholders. The NMB is calculated as:

NMB = WTP threshold*ΔE – ΔC

When comparing multiple options, calculating the incremental net monetary benefit and the value ranking of the mutually exclusive treatment options can give us an optimal choice as we would like to maximize the incremental net monetary benefit (iNMB).

In economic evaluation, defining the perspective of the evaluation is necessary to inform the stakeholders. The cost and health gains in the societal perspective include the wider scope of the costs and the health benefits in the society, not only being concerned with the direct health care costs and direct health benefits as in the healthcare perspective. For instance, considering productivity costs as an element in the analysis presents a comprehensive view of an intervention's impact on society. However, this inclusion may inadvertently introduce a discriminatory aspect into the model, potentially favoring employed individuals over those who are not working, depending on how costs are factored in. Policymakers should exercise caution during decision-making processes, carefully scrutinizing the selected input cost and utility parameters. There is a risk of under-representation of certain demographics, such as women, seniors, or neonates, particularly when productivity costs are taken into account. This issue becomes particularly pertinent in developing countries where, for example, women are less likely to be engaged in formal employment (Briggs et al., 2006; Drummond, 2015).

Health economic decision modeling:

Decision modeling is a tool to support economic evaluation and get to an established framework to inform decision-making under conditions of uncertainty. Decision modeling allows for variability and uncertainty with all decisions. Decision-making in the face of uncertainty is a crucial consideration in economic evaluation, as there is often uncertainty

regarding the costs and outcomes of different treatment options. Health-economic models can help explore and identify the most appropriate course of action in diverse scenarios.

Models can be categorized along two main dimensions. Firstly, decision analytic models may be grounded either in the interests of the average patient within a cohort, referred to as a cohort model, or in individual-level data, taking into account the distinctive characteristics of each patient, known as an individual sampling model. Secondly, models can be formulated as (health) state transition models or as more adaptable dynamic transition models (Briggs et al., 2006; Drummond, 2015).

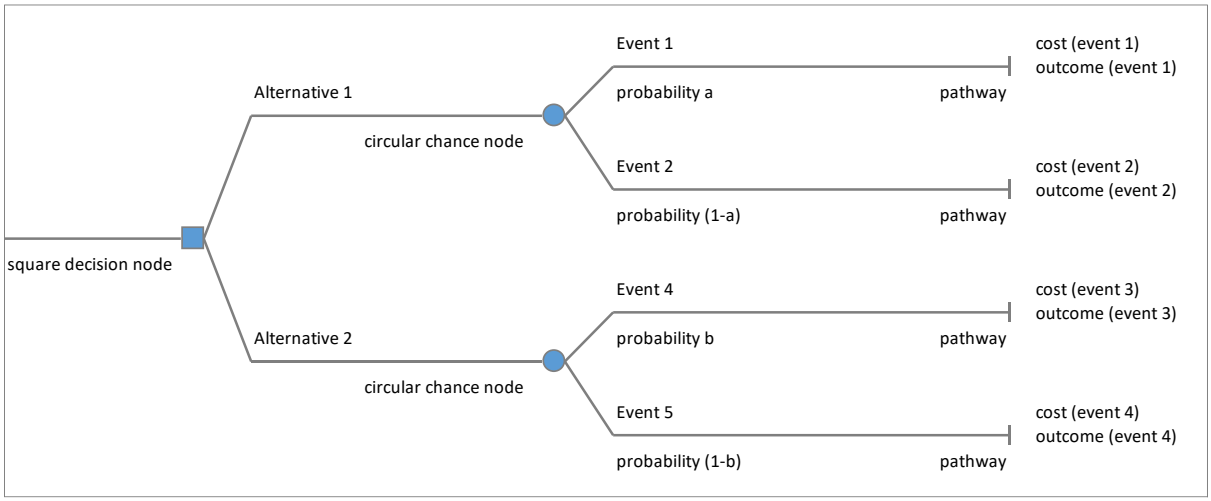
Markov models and decision tree models are two commonly used types of health-economic models that support decision-making in the face of uncertainty. Markov models are particularly valuable for modeling longer time periods, capturing time-dependent factors in a structured manner. Decision tree models are simplified Markov models that involve different branches/pathways representing treatment options and possible treatment prognoses (Briggs et al., 2006).

Decision tree model

Decision tree models are appropriate for modeling interventions with short-term costs and consequences.

- Square decision node: The decision node represents a tree's starting point and decision-making point among different treatment options.
- Circular chance node: A chance node represents a point where different alternative events might happen. Branches, according to the number of alternatives, emerge from the node.
- Pathways: Pathways represent the different and mutually exclusive journeys through the decision tree.
- Probabilities: Probabilities represent the likeliness of an event to occur on a cohort level. Navigating through the tree, decision tree models use conditional probabilities, i.e., probabilities based on the probability of earlier events having occurred.

Figure 42 *Decision tree model*



To ascertain the probability and subsequent impact of a particular pathway, the probabilities along that path are multiplied and expected costs and effects (payoffs) are attributed to each pathway of a treatment alternative. Furthermore, the outcomes of different pathways are aggregated to calculate the overall impact of a treatment option. Decision tree models can be adapted in appearance, and particularly for sequential screening and diagnostic scenarios, more sophisticated disease-based approaches can serve as frameworks for modeling conditional probabilities and multiple tests in a structured manner. Additionally, there are instances where combining specific types of models, such as decision trees and Markov models, can be beneficial. This integration enhances decision tree models by providing a more extended-term perspective and allows for a more detailed representation of treatment options (Briggs et al., 2006).

Uncertainty

Economic evaluation models often encounter various forms of uncertainty stemming from sources such as variability, parameter uncertainty, decision uncertainty, and heterogeneity (Briggs et al., 2006). This uncertainty can arise from a variety of sources, including:

- Variability (or first-order uncertainty) refers to the inherent variation among patients in a cohort or population, such as differences in clinical events. This type of uncertainty is unavoidable in cohort models.
- Parameter uncertainty (or second-order uncertainty) occurs when the input parameters to a model are estimated from data, such as randomized controlled trials or observational studies. The imprecision of these estimates can lead to uncertainty in the model's output.

- Decision uncertainty (or structural uncertainty) arises from the uncertainty of the decision being made based on the model's output.
- Heterogeneity refers to uncertainty specific to certain patient characteristics, such as the risk of adverse pregnancy events in patients with different attributes (e.g., age).

Health-economic models can partially address these uncertainties, with Probabilistic Sensitivity Analysis (PSA) being a common technique. PSA entails assigning statistical distributions to input parameters and varying all parameters simultaneously to generate a distribution of potential model outcomes. This enables the visualization of uncertainty in economic evaluations, for instance, through incremental cost-effectiveness scatterplots, cost-effectiveness acceptability curves (CEACs), and cost-effectiveness acceptability frontiers. CEAC and CEAF were produced using a net-benefit approach based on the simulations of cost-outcome data from PSA (Briggs et al., 2006). CEAC provides the probability that an intervention is cost-effective according to increasing threshold values. CEAF provides the probability that optimal intervention is cost-effective according to increasing threshold values. The highest expected net benefit determines the optimal option, whereas CEAC represents the proportion of iterations in which each option had the highest net benefit. By utilizing PSA and other techniques to address uncertainty in health-economic modeling, decision-makers can make more informed choices regarding healthcare resource allocation (Briggs et al., 2006).

The final section delves into the Value of Information Analysis (VOI) concept and its utility in quantifying the value of additional research in health-economic modeling.

Value of information (VOI)

VOI is a method for evaluating the value of additional research to reduce uncertainty in decision-analytic models. It centers on the notion that enhanced information on model parameters leads to improved decision-making.

The initial step in computing VOI is to ascertain the Expected Value of Perfect Information (EVPI), representing the maximum sum worthwhile for conducting further research to eliminate all uncertainties in the model. EVPI is computed by subtracting the expected net benefit of the best decision under the current level of uncertainty from the expected net benefit of the best decision under perfect information (Briggs et al., 2006). The maximum net health benefit of different interventions (j) and its range of values of uncertain parameters (θ). EVPI is calculated as indicated by subtracting the average expected net benefit (NB) under

the current evidence from the maximum (max) expected net benefit (NB) in each iteration under perfect information.

$$EVPI = E_\theta \max(j)\, NB(j, \theta) - \max(j)\, E_\theta\, NB(j, \theta)$$

If EVPI surpasses the expected cost of additional research, it justifies the research investment. If not, the research may not be cost-effective.

Understanding the EVPI curve, EVPI increases up to a local maximum at the point where the threshold value is equal to the value of the ICER for the adoption of the alternative approach relative to the comparator. Up to this value, EVPI is increasing. This is because the uncertainty surrounding the adoption decision is increasing (error probability increasing), as is the value applied to the consequences of making an incorrect decision. After this point, the uncertainty surrounding the adoption decision begins to fall, and the consequences associated with making an incorrect decision continue to rise. The overall EVPI depends on the interaction between these terms (error probability and the consequence of making an incorrect decision).

EVPPI and EVSI: If EVPI justifies the research, the next step involves calculating the Expected Value of Perfect Parameter Information (EVPPI) and the Expected Value of Sample Information (EVSI). EVPPI is the maximum amount justifiable for further research to reduce uncertainty about a specific parameter in the model, while EVSI is the maximum amount worthwhile to spend on collecting data to reduce uncertainty about a specific parameter in the model (Briggs et al., 2006).

Relevant topics

Screening:

Sensitivity and specificity are commonly referred to as representing the precision of the screening/diagnostic tests.

$$Sensitivity = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)}$$

Test sensitivity states the correct identification of patients with a disease. For sensitivity, the absolute values of true positives and false negatives are introduced: true positive(TP) = the number of disease-positive patients who are correctly identified as positive by a test;false negative (FN) = the number of disease-positive patients, which are wrongly categorized as negative by the screening/diagnostic test.

$$Specificity = \frac{True\ Negatives\ (TN)}{True\ Negatives\ (TN) + False\ Positives\ (FP)}$$

On the other hand, test specificity expresses the share of correctly identified negative patients. The absolute values of true negative and false positive tested patients are required to calculate the test specificity.

References

Adhikari, B. (2023). *Organogram of Health system in Federal context of Nepal*. Public Health Globe. Retrieved 1/11/2023 from https://publichealthglobe.com/structure-of-health-system-in-federal-context-of-nepalhealth-system-in-nepal/

Bauer, A., Knapp, M., & Parsonage, M. (2016). Lifetime costs of perinatal anxiety and depression. *Journal of Affective Disorders*, *192*, 83-90. https://doi.org/https://doi.org/10.1016/j.jad.2015.12.005

Briggs, A., Claxton, K., & Sculpher, M. J. (2006). *Decision modelling for health economic evaluation* (Reprint 2011. ed.). Oxford University Press.

Cox, E. (2021). Perinatal Patients with Symptoms of Depression. In (pp. 133-138). Switzerland: Springer International Publishing AG. https://doi.org/10.1007/978-3-030-71497-0_11

Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry*, *150*(6), 782-786. https://doi.org/10.1192/bjp.150.6.782

Drummond, M. (2015). *Methods for the economic evaluation of health care programmes* (Fourth edition. ed.). Oxford University Press.

El-Den, S., Pham, L., Anderson, I., Yang, S., Moles, R. J., O'Reilly, C. L., Boyce, P., Raine, K. H., & Raynes-Greenow, C. (2022). *Perinatal depression screening: a systematic review of recommendations from member countries of the Organisation for Economic Co-operation and Development (OECD)* [871-893]. Wien ; New York :.

Fuhr, D. C., Weobong, B., Lazarus, A., Vanobberghen, F., Weiss, H. A., Singla, D. R., Tabana, H., Afonso, E., De Sa, A., D'Souza, E., Joshi, A., Korgaonkar, P., Krishna, R., Price, L. N., Rahman, A., & Patel, V. (2019). Delivering the Thinking Healthy Programme for perinatal depression through peers: an individually randomised controlled trial in India. *Lancet Psychiatry*, *6*(2), 115-127. https://doi.org/10.1016/S2215-0366(18)30466-8

Joshi, D., Shrestha, S., & Shrestha, N. (2019). Understanding the antepartum depressive symptoms and its risk factors among the pregnant women visiting public health facilities of Nepal. *PLoS One*, *14*(4), Article e0214992. https://doi.org/10.1371/journal.pone.0214992

Meltzer-Brody, S., & Rubinow, D. (2021). An Overview of Perinatal Mood and Anxiety Disorders: Epidemiology and Etiology. In E. Cox (Ed.), *Women's Mood Disorders: A Clinician's Guide to Perinatal Psychiatry* (pp. 5-16). Springer International Publishing. https://doi.org/10.1007/978-3-030-71497-0_2

Ministry of Health and Population [Nepal]. (2023). *Female Community Health Programme*. Ministry of Health and population [Nepal]. Retrieved 26/10/2023 from https://mohp.gov.np/program/female-community-health-programme/en

Morrell, J., Sutcliffe, P., Booth, A., Stevens, J., Scope, A., Stevenson, M., Harvey, R., Bessey, A., Cantrell, A., Dennis, C.-L., Ren, S., Ragonesi, M., Barkham, M., Churchill, D., Henshaw, C., Newstead, J., Slade, P., Spiby, H., & Stewart-Brown, S. (2016). A systematic review, evidence synthesis and meta-analysis of quantitative and qualitative studies evaluating the clinical effectiveness, the cost-effectiveness, safety and acceptability of interventions to prevent postnatal depression. *Health Technology Assessment*, *20*, 1-414. https://doi.org/10.3310/hta20370

Motrico, E., Moreno-Peral, P., Uriko, K., Hancheva, C., Brekalo, M., Ajaz, E., Apter, G., Bramante, A., Conejo-Cerón, S., Christoforou, A., Dikmen-Yildiz, P., Evagorou, O., Fonseca, A., Lupattelli, A., Radoš, S. N., al Maach, N., Roiguez-Muñoz, M. F., Žutić, M., & Lambregtse - van den Berg, M. P. (2022). Clinical practice guidelines with recommendations for peripartum depression: A European systematic review. *Acta Psychiatr Scand*, *146*(4), 325-339. https://doi.org/10.1111/acps.13478

NHRC. (2022). *Towards Universal Health Coverage: Addressing Financial Hardship and Improving Access to Healthcare in Nepal (Policy brief)*. Nepal Health Research Council.

https://nhrc.gov.np/wp-content/uploads/2023/04/Policy-brief-Towards-Universal-Health-Coverage.pdf
Riecher-Rössler, A., & Steiner, M. (2005). *Perinatal stress, mood and anxiety disorders : from bench to bedside* (Vol. No. 173). Karger.
Runkle, J. D., Risley, K., Roy, M., & Sugg, M. M. (2023). *Association Between Perinatal Mental Health and Pregnancy and Neonatal Complications: A Retrospective Birth Cohort Study*. Washington, D.C. :.
Shrestha, S., Adachi, K., Petrini, M. A., & Shrestha, S. (2014). Factors associated with post-natal anxiety among primiparous mothers in Nepal. *International Nursing Review*, *61*(3), 427-434. https://doi.org/10.1111/inr.12118
Slomian, J., Honvo, G., Emonts, P., Reginster, J.-Y., & Bruyère, O. (2019). Consequences of maternal postpartum depression: A systematic review of maternal and infant outcomes. *Womens Health (Lond)*, *15*, 1745506519844044-1745506519844044. https://doi.org/10.1177/1745506519844044
Spitzer, R. L., Kroenke, K., Williams, J. B. W., & and the Patient Health Questionnaire Primary Care Study, G. (1999). Validation and Utility of a Self-report Version of PRIME-MD: The PHQ Primary Care Study. *JAMA : the journal of the American Medical Association*, *282*(18), 1737-1744. https://doi.org/10.1001/jama.282.18.1737
UNICEF, Nepal CO, & SPEE. (2021). *Health Budget: FY 2021/22* (PF4C - Budget Brief Updates No. 3). https://www.unicef.org/nepal/media/14426/file/Budget%20Brief%20-%202021-22%20-%20Health.pdf
Waqas, A., Koukab, A., Meraj, H., Dua, T., Chowdhary, N., Fatima, B., & Rahman, A. (2022). Screening programs for common maternal mental health disorders among perinatal women: report of the systematic review of evidence. *BMC Psychiatry*, *22*(1), 54. https://doi.org/10.1186/s12888-022-03694-9
Waqas, A., Zafar, S. W., Meraj, H., Tariq, M., Naveed, S., Fatima, B., Chowdhary, N., Dua, T., & Rahman, A. (2022). Prevention of common mental disorders among women in the perinatal period: a critical mixed-methods review and meta-analysis. *Glob Ment Health (Camb)*, *9*, 157-172. https://doi.org/10.1017/gmh.2022.17
World Health Organization. (2015). *Thinking healthy: a manual for psychosocial management of perinatal depression, WHO generic field-trial version 1.0, 2015* (9754004110).
World Health Organization. (2022a). *Guide for integration of perinatal mental health in maternal and child health services*. World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/362880/9789240057142-eng.pdf?sequence=1
World Health Organization. (2022b). *WHO recommendations on maternal and newborn care for a positive postnatal experience*. World Health Organization.
World Health Organization. (2023). *Health Expenditue Profile Nepal* Global Health Expenditure Database. https://apps.who.int/nha/database/country_profile/Index/en