UNIVERSITY *of the*
WESTERN CAPE

MEASURING SUBSTANCE USE IN PREGNANT WOMEN: A SYSTEMATIC REVIEW

NYANISA GQWEDE

A full thesis submitted in fulfilment of the degree

Master of Arts (Psychology) in the

Department of Psychology

University of the Western Cape

Bellville

2024

Supervisor: Professor Maria Florence

Co-supervisor: Professor Sabirah Adams

ABSTRACT

Introduction: Substance use during pregnancy is a pressing global public health and clinical concern. Tools to measure substance use during pregnancy can be used to ascertain harmful use of substances and related complications for the unborn child. Past studies have identified poor rates of detection of prenatal substance use. Consequently, prenatal substance use screening is a common feature to target prevention, treatment, and best practice guidelines. However, there is little information to support the selection of appropriate substance use screening tools for use with pregnant women. **Study aim:** The study aimed to systematically review the existing empirical research on screening methods (biological or self-report) used to detect substance use in pregnant women. The objectives of the study were: to identify types of screening methods used in empirical research studies to identify substance use in pregnant women, to determine the utility of these screening methods; and to determine the limitations of these screening methods.

Study method: The review search included EBSCOhost, PubMed, Google Scholar, CENTRAL, Embase, SAGE, and JSTOR. Additional studies were identified through reference mining. The timeframe for the review was studies published between 2010 to 2021. Only studies that focused on maternal 'substance use' in general were included. Grey literature and commissioned reports were excluded from the review. Studies that included cohorts of participants using substances in general, and not specifically maternal substance use, were also excluded. The Preferred Reporting Items for Systematic review and Meta-Analysis Protocols (PRISMA) Statement Checklist and the PRISMA-P (for Protocols) were employed in the study. The study used the Evaluation Tool for Quantitative Research Studies for study appraisals. Permission to conduct the study was obtained from the Humanities and Social Science Research Ethics Committee of the University of the Western Cape. Ethics considerations in terms of the systematic review process, including reporting, were adhered to.

Results: It was found that fifteen studies met the inclusion criteria for the review. All articles were screened by three reviewers who independently reviewed the study title, abstract, and eligible full-text articles. Of the fifteen articles evaluated, nine screening tools were identified. Only five of these screening tools have been found used specifically with pregnant women, these are the: Tolerance, Annoyance, Cut down, and Eye-Opener (T-ACE); Tolerance, Worried, Eye-openers, Amnesia, Cut down (TWEAK); Alcohol Use Disorder Identification Test (AUDIT); NIDA Quick Screen, Substance Use Risk Profile – Pregnancy (SURP-P); and the Car, Relax, Alone, Forget, Friends, Trouble (CRAFFT). The study findings highlight the importance of biological and self-report screening methods to gain a more holistic understanding of the mother's contextual realities impacting on her substance use. Further, the study findings imply that measuring substance use of pregnant women will help to address mitigation and prevention efforts.

Keywords: maternal substance use; systematic review; pregnancy; biological assessments; drug toxicology tests; meconium; screening tools; self-report; Evaluation Tool for Quantitative Studies; PRISMA

ii

ACKNOWLEDGEMENTS

DECLARATION

I hereby declare that the thesis entitled *Measuring Substance Use in Pregnant Women: A systematic review* is my own work, that it has not been submitted before for any degree at any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Student Signature: *Nyanisa Gqwede*

<h1 style="text-align:center">Contents</h1>

Chapter 1: Background and Rationale

Substance use during pregnancy is a pressing global public health and clinical concern (Coleman-Cowger et al., 2019). Several health bodies, such as the American Congress of Obstetricians and Gynaecologists, the Centres for Disease Control and Prevention, and the World Health Organization (WHO), advocate for the screening of prenatal substance use as crucial to determine substance use during pregnancy and subsequent pregnancy outcomes. This is essential as several risks are associated with prenatal substance use, not only for the mothers' health, but also for the health of the unborn child. This includes numerous health and social challenges, such as stillbirth, miscarriage, low birth weight, prematurity, physical abnormalities, and neurological damage (WHO, 2014). Prenatal substance use is further exacerbated by contextual and socio-historical factors such as poverty, psychiatric disorders, a history of child sexual abuse, and domestic abuse in current relationships (Lander et al., 2013).

A focus on screening or measuring substance use during pregnancy is fundamental to strengthening the first 1000 days of life, 'The critical window to ensure that children survive and thrive' (United Nations Children's Fund [UNICEF] South Africa, 2017, p.1). The first 1000 days of a child's life denotes the time from pregnancy up to the child's second birthday, which is a period of developing one's optimal health trajectory. The mother's appropriate nutrition and care during this time will impact the child's survival, as well as the mother's immediate and long-term health, stability, and prosperity.

Haefele (2012) defines maternal substance use as the use of alcohol and/or other substances during pregnancy. Substance use can develop into dependence through the use of illicit and prescription drugs consistently over time. It is characterised by a loss of control and increased use of the substance by an individual. According to the International Classification

of Diseases, 10th Revision (ICD-10) and the American Psychiatric Association (2013) Diagnostic and Statistical Manual-V, the diagnosis of substance use disorder requires at least three of the following criteria: a strong desire to take the substance; impaired control of use; a withdrawal syndrome on ceasing or reducing use; tolerance to the effects of the substance/s; a disproportionate amount of time spent by the user obtaining, using and recovering from substance use; and continuing to use the substance/s despite the problems that occur (Degenhardt et al., 2018). Substance addiction can be defined as a neuropsychiatric disorder characterised by a recurring desire to continue taking drugs despite harmful consequences. Furthermore, this drug-seeking behaviour is associated with craving and loss of control over drug use (Zou et al., 2017).

It is important to note that there is contention around the definition of substance use and abuse, with growing consensus to shift from *abuse* to *use*. In the past, addiction was largely seen as a moral failure in willpower. During the late 18th century, Benjamin Rush argued that addiction was a disease of the will. According to his view, addiction was seen as a moral condition induced by an addict's weakness of will. There was a significant change in the way addiction was perceived from the DSM-I to the DSM-V. The DSM-I defined addiction as a product of aberrances in personality. In the DSM-II the wording changed to dependence and focused more on psychobiological constructs, such as evidence of habitual use, or a clear sense of need for the drug. With the DSM-III, a distinction was made between 'substance dependence' and 'substance abuse', with the former characterised by physiological dependence such as tolerance and withdrawal, and again rooted in the framework that it is a psychobiological disorder rather than a problem of the personality or the mind. In the DSM-IV, factors contributing to addiction were not only psychophysiology (tolerance and withdrawal) but also cognitive. The DSM-V provides a more holistic definition of substance dependence,

7

emphasising the psychobiological changes that occur with drug abuse, which promotes a lack of cognitive control over the use of the drug (Zou et al., 2017). Furthermore, according to the American Psychology Association (2022), substance abuse denotes a particular pattern of "compulsive substance use characterised by recurrent significant social, occupational, legal, or interpersonal adverse consequences, such as repeated absences from work or school, arrests, and marital difficulties."

Therefore, it is imperative to understand pregnant women's past and current alcohol, nicotine, prescription substance use, and any illicit substance use to determine appropriate mitigation and treatment (Wong et al., 2011). This is especially relevant given the long-term consequences of prenatal substance use for women, such as an elevated risk of mortality, morbidity, and social and legal consequences (Guttmann et al., 2019; Hostage, 2020). In addition, if substance use is detected, a more complete assessment by perinatal healthcare providers is recommended to determine if there is a history of substance abuse or dependence. This can be verified using biological assessments, such as urine and hair assessments. The use of biological assessments is important for early intervention if the expecting mother has access to and visits the healthcare facility regularly. Early intervention can include the treatment of maternal and neonatal withdrawals, counselling, and referrals for outpatient/inpatient treatment (Wong et al., 2011). This review study aimed to examine the existing empirical research on the screening tools used to identify substance use in pregnant women.

Review question and objectives

The review question is to determine: What empirical research exists on screening methods (biological or self-report) used to detect substance use in pregnant women? The specific objectives are:

1. What type of screening methods are used in empirical research studies to identify substance use in pregnant women?

2. What is the utility of these screening methods?

3. Are there any limitations of the identified screening methods?

Prevalence of maternal substance use

Prevalence rates of prenatal substance use have been shown to vary globally. A critical review study by Tavella et al. (2020) aimed to establish the prevalence of illicit substance use during gestation across global empirical studies. They found that 70 studies were published on the topic between 2000 and 2019, encompassing a pooled sample of 689 553 participants across 14 countries. The average prevalence rate was 1.83% across these studies, taking into consideration the sample size per study. Marijuana was found to be the most frequently used substance by an average of 42.85% of women across the studies, followed by cocaine (14.29% of studies). Synonymous with the United Nations Office on Drugs and Crime (UNODC, 2016), marijuana was found to be the most frequently used illicit substance amongst pregnant women globally. They also found high prevalence rates among adolescent pregnant women. Regarding the accuracy or sensitivity of different methods of screening using questionnaires or interviews, Tavella et al. (2020) found a mean prevalence of illicit substance use at 1.65% for pregnant women. However, a toxicological analysis (biological matrices of various methods) provided a substantially higher prevalence of 12.28%, which is a 7.4 times higher rate. This points to the increased sensitivity of toxicological analysis when compared to questionnaires/interviews. Ondersma et al. (2019) found that some self-reporting measures were highly specific, and others were highly sensitive, but none were both. Sensitivity and specificity are terms that are used to assess a clinical test. Cowger et al. (2019) explains that sensitivity is calculated as the percentage of individuals with positive reference test and those who also have positive index

tests, and specificity is the percentage of persons with a negative reference test who also had negative index tests. While they identified studies conducted on almost every continent, prevalence rates for Africa were less forthcoming (Tavella et al., 2020).

There is little information available on the extent of substance use other than tobacco among pregnant women in low-to-middle income countries (LMIC's). Overall, the prevalence of tobacco use in these countries is estimated at 2.6%, with some countries having higher maternal tobacco use rates of up to 15%. According to the WHO, cannabis is the most common illicit drug worldwide, followed by amphetamine-type stimulants and opiates, and they are all likely to be used by women of reproductive age. The limited data available for Africa is from South Africa, and indicates that between 3.6% and 8.8% of pregnant women use illicit substances and 19.6% use alcohol. The illicit substances used most frequently in South Africa are methamphetamine and cannabis. Opiate use has also increased in places such as Africa and Asia.

Most studies in Africa have focused on specific substances, for example, alcohol and its deleterious effects on the individual and the prevalence of Fetal Alcohol Spectrum Disorder (e.g., Myers et al., 2018; Vythilingum et al., 2012). Vythilingum et al. (2012) conducted a study aimed at assessing the prevalence of alcohol and substance use in a South African antenatal population. The study further examined the relation between alcohol and substance use and sociodemographic factors, depression, and perceived stress in this sample. The study included 323 women from the Metropole East district, Cape Town. They found high rates of both alcohol abuse and antenatal depression, and a significant association between depression, substance use and alcohol abuse, and that many women using alcohol and/or substances stop their use once they discover they are pregnant.

10

Myers et al. (2018) assessed the unique contribution of hazardous alcohol use and examined its effect on infant growth outcomes. The study formed part of Drakenstein Child Health Study between 2012 and 2015 with South African women enrolled in the study, controlling for maternal psychosocial stressors and hazardous tobacco and drug use in a cohort of 986 pregnant. The study found that a sizable proportion of women reported hazardous alcohol use during pregnancy. Hazardous alcohol use was associated with infant growth restriction at birth even after controlling for maternal body mass index (BMI). Although the study used the Alcohol, Smoking and Substance Involvement Screening Test (ASSIST) to assess substance use among pregnant women, it did not focus on how the screening tool is utilised and the effectiveness as well as limitations thereof.

Cook et al. (2017) indicated that several studies conducted in South Africa have reported prevalence rates of substance use (specifically illicit substances) in pregnant women ranging from 3.6% to 8.8%, with approximately 19.6% of pregnant women using alcohol. Methamphetamine and cannabis are the most frequently used substances in South Africa (Forray, 2016). South African literature has not included a national sample resulting in underreporting of substance use in this context. Overall, there is a dearth of data on the prevalence of substance use specifically from LMIC's). For this reason, further research is necessary, particularly in LMIC's that provide prevalence rates in terms of trends and patterns of substance use during pregnancy to identify appropriate screening tools and early intervention and treatment.

An important aspect that needs to be considered is the context-specific cultural beliefs and norms that have been shown to influence prevalence (Cook et al., 2017). Francis et al. (2019) states that one of the potentially important modifiers of risky behaviours is religiosity,

11

which continues to be highly prevalent in South Africa despite rapid social change. On the other hand, Chen et al. (2018) indicates that the United States of America (USA) is also highly religious. Religiosity played an important role in the lives of South Africans during the apartheid period where churches and other places of worship provided people with an opportunity and place for both worship and for meeting to address the political challenges of the era. During the post-apartheid era, religiosity and spirituality have continued to provide solace and meaning to vast proportions of the population, and particularly, members of disenfranchised communities, which face high levels of unemployment and poverty and other social stressors, such as substance use. Further, Chen et al. (2018) found that religious involvement has also been linked to a wide range of other outcomes, such as greater psychological well-being, character strengths, reduced mental illness, and healthier behaviours. Additionally, these authors indicate that religious involvement may protect against certain behaviours and promote positive practices. These studies are, however, subject to certain limitations. There is often limited control for baseline characteristics, and reverse causation often cannot be ruled out.

Methods of testing substance use

Screening instruments and assessment of substance use in pregnancy is essential to provide accurate, reproducible, and relevant measurements (Gaya et al., 2009). The use of various screening instruments is necessary for health facilities, as they provide richer information and enable healthcare workers to tailor treatments and interventions and compare results across groups (Gaya et al., 2009). There are two main methods of detecting substance use in pregnant and non-pregnant populations, namely biological assessment and self-report evaluations (Tavella et al., 2020). Although many health care providers use biological testing to determine substance use in pregnancy, such as positive urine toxicology, this does not

12

delineate the temporality or misuse (Coleman-Cowger et al., 2019). Tavella et al. (2020) found that different biological matrices have been used in the literature to examine the toxicological analysis of substance use in pregnancy, which included hair, urine, serum, whole blood, umbilical cord, and meconium testing. However, Wong et al. (2011) maintain that drug toxicology testing is not recommended for universal screening as it has numerous challenges, which should be considered only after a comprehensive assessment and if there is a clinical indication. Urine drug screening can only detect recent substance exposure, while neonatal hair and meconium testing can assess intrauterine use as meconium and hair form in the second and third trimester, respectively (Wong et al., 2011).

These various types of biological testing are often used with self-report methods. Self-report evaluation methods can take the form of maternal interviews and/or the use of screening instruments to determine prenatal substance exposure. Maternal interviews employ open-ended, non-judgemental questioning, and have been shown to elicit disclosure of perinatal maternal substance use (Wong et al., 2011). Further, while several validated screening instruments are available for alcohol and tobacco use, there have been fewer studies that have developed and validated screening instruments to detect substance use (prescription or illicit) in pregnant women (Coleman-Cowger et al., 2019).

Here Tavella et al. (2020) make a crucial point based on findings from their critical review. They argue that the prevalence of substance use in pregnancy is underestimated when solely utilizing self-report measures, such as questionnaires and interviews in comparison to empirical research studies that use toxicological methods. Therefore, when selecting an appropriate and effective screening instrument for use among pregnant women it is important

13

to ensure that it is valid and reliable in all sub-groups who have shown variations in substance use such as 'race', age, and trimester (Coleman-Cowger et al., 2019).

As the majority of pregnant women seek and receive medical care during their pregnancy, it is an important to reach, detect, and treat those using substances. Notwithstanding this, it is important to note that many women who use substances, whether illicit, tobacco, or alcohol, during pregnancy are less likely to visit antenatal clinics, compared to non-users of illicit drugs (Arrie et al., 2006). This inability or reluctance to visit antenatal clinics can be related to several factors, including the use of substances before pregnancy, living in disadvantaged circumstances characterized by poverty, low socio-economic status, lack of access or proximity to health care clinics, or lack of awareness of treatment facilities and fear of stigma from being treated for a substance use disorder (Jones et al., 2011).

Measures to assess maternal substance use

Several screening instruments have been used to identify substance use in pregnant women (WHO, 2014). These include the:

- Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST)

- 4P's (Parent, Partner, Past, Pregnancy) Plus

- Hospital Screening Questionnaire (HSQ)

- Pregnancy Information Program (PIP)

- Substance Use Risk Profile – Pregnancy (SURP-P)

- Alcohol Use Disorder Identification Test (AUDIT)

- Alcohol Use Disorder Identification Test – Consumption (AUDIT-C)

- Cut, Annoyed, Guilty, and Eye questionnaire (CAGE)

- Normal drinker, Eye-Opener, Tolerance (NET)

14

- Short Michigan Alcohol Screening Test (SMAST)

- Ten Question Drinking History (TQDH)

- Tolerance, Annoyance, Cut down, and Eye-Opener (T-ACE)

- Car, Relax, Alone, Forget, Friends, Trouble (CRAFFT)

- Tolerance, Worried, Eye-openers, Amnesia, Cut down (TWEAK)

The ASSIST, HSQ, and PIP were developed to screen for tobacco, alcohol, and other substances, while the 4Ps Plus and SURP-P detect alcohol and other substances. The rest and the majority of tools specifically assess alcohol use (AUDIT, AUDIT-C, CAGE, NET, SMAST, TQDH, T-ACE, and TWEAK). According to Gaya (2009), the AUDIT is used to identify risky or harmful alcohol consumption as well as alcohol dependence and abuse.

A key point of concern is that not all the instruments mentioned above have been tested with samples of pregnant women. However, the AUDIT-C was specifically developed to assess alcohol consumption and has been validated for use with pregnant women. The SURP-P and 4P's Plus were found to be more ideal screening tests during pregnancy owing to its high sensitivity and negative predictive values. The SURP-P is a simple and flexible tool for identifying possible substance use in pregnancy, however, a further screen is required for identifying those who would require treatment (Coleman-Cowger et al., 2019). According to the WHO (2014), the HSQ, PIP, and TQDH have not been validated for pregnant women given insufficient information on their sensitivity and specificity. The HSQ includes 18-40 items, and is used to assess tobacco, alcohol, and substance use. The PIP consists of 200 items, and it is also used to assess tobacco, alcohol, and substance use. Whereas the TQDH is used to assess alcohol use and consists of 10 interview questions, however, it does not include specific questions about pregnant women. Furthermore, the WHO reported that the ASSIST has

15

typically been used clinically as a follow-up to positive screening results of the NIDA Quick Screen (Oga et al., 2020; WHO, 2014), with no known studies on pregnant populations.

The mode of administration of these instruments also differs, with most paper-and-pencil tests to be completed by the mother, two interviewers administered (ASSIST and TQDH), and one computer-based (the PIP). Most of the measures are brief with 10 items or less, with the PIP being the longest instrument including 200 items. While several screening tools have been developed for identifying substance use in pregnancy, there is not one specific instrument recommended for universal use as this would depend on the context in which it is used, and necessary adaptation and translation (Wong et al., 2011).

The ASSIST is a two-part screening tool that is used to measure tobacco, alcohol, and substance use, and is used only if the participant endorsed the use of prescription drugs for nonmedical reasons or illegal drugs in the past year (Coleman-Cowger et al., 2019). The ASSIST was developed by the WHO; it is recommended for use by the National Institute on Drug Abuse (NIDA) and the Substance Abuse and Mental Health Services Administration (SAMHSA). It is a screening instrument that consists of 8 items that identify lifetime use, current (past 3 months) use, and level of risk (low, moderate, and high-risk use) of tobacco, alcohol, and 7 classes of commonly used drugs. A score from 4-26 indicates moderate risk use and is classified as unhealthy use, whereas a score from 27 + indicates high-risk use (Han et al., 2017).

The 4Ps plus consists of 7 questions and is used to measure alcohol and general substance use, it has been previously validated in a sample of pregnant women. The 4Ps Plus was specifically designed to quickly identify obstetrical patients in need of in-depth substance

16

use assessment or follow-up monitoring. It provides a clinically valuable methodology for integrating screening for risk from substance use and abuse into primary prenatal care settings (Chasnoff et al., 2007; Schaubergeret et al., 2014). The instrument has demonstrated moderate high sensitivity (83%) which means a calculation of number of true positives plus 20% of false negatives and specificity (80%) which means a calculation of number of true negatives plus 20% of false positives and its goal is to identify women with risky drinking patterns prior to learning of their pregnancy (Armstrong et al., 2010).

The SURP-P is used to measure alcohol and substance use and consists of three items. It involves classifying the number of alcoholic drinks consumed months before pregnancy (Coleman-Cowger et al., 2019). The SURP-P was developed in a sample of pregnant women and cross-validated with a sample of pregnant women using the high-risk cut-off. Its validation on alcohol use in terms of sensitivity was estimated at 48% and specificity at 85%, and for cannabis use with a sensitivity of 68% and specificity of 86% (Ondersma et al., 2019).

The AUDIT is a 10-item self-reported screening tool that was developed based on the results of a cross-national study by the WHO. to detect patients with patterns of hazardous and harmful consumption of alcohol (Seib et al., 2012). According to Gaya (2009), the AUDIT is used to identify risky or harmful alcohol consumption, as well as alcohol dependence and abuse. It comprises 10 items with scoring from 0-7 which indicates low risk, 8-15 which indicates increasing risk, 16-19 which indicates higher risk and 20 or more indicates possible dependence and the AUDIT is administered by healthcare workers. According to Wendy et al. (2019), the AUDIT is both sensitive and specific in discriminating alcoholic patients from non-alcoholic patients. Patients who scored 8 or higher in the AUDIT have been found to experience a greater

17

amount of alcohol and related medical disorders and are hospitalized within 2-3 years of the initial screening.

The AUDIT-C is the short version of the AUDIT screening instrument, which includes 3 of the 10 AUDIT items. It was introduced in June 2008 into public antenatal services throughout Queensland, Australia. As with the AUDIT, it is used to identify hazardous alcohol drinkers, with scores ranging from 0-12 (Macleod et al., 2021). The purpose of the AUDIT-C was to allow consistency in data collection across various antenatal settings, and to accurately identify substance use in pregnant women to allow opportunities for early intervention. It is an effective screening tool to identify harmful or hazardous drinking in the primary care setting. Because of its brief nature, it enables clinicians to accurately detect patients who consume alcohol at hazardous levels (Seib et al., 2012).

The TWEAK is a five-item scale developed to screen for risky drinking during pregnancy. Two points are given if a woman reports that she requires 3 or more drinks to feel the effect of alcohol, the worry question is also given 2 points for a positive response, and if the last 3 questions are positive it is given 1 point. On a scale of 0-7 if a woman scores a total of >2 points she is likely to be a risky drinker (Sarkar et al., 2010). The CAGE is used to detect problem drinking. It asks about alcohol problems ever experienced and is feasible to use and easily applied in clinical practice owing to its brevity (Choe et al., 2019; Tricia et al., 2017).

The T-Tolerance, A-annoyance, Cut down, and Eye-Opener (T-ACE) is a four-item screening tool that includes the screening of alcohol use during pregnancy (Sarkar et al., 2010). The T-ACE is a measurement tool that consists of four questions that identify the risk of alcohol

18

intake for sufficient potential damage to the foetus. It is completed at intake; its score has a range of 0-5 and the value of each answer to the four questions is totalled to determine the final score. T-ACE is based on the CAGE, but it was modified to improve the identification of risky drinking during pregnancy. This screening tool identifies 90% of potential risky drinkers. T-ACE substitutes the guilt question from the CAGE with the question about tolerance to the effects of alcohol. It was considered positive with any score >0 and was combined with other risk factors such as the 'regular use of two or more over the counter drugs', 'consumption of a large amount of coffee', and use of alcohol to fall asleep to determine alcohol drinking at least 15 days to months (Chang et al., 2010).

In a study conducted by Yonkers et al. (2010), the T-ACE and the TWEAK were used. The study aimed to report on the development of a questionnaire to screen for hazardous use in pregnant women and compare the performance of the screening tool for other drugs and alcohol measures. The sample included pregnant women who presented for obstetric care at the hospital-based reproductive health clinics, with N = 2,684 participants. The findings of this study illustrate that the substance use risk profile pregnancy scale algorithm for a low-risk population was mostly highly predictive of substance use validation and with high sensitivity of 91% and adequate specificity of 67%. The study, however, also found that the SURP-P (for high-risk populations) had lower sensitivity of 57% and higher specificity at 88%, this basically means that SURP-P had a higher negative reference test for high-risk populations, which basically means that SURP-P tests showed that there were more true negative tests in comparison to the false positives. These instruments are known for performing acceptably in testing alcohol use in pregnant women. The other screening tools applied in the study were four questionnaire items from the Addiction Severity Index, and two questions about domestic

violence (4Ps and 4Ps plus) used to detect pregnant women with lower levels of alcohol and drug use.

Coleman-Cowger et al. (2019) found that the SURP-P and 4P's Plus performed similarly with more than 90% sensitivity and more than 85% negative predictive values which makes them ideal screening tests, rather than the NIDA Quick Screen-ASSIST, which had a lower sensitivity with a similar negative predictive value. Chang et al. (2019) conducted a study to validate five self-report, non-proprietary questionnaires in the identification of substance use among pregnant women. The study included a total of 1220 pregnant women from outpatient prenatal care. This study found that over 15% met diagnostic criteria for substance use disorder. The study suggested that the CRAFFT and SURP-P showed the modest ability to identify substance use disorders among pregnant women with more than 95% sensitivity.

Ghrekin et al. (2010) looked at validating the Drug Abuse screening test (DAST-10) for use during pregnancy. The study aimed at examining the ability of the drug abuse screening test (DAST-10) to identify prenatal drug use. The participants completed computerised assessments that include the DAST. They were also asked to provide hair and urine samples that were used to assess recent drug use. The study included 300 women who had given birth at an obstetric hospital in Detroit. They found that 18% of the participants had positive hair or urine screening for marijuana, 8.5% had a positive screen for cocaine, amphetamines, or opiates, 23 and 7% tested positive for other drugs. The study suggested that hair screens detected more drug use than urine screens did. Furthermore, 12.2% were positive for any drug using hair screening and 2.5 % were positive on any drug use using urine samples. The results also suggested that underreporting of illicit drug use remains a substantial concern despite the use of anonymity. However, the DAST-10 still identified less than half of women using drugs

20

during the last trimester of pregnancy. Nearly 20% of the sample had a positive toxicology screen but denied drug use on the DAST-10. The sensitivity of the DAST was only 47%. These findings suggested that brief drug use screeners might have limited utility for pregnant women, as the efforts to facilitate disclosure via reassurance alone are unlikely to be sufficient in the utilized population.

There is a wide range of biological tissues used for screening substance use during pregnancy. Maternal tissues including blood, hair, and urine can be screened before and after birth, and fetal tissues such as blood, urine, hair, and meconium can be screened following birth. According to Taplin et al. (2014), hair and sweat testing are generally used to detect use over a period of weeks or months. Urine tests detect use during previous days or weeks, while oral fluid testing detects use within a shorter period of hours or days. Taplin et al. (2014) stated that the use of frequent urine testing, which is best conducted on a random basis, is a preferred method; but is expensive. Less expensive is hair testing, as hair can be easily harvested and analysed within a 16-hour observation window. Hair testing has important benefits that should be weighed up in terms of convenience. Regular, frequent, monitored urine testing is one of the best practice models. Both hair and urine testing are best practices and are valid, reliable methods for assessing the extent of ongoing unlawful drug use. Hair testing, however, seems to be a viable and useful tool for the management of these cases where children are put at risk by maternal substance use. However, hair testing is currently not common practice in drug treatment health programs and is mainly used in the workplace and for court purposes, within the context of the USA (Taplin et al., 2014).

Furthermore, Coleman-Cowger et al.'s (2017) study used both self-reported and biochemical measures to test for various drugs. Participants were also asked to consent for hair

21

testing that involves the cutting of approximately 100 strands of hair from the head or other body hair samples, which were then shipped to an external laboratory on the same day for drug testing using mass spectrometry (Coleman-Cowger et al., 2017). As the study used both urine and hair testing, it enabled them to examine past 90-day substance-use history with precision. The study has also conducted a direct, head-to-head comparison of multiple screening tools for prescription-drug and illicit drug use among pregnant women, while also using biological measures as a gold standard for comparison.

Williams et al. (2014) conducted a study to describe three assessments of alcohol use and other substance use. Among these assessments, were the biological urine tests, whereby the prevalence of substance use among pregnant women attending public sector antenatal birthing units in the Cape Metropole, could be estimated. The study population consisted of pregnant women who attended 11 public sector, or community-based clinics, scattered across eight health sub-districts, for antenatal care and delivery in the greater Cape Town. Urine samples were tested for alcohol, cannabis, and heroin, using drug screening called immunoassays. The tests provided a measure of the metabolites of the specific drugs for which participants were tested. Ethyl glucuronide is a direct metabolite of alcohol; its presence in urine may be used to detect recent alcohol consumption. Immunoassay urine drug screens are commonly used as they are relatively inexpensive and rapid. For this reason, standard immunoassay testing is the preferred initial test for urine drug screening (Williams et al., 2014).

Rationale for the study

Prenatal substance use is a persistent public health concern. Previous empirical studies have demonstrated that prenatal substance exposure, whether direct or indirect, is associated with increased risk factors and negative outcomes for both the mother and foetus (Polak et al.,

22

2019). These negative outcomes include a high-risk pregnancy, and adverse developmental outcomes such as congenital malformations, low birth weight, cognitive, and neurobehavioral deficits (Shogren et al., 2017). Therefore, screening for, measuring, and determining substance use during pregnancy is a crucial first step in treating and mitigating the adverse effects of maternal substance use (Chang, 2001., Hostage et al. 2020). In this regard, Hostage et al. (2020) advance 'universal screening for substance use' during pregnancy, which should include brief intervention and referral to treatment. While several health bodies advance the use of biological testing to determine maternal substance use, this method alone does not provide the context that the mother is living in, nor the temporality of substance use or evidence of 'problematic' use (Coleman-Cowger et al., 2019). Therefore, several studies have used self-report measures that assess substance use during pregnancy, with some studies using both biological testing and self-report measures (Tavella et al. 2020, Wong et al., 2011).

Despite this emerging body of research examining the prevalence of substance use during pregnancy, there is no measuring instrument that has been developed specifically for pregnant women. However, a number of instruments have been adapted for use with pregnant samples, such as the AUDIT-C, which was specifically developed to assess alcohol consumption, and has been validated for use with pregnant women. The SURP-P, 4P's, TWEAK, and ASSIST have also been validated for use with pregnant woman although they were not specifically developed for use with pregnant woman (Coleman-Cowger et al., 2019, Seib et al. 2012, Han et al. 2017), and were found to be valid for this cohort. Moreover, there is a tension in the literature regarding biological testing and self-report measures. Recent research found that while biological testing is deemed to be the 'gold standard', it too has weaknesses. For example, both urine and hair testing can produce false-positive results, and do not provide the temporality of substance use. For this reason, Coleman-Cowger et al. (2019)

maintain that self-report screening measures provide useful information, are feasible, and can provide the medical practitioner with context and build rapport with their client – both of which are essential in treating substance use. The evidence is in fact mixed regarding which measurement type is superior.

The current study therefore aims to make a key contribution in this regard, by conducting a systematic review to identify existing empirical research on the screening methods (biological or self-report) used to detect substance use in pregnant women. The study further aims to identify the effectiveness of these methods and measures and ascertain any limitations of the screening methods used to measure substance use in pregnant women. The study aligns to the 2030 Sustainable Development Goals (United Nations, 2015). Overall, the study addresses *Goal 3: Good Health and Well-Being*, by contributing to knowledge on foetal and maternal health, and in particular Target 3.5 that aims to "Strengthen the prevention and treatment of substance abuse, including narcotic drug abuse and harmful use of alcohol", and 3.7 which focuses on universal access to reproductive healthcare services. The review aims to provide an evidence-based meta-synthesis on studies using these screening methods, using a narrative synthesis approach. In this way, it engenders a comprehensive appraisal and account of available literature on the topic. The findings of the review can be employed to inform and integrate appropriate and relevant screening methods for maternal substance use. Moreover, the review can be used as a motivation to embed reproductive health and preventative interventions for substance use among women of child-bearing age, into national strategies and programmes within various country contexts, particularly in LMIC's.

Chapter 2: Method

Research design

The study utilised a systematic review design, focusing on empirical peer-reviewed literature. According to Kim et al. (2017), a systematic review refers to a specific methodology in research, developed to gather and evaluate the available evidence pertaining to a focused topic. It is therefore evidence-based research. According to Wong et al. (2019), the process of conducting a systematic review includes five steps, namely: framing of questions for review; identifying relevant studies; assessing the quality of the studies; summarising the evidence; and interpreting the findings. Framing questions for reviews involve the specification of problems that need to be addressed in a more structured format (van Mastrigt et al., 2019). The remaining four stages must be extensive, apply assessment tools, and tabulate the studies assessed and be able to interpret the findings to determine if the stages and requirements have been met (Van Mastrigt et al., 2019). The rationale for using a systematic review in the current study is to provide a methodologically rigorous opportunity to evaluate the state of research on measurement instruments that assess substance use among pregnant women.

In the current study, the PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocols) Statement Checklist and the PRISMA-P (for Protocols) were employed (Moher et al., 2015). The PRISMA tools have been widely used across disciplinary contexts, particularly within health and social sciences. The PRISMA-P (2015) Checklist was developed during the proposal phase of the study for use in the planning and framing of the review. It includes 17 items classified into three key sections, namely: administrative information, introduction, and methods (Moher et al., 2015). For the current study, the *administrative information* section delineated the study as a systematic review, the student, and supervisors' (reviewers') details and their respective roles, as well as the title of the review.

25

The *introduction* section entailed the rationale for the current review owing to a particular gap in the literature of an integrative review on measuring maternal substance use. Importantly, this section also encompassed the research question employed in the current study, namely the SPIDER (Sample, Phenomena of Interest, Design, Evaluation, and Research type), which is discussed below. The SPIDER was used as the guiding framework to develop the research question. Finally, the methods section included in the review inclusion criteria and search strategy comprises the sources used, data management, selection process, the data extraction process, risk of bias, and data synthesis that was employed.

The PRISMA Statement was developed by Moher et al. (2009). It encompasses a 27-item checklist and a flow diagram that represents the review search process (Moher et al., 2009). The main aim of the PRISMA Statement is to assist scholars in reporting systematic reviews and meta-analyses. Although this tool was developed for research on randomised controlled trials, the developers note that it can be used for systematic reviews more broadly. However, while the PRISMA Statement is helpful in the 'critical appraisal' of published reviews, it should not be used as a quality appraisal tool to assess the 'quality of a systematic review' (Moher et al., 2009). The completed PRISMA Statement checklist is included as an appendix to this review (Appendix A). Further, the reporting practice of the PRISMA Statement has been followed in this review.

The role of a search tool in any systematic review is important to consider. As noted above, the SPIDER was used to guide the development of the review question and informed the search strategy. The SPIDER has been used in other systematic reviews within the health and social sciences (see Cooke et al., 2012; Eriksen & Frandsen, 2018; Lawlis et al., 2018; Methley et al., 2014). This aligned to two of the five steps put forward by Khan et al. (2003),

26

thus: *Step 1: Framing the question* and *Step 2: Identifying relevant publications*, which we undertook to conduct the systematic review (see Khan et al., 2003). Notwithstanding the numerous search tools (or frameworks) that exist to conduct a systematic review to conceptualise and structure systematic review questions, including the PICO (Population, Patient, Problem; Intervention; Control), PEO (Population and problems, Exposure, Outcomes) and FINER (Feasibility, Interesting, Novel, Ethical, Relevant), the SPIDER was best suited to the focus of the current study. This is often used in the social sciences (Methley et al., 2014). Noting the focus of the current review, we encapsulated the review focus by addressing the focal areas of the SPIDER. This includes the sample (women), phenomenon of interest (maternal substance use), design (systematic review), evaluation (self-report and biological testing), and research type (qualitative, quantitative, and mixed methods). This is evident in the intentional broad-based nature of the review question to capture both the breadth and depth of empirical studies focusing on the topic.

Article search strategy

The article search strategy for the review included electronic resources accessed via electronic meta-databases and databases, namely EBSCOhost, PubMed, Google Scholar, CENTRAL, Embase, SAGE and JSTOR, which index peer-reviewed journal articles published in English. Given the general nature of the review question, the review intended to distil all empirical literature on measuring maternal substance use within the period of 2010 to 2021. The following keywords were used in the noted database searches: 'substance use', 'substance use measuring tool during pregnancy', 'effectiveness of substance use screening tools', and 'assessment of substance use during pregnancy' (see Appendix A for Keyword strings searched). While the initial article search was completed in December 2020. An additional search was conducted from the period of January to August 2021, to identify any recent studies

27

on the topic that were published during this time.

Inclusion and exclusion criteria

The inclusion criteria for the study encompassed studies using a quantitative approach. The key motivation for the inclusion of studies using this approach was to detect empirical research on maternal substance use screening using a standardised measuring instrument or employing biological testing methods. We included studies that focused on maternal 'substance use' in general. However, the study included empirical articles published between 2010 and 2021.

The study exclusion criteria encompassed studies that included cohorts of participants using substances in general, and not specifically maternal substance use, studies that were published outside of the specified period (2010 - 2021), and studies using a qualitative methodological framework. Qualitative studies were excluded as the focus of the current systematic review was to ascertain the quantitative screening tools that have been used to assess maternal substance use. Grey literature and commissioned reports were also excluded from the review. All relevant studies yielded from the database searches were available through the University of the Western Cape (UWC) subscription or were publicly available. Reference mining was used to access studies not yielded by initial database searches. The table below indicates the inclusion and exclusion criteria of the study.

28

Table 1

Inclusion and exclusion criteria

	Included	**Excluded**
Publication type	2010 to 2021	Non-English language
	From any country	Commissioned Reports,
	Peer-reviewed articles	Systematic Reviews,
	Published articles	Commentaries
		Grey Literature
Study population	Pregnant women	General population
Measurement method	Quantitative measurement methods or biological testing	Qualitative methods
Study design	Quantitative studies	Qualitative studies

Quality appraisal

The study used an adapted version of the Evaluation Tool for Quantitative Research Studies as the quality appraisal (see Appendix A; Long et al., 2002; Long & Godfrey, 2004). The tool was developed to assist researchers with the critical appraisal of studies, serving as a template consisting of key questions for analysing the key areas of each article (Hoosen et al., 2021; Long et al., 2002). The adapted version of the appraisal tool was revised to align to the focus of the current study. Sections of the critical appraisal tool that focused on clinical features not relevant to this systematic review were removed. The sections removed included methodological frameworks, such as comparable groups, outcome measure, time scale (quantitative), and theoretical framework (qualitative) were removed. The existing sections included in the study are the study overview, study setting, sample, ethics, and policy and practice implications. The tool was further adapted for use with mixed-methods studies. For each article included in the full-text review process, the quality appraisal tool was applied to

29

assess the quality, summarise the key points, and evaluate whether the article should be included or excluded (Long et al., 2002). As the critical appraisal of empirical research is crucial for a systematic review, it was important to consider if the quality of the study is of a high calibre, but also whether the data collected have an organisational and culturally relevant context (Long & Godfrey, 2004). Therefore, quality appraisal is pertinent in every step of the review.

Long et al. (2002; 2004) provide permission for use of the critical appraisal tools in their respective articles. They (Long & Godfrey, 2004; Long et al., 2002) contend that their quality appraisal tool focuses on the quality and rigour of each study and does not while there is contention in the literature regarding the use of scored and unscored appraisal tools for systematic reviews, the current adapted appraisal tools emphasise the quality and rigour of each study. Therefore, using this tool, reviewers are able to make evaluative comments regarding the strengths and weaknesses of each study rather than the use of scoring, to include or exclude a study. The review was conducted by the student (first reviewer), and supervisors (subsequent reviewers) during the process of article searching (title and abstract search) and full-text article appraisal to ensure rigour of the review (Wager & Wiffen, 2011).

Data extraction

The study used an adapted data extraction tool to summarise the results of the key studies included in the review. This included: Author; Context; Content – Aims and Objectives; Method and Data analysis; Sample size and age cohort; Key findings; and Measurement tool used (See Appendix C: Data extraction table template).

Data synthesis

Narrative synthesis was used to synthesise the findings of the studies included in the review (Schwarz et al., 2019). According to Popay et al. (2006), narrative synthesis refers to an approach to the systematic review and synthesis of findings from different studies that relies primarily on the use of words and text to summarise and explain the findings of the synthesis. It is a type of meta synthesis, whereby the systematic review findings will be analysed. While narrative synthesis can involve the manipulation of statistical data, the defining characteristic is that it adopts a textual approach to the process of synthesis to tell the story of the findings from the included studies. Narrative synthesis can be used in a systematic review focusing on a wide range of questions, and not only those relating to the effectiveness of interventions. Narrative synthesis was used to synthesise the findings, as it aligns to the focus of the study to examine the screening tools used to assess maternal substance use, its effectiveness, and limitations of those screening tools.

Procedure and ethics

Wager and Wiffen (2011) explained key considerations for addressing the ethics of a systematic review. This includes specifying how databases are accessed, stating the reviewers of the study, good practice in publishing a systematic review in terms of avoiding plagiarism through citing all sources used as full references, and ensuring accuracy (by extracting data accurately and not to slant the results in a particular way). Unlike primary researchers, systematic reviewers use publicly accessible documents (Suri, 2020). In the current review, three reviewers (student and two supervisors) conducted the review. The student worked independently, and the supervisors deliberated the findings of the review to reach consensus to ensure accuracy of the studies appraised, inclusion of articles, and extraction of information. The student had access to the specified databases through the UWC library to conduct the

31

searches. Furthermore, all references obtained through the review were cited in full, with clear acknowledgement of the respective authors. Lastly, UWC is acknowledged as the institutional resource that enabled the review to be completed.

Ethics clearance was obtained from the Humanities and Social Sciences Research Ethics Committee of the UWC. Ethics considerations regarding this systematic review emphasised the importance of being methodical, thorough, and precise. It was also important to verify the findings and to minimise bias to ensure accuracy of results and findings (Leibovici & Reeves, 2005).

32

Chapter 3: Results

The review aimed to determine the existing empirical research on screening methods (biological or self-report) used to detect substance use in pregnant women. This section presents the results of the review using narrative synthesis. The studies included in the review were synthesised and discussed under the key headings of:

● Context: Location

● Content: Aim, objective, and design

● Sample size and age cohort

● Key findings and Evaluation summary

Article search procedure

The initial title search was using the identified keywords. It yielded 487 results from EBSCOhost, and 226 articles from SAGE and JSTOR duplicates were from SAGE and JSTOR, it yielded 322 articles. These articles were saved and managed in EBSCOhost. A thorough title search was then conducted by two reviewers, which resulted in 40 articles being selected for the next stage of abstract appraisal. Therefore, 282 articles were excluded as they did not meet the inclusion criteria. Of these 40 articles, 25 were excluded at this stage of the review, as 19 articles did not give clear information on the utility, effectiveness, and limitation of the screening tools, 3 studies did not indicate the type of study or the findings did not have information about substance use measuring tools, and 3 studies were excluded as they were published before 2010. The articles that were included at this stage were based on the study inclusion criteria, reflected in the focus on the screening tool/s used, the utility of the screening tool. Eleven screening tools were identified in the studies included in this systematic review, from the 15 articles included in the review.

The full article screening using the Critically Appraised Topics (CAT) was conducted on the 15 included articles and was divided between the two reviewers. Each of the articles underwent a rigorous review and was evaluated based on the critical appraisal tool. All 15 studies qualified for inclusion in the review. Figure 1 below provides a summary of the overall selection process of articles by means of a flowchart.

Figure 1: *Article selection flow diagram*
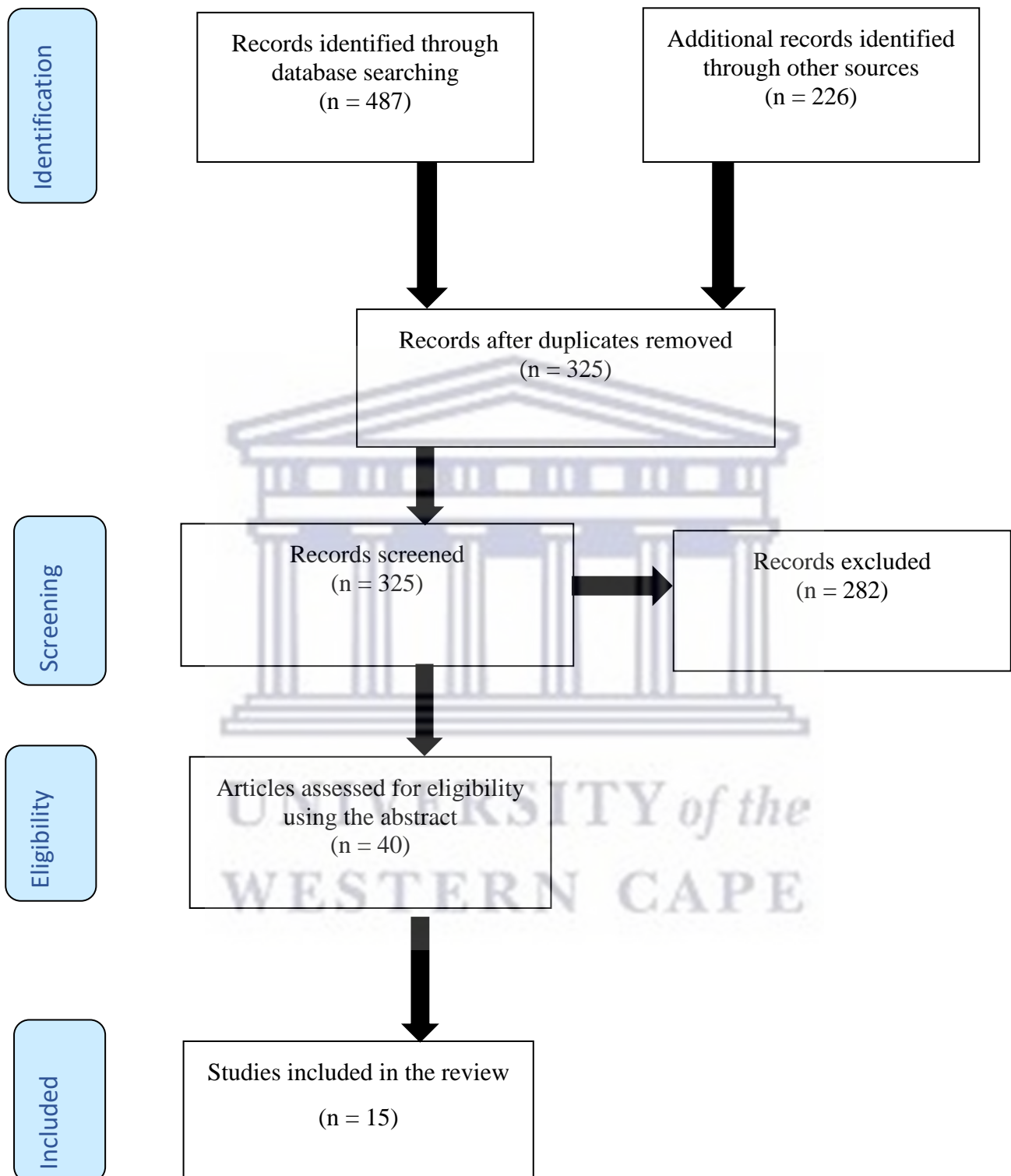
Source of Articles

The articles for the review study were initially searched for within the databases of EBSCOhost, PubMed, Google Scholar, CENTRAL, Embase, SAGE and JSTOR. Of the 15 included articles in the systematic review, 10 were from the PsychINFO database and 5 from CINAHL. The key findings across the studies are summarised in Table 3 below.

Key findings

Context: Location

All the included articles targeted pregnant women and the tests conducted were done after each woman had given birth (post-partum) (see Table 3). Most of the articles accepted into the systematic review were conducted outside the African context, namely in Ibiza (Spain), Canada, Marylyn (USA) and Estonia. This indicates that there is a lack of such studies conducted in the African context, this is only highlighted because in the selected articles only two studies were conducted in Cape Town, South Africa (Williams et al., 2018) and Lusaka, Zambia (Moise 2019). Studies that were conducted at Baltimore, Maryland, USA include the study by Cowger et al.,2018; Oga et al.,2019; Marie 2018, respectively. The study conducted by Garg et al. (2016) was done at Manchester Specialist Midwifery Service (MSMS), Manchester, England. While the study done by Chiodo et al. (2019) was completed at American College of Nurse-Midwives, USA. In addition to the list of studies conducted within the USA, Chang et al. (2019) study indicated that the Recruitment of the research took place at three sites offering outpatient prenatal care in the USA (the Yale New Haven Health System in New Haven, CT; Massachusetts General Hospital in Boston, MA; and Henry Ford Health System in Detroit, MI).

The studies conducted by (Ondersma et al., 2019; Grekin et al., 2010) were both conducted in Detroit, Michigan. Breanna (2020) conducted his study in Southern California. The study of Moise (2019) was conducted in Two public health clinics of Lusaka, Zambia, Kalingalinga urban clinic and Mtendere peri-urban clinic. Præstegaard et al. (2018) conducted their research in Aarhus University Hospital and Fredericia Hospital, Denmark. The study of Grekin et al. (2010) was conducted in Obstetric hospital in Detroit. The study by Yonker et al. (2011) was conducted at an inner-city hospital clinic and was predominantly African American

37

and Hispanic. Nokoo et al. (2016) conducted a non-systematic analysis using the database of Medline and Embase. The study conducted by Sarkar et al. (2010) was done in Toronto, Canada through the Mother Risk Alcohol Helpline (MRAH).

Content: Aim, objectives, and design

Each of the articles included in the systematic review possess their own set of aims and objectives. Coleman-Cowger et al. (2018) conducted a study aimed at determining the most effective screening tool to identify prescription drugs and illicit drugs abused by pregnant women which is acceptable among patients and clinicians so that evidence-based guidance may be offered utilising hair and urine samples as a criterion variable.

A cohort study conducted by Garg et al. (2016) aimed to assess the validity of self-report for the use of major classes of illicit drugs and opioid-maintenance therapy among pregnant women at a substance abuse treatment programme. Another cohort study was conducted by Sarkar et al. (20101) which aimed to compare the effectiveness of the TWEAK and T-ACE screening tools in identifying problem drinking using traditional cut-points (CP). The purpose of the descriptive analysis study conducted by Chiodo et al. (2019) was to describe, in a national sample of professional American midwives and nurse-midwives, their self-reported knowledge of the prevalence, levels and risks of drinking alcohol during pregnancy, attitudes toward drinking during pregnancy, knowledge of the prevalence and characteristics of FASD, awareness of and current practice in using standardised clinical screening tools during pregnancy, perceived barriers to in-pregnancy alcohol screening and intervention, and the relationships among midwives' knowledge in these areas and their clinical practice.

The cross-sectional study conducted by Moise (2019) determined individual correlates and the prevalence of alcohol use in pregnant women attending prenatal care at two health clinics in Lusaka, Zambia. Præstegaard et al. (2018) conducted a cross- sectional study which aimed to investigate the ability of TWEAK to identify alcohol risk drinkers among pregnant Danish women. Another Cross-Sectional study by Chang et al. (2019) evaluated five self-report, non-proprietary questionnaires in the identification of substance use disorders (including alcohol, cannabis, opioids and stimulants) among pregnant women to develop an ideal questionnaire set in the complicated societal context, which includes increasing rates of use and potential sanction. Additional to some of the Cross- Sectional studies conducted includes a study that aimed to determine predictors of alcohol use among pregnant women (Williams et al. 2018). Yonkers et al. (2011) also added a cross-sectional study that aimed to report on the development of a questionnaire to screen for hazardous substance use in pregnant women and to compare the performance of the questionnaire with other drug and alcohol measures.

Nikoo et al. (2017) conducted an analytical review with the aim to examine the existing evidence for an achievable new direction for medical school curricula teaching physicians the required skills to conduct a critically conscious, compassionate, and safe prenatal screening for substance use using a non-systematic literature search.

Breanna (2020) conducted a study using semi-structured interviews with the aim to determine the frequency of substance use screenings in care and assess how well the existing services in Marin County are serving the needs of pregnant women. A cross-sectional study of Marie (2018) aimed to explore pregnant women's acceptability of alcohol, tobacco and drug use screening and willingness to disclose their use in prenatal care. Ondersma et al. (2012)

39

conducted a study aimed to develop and begin validation of an indirect screener for identification of drug use during pregnancy, without reliance on direct disclosure.

Gerkin et al. (2010) examined the ability of the Drug Abuse Screening Test (DAST-10) to identify prenatal drug use using hair and urine samples as criterion variables.

Lastly the quantitative analysis study of Oga e al. (2019) aimed to utilise the 4Ps Plus screener for three purposes; To examine the prevalence of self-reported problematic parental and partner substance use and the associations with substance use during pregnancy, to evaluate the utility of self-reported perceptions of parental and partner substance use as proxies for prenatal drug use and to quantify the degree to which the sensitivity of the screener can be improved if self-reported problematic parental and partner substance use are included in the evaluation of screening results.

Sample size and age cohort

Out of the 15 articles included in the paper, five of these included utilised sample sizes more than 1000 participants and thus demonstrated good methodological quality in terms of sample size (see Table 3) (Chang et al., 2019; Chiodo et al., 2019; Præstegaard et al., 2018; Williams et al., 2018; Yonkers et al., 2011). The study by Chang et al. (2019) included 1220 pregnant women. While Chiodo et al.'s (2019) study had a total of 6000 participants, which included certified nurse-midwives (CNMs), certified professional midwives (CPMs), certified midwives (CMs), women's health nurse practitioners (WHNPs), and nurses (RNs) who provide prenatal care in the United States were invited to participate. Further after one month of recruitment, survey response was low (N = 26) (Chiodo et al., 2019).

In the study by Præstegaard et al. (2018), all pregnant women participated in the

40

antenatal care program. Danish-speaking pregnant women referred to the Midwife Centre at Aarhus University Hospital (n = 1554, representing every second woman attending every day) and at Fredericia Hospital (n = 499, all pregnant women) for routine antenatal care were invited to participate in this study at their first visit. There were two groups who shared the same characteristics excluding location, namely Aarhus University Hospital (n = 1554, representing every second woman attending every day) and Fredericia Hospital (n = 499, all pregnant women).

The study conducted by Williams et al. (2018) included 5231 pregnant women attending 11 Midwife Obstetric Units (MOUs) in Cape Town. The study population comprised pregnant women 16 years or older, presenting for their first antenatal visit and non-pregnant women and women who had previously presented for antenatal visit. Yonkers et al. (2011) study included pregnant women who presented for obstetric care at three hospital based reproductive health clinics, the sample included a number of 1610 who were selected for training and a total of 1074 who were selected for validation.

Ten of the articles included in the review utilised sample sizes less than 1000 individuals and thus demonstrated limited methodological quality in terms of sample size (Breanna, 2020; Coleman-Cowger et al., 2018; Nikoo et al., 2017; Ondersma et al., 2019; Moise, 2019; Marie, 2018; Sarkar et al., 2010; Nikoo et al., 2016; Gerkin et al., 2010; Oga et al., 2019). In the study by Cowger et al. (2018), there were trials expected to be done within 2 clinics. Further Cowger et al. (2018) indicated that at the first clinic there would be an estimated number of 950 individual obstetrical patients that were cared for on an annual basis, where they had anticipated approaching 403 which would constitute nearly 50% and expected 322 which would constitute 80% or more to agree to participate in this study. In the second clinic

41

which approximately has 500 pregnant women aged 18 years or older cared for annually, they would expect at least 450 (90%) to meet eligibility criteria. For this second clinic, they had estimated approaching 225 (50%) pregnant women and expect 180 (80%) or more to agree to participate in this study (Cowger et al., 2018).

In the study conducted by Garg et al. (2016), 102 pregnant women aged18 years of age or older, and had reached 35 weeks' gestation or less were included in the study. These pregnant women were recruited from a University of New Mexico Hospital (UNMH) affiliated specialty prenatal care clinic serving patients with current or history of substance use. Ondersma et al. (2019) included a sample of 50 pregnant women aged from 18 years and above, who were seeking prenatal care services and scoring positive for marijuana use risk.

On the other hand, the study conducted by Breanna (2020) included five healthcare professionals (perinatal case manager, substance abuse counsellor, obstetrician, and two paediatricians) that worked at the local Marin Community Clinic and local organizations were purposefully chosen via purposive sampling from the preceptor's recommendation. These selections were based on the relevance of the provider and staff experience because of their interests in women's maternal health per the MCAH's close work with them. A sample of pregnant women aged between 18 and above in the study by Moise (2019) were recruited from two public health clinics of Kalingalinga urban clinic and Mtendere peri-urban clinic, which provide free prenatal care to pregnant women in these parts of Lusaka province in Zambia. At Kalingalinga urban clinic, 90 women were approached and 79 (87.78%) completed the tool opposing 109 of 120 women approached at Mtendere clinic (90.83%).

In the study conducted by Marie (2018), 589 pregnant women aged 18 and older, were

recruited at their first prenatal care visit from four university-affiliated prenatal care facilities in Baltimore, Maryland and Southern Louisiana using. A total of 75 pregnant problem drinkers were included in the study of Sarkar et al. (2010) who called the mother risk Helpline to report any alcohol consumption during pregnancy and at least 2 months prior were invited to participate in our study. In the study conducted by Gerkin et al. (2010) participants were 300 women who had recently given birth at an obstetric hospital in Detroit. Gerkin et al. (2010) conducted a non-systematic literature review which included 47 articles included on the study focusing on drug use during pregnancy, in addition the study looked at validating the drug abuse screening test against physiological measures. Oga et al. (2019) conducted a study that included a total of 500 pregnant women aged 18 years or older, who speak and understand English well enough to provide informed consent.

Key findings and evaluation summary

This section will focus on the evaluation of each study included in the results (see Table 3). It considers the recommendations, implications and limitations identified authors of each study included in this review results. Some of these studies have highlighted the importance of education, inclusivity, and volunteer work in ensuring the usefulness of measuring substance use during pregnancy. Cowger et al. (2018) argue that a clear recommendation for a clinically useful screening tool for prenatal substance use is crucial to allow for prompt and appropriate follow-up and intervention. While on the other hand, Garge et al. (2016) has strongly emphasised that the strength of their study is the inclusion of Hispanic and Native American pregnant patients, who are underrepresented in other cohorts. Chiodo et al. (2019) further mention that effective screening would identify patients at the highest risk for drinking during pregnancy, or drinking at higher foetal risk levels, so that education and intervention can be efficiently targeted. This study (Chiodo et al., 2019) also included how overcoming the barriers to screening and improving midwives' knowledge about FASD and rates of maternal drinking

43

in their own communities and practices could have a positive impact on birth outcomes and child health and development. Screening for alcohol use during pregnancy by primary care providers, including nurse midwives, is a necessary first step in the primary prevention of FAS and other FASDs.

Breanna (2020), in support of the recommendations by Chiodo et al. (2019), suggests that more training and a cohesive approach to screening should be implemented for both hospital and clinic settings. More understanding is needed for other care physicians as well as a need to address the gaps in care for the younger adolescent population, African Americans, and changing the negative perception of healthcare maternal patients have toward them. Examining other social determinants of health are also future implications to consider in perinatal and postnatal care. Moise (2019) indicates that understanding barriers, risks, and protective factors associated with at risky drinking in pregnancy, early screening, counselling, including the provision of additional treatment to those who screen in need of additional services, and other factors (e.g., personal, and environmental) may be useful in the development of policies and interventions aimed at screening and reducing drinking during pregnancy and related consequences.

Marie (2018) advocates that as healthcare providers, public health officials, policy makers, researchers, and community members we need to be actively advocating for upstream decisions and interventions that lead to equitable individual and community health. Individual health care providers, public health officials, policy makers, and community members must remain critical of how our actions, policies, and compliance with this system frequently cause and perpetuate harm to vulnerable women, families, and communities.

44

Williams et al. (2018) also supports the notion that the training and education of health providers regarding the factors influencing continued alcohol use during pregnancy is important. These findings are useful in guiding understanding of contextual issues surrounding continued alcohol use while pregnant and guiding our efforts in the prevention of FASD. It also aids in the achievement of Sustainable Development Goal (SDG) 3, and specifically SDG 3, target 3.5 (Substance abuse: Strengthen the prevention and treatment of substance abuse, including narcotic drug abuse and harmful use of alcohol), which aims to prevent harmful use of alcohol. Nikoo et al. (2016) argues that further research is recommended on the availability and efficacy of training in the form of postgraduate continuous medical education programs and workshops. This can increase the quality of Prenatal Screening for Substance Use (PSSU) through a broad spectrum of graduated care providers who are currently in practice.

On the issues regarding limitations raised on the studies reviewed, there seems to be discussions on timeframe and sample size. Garg et al (2016) indicates that the reported time frames of urine drug screens and self-report did not perfectly overlap for all participants. Imperfect overlap between Urine Drug Screens and self-report was that patients who were enrolled later in the third trimester could have had positive Urine Drug Screens prior to enrolment (which they abstracted from their medical records) but may not have reported that drug use due to recall. Ondersma et al. 2019) states that the limitations of their study centre on its small size and exploratory nature. The sample size of 45 was seen as adequate to inform basic feasibility and acceptability questions but is too small to support broader inferences or to make valid comparisons between conditions. Also, as noted, participant self-selection of texting frequency limits comparability in terms of retention/opting out of text messaging.

Sarkar et al. (2010) indicates that the methodological strength of their study is based

45

on the fact that in this cohort, women made the initial contact of their own volition and provided a full account of their exposure. This completeness of data provision is unique in research involving high-risk pregnant women. This is the first study comparing the effectiveness of TWEAK and T-ACE among women who voluntarily admitted to drinking alcohol at risk levels. As a result, any inherent concerns associated with underreporting are minimal in this cohort. Yonkers et al. (2011) mentions that the Substance Use Risk Profile-Pregnancy scale can potentially detect a range of substances that may be abused. Clinicians need to further assess women with a positive screen to identify those who require treatment for alcohol or illicit substance use in pregnancy. Oga et al. (2019) states that future research should further examine problematic partner use, how it aligns with actual partner substance use, and how prenatal care interventions can address partner use within the context of maternal substance use.

While the studies above have highlighted the importance of education, training and ensuring that policy within measuring substance use in pregnant women is strengthened, Oga et al. (2019) asserts that future research should further examine problematic partner use. This should also consider how it aligns with actual partner substance use, and how prenatal care interventions can address partner use within the context of maternal substance use.

The study by Cowger et al. (2019) focused on comparison and validation of screening tools for substance use in pregnancy. Furthermore Cowger, et al., (2019) explains the sensitivity and specificity results which were identified during application of the tools 4P's Plus, NIDA Quick screen Assist and SURP-P, those results found that for the 4P's Plus sensitivity amounted to 90.2% (84.5, 93.8) and specificity amounted to 29.6% (24.4, 35.2). For the NIDA Quick screen - ASSIST sensitivity amounted to 79.7% (71.2, 84.2) and specificity

amounted to 82.8% (78.1, 87.1). For the SURP- sensitivity amounted to 92.4% (87.6, 95.8) and specificity amounted to 21.8% (17.4, 27.2). Test-retest reliability (phi correlation coefficients) was 0.84, 0,77 and 0,79 for the 4P's plus, NIDA Quick Screen ASSIST and the SURP-P, respectively. For all screening tools, there were differences in validity indices by age and race, but no differences by trimester.

The study conducted by Garg et al. (2016) on the validity of self-reported drug use information among pregnant women where the sample included a large proportion of ethnic minorities (80% Hispanic/Latina and 7% American Indian) and socially disadvantaged (50% < less than high school education and 94% Medicaid-insured) pregnant women. The study found that on average, patients had 4.8 ± 3.0 urine drug screens during the third trimester. Sensitivity of self-report was low (<47%).

The research study by Chiodo et al. (2019) found that 37.7% of the respondents believe drinking alcohol is safe during at least one trimester of pregnancy. Only 35.2% of respondents reported screening to assess patient alcohol use. Only 23.3% reported using a specific screening tool, and few of those were validated screens recommended for use in pregnant women. Respondents who believe alcohol is safe at some point in pregnancy were significantly less likely to screen their patients.

The study conducted by Ondersma et al. (2019) reported that all participants assigned to receive the e-SBI (n = 30) were able to complete it during the in-clinic baseline session. Participants assigned to text messaging (n = 30) received a median of 24 text messages before giving birth; 6 of 30 (20%) chose to end text messages before giving birth. Acceptability ratings for the e-SBI were high, with most being above 4.4 on a 5-point scale (e.g., for ease of use, respectfulness, and helpfulness). Acceptability ratings for some aspects of the texting

47

intervention were also high (e.g., for ease of understanding and respectfulness). Participants in the combined e-SBI plus texting condition chose to stop messages at a rate of 13.3% (2 of 15), versus 26.7% (4 of 15) in the texting-only condition (p = 0.37).

The study by Breanna (2020) found that there is no technical universal screening tool used. Participants also noted several common themes in gaps of care, in terms of patient's views toward health care, needed improvements, common substances seen, the difference between the hospital and clinic protocol, adolescents, and African Americans.

While the study by Moise (2019) on the alcohol use, pregnancy, and associated risk factors: a pilot cross-sectional study of pregnant women attending prenatal care in an urban city suggested that about 40 (21.2%) pregnant women were identified by the T-ACE as at-risk for problem drinking during pregnancy. Except for regular prenatal care and distance, there was no difference in the demographic factors between pregnant women who scored < 2 on the T-ACE and those that scored > 2 points (all p's > 0.05). A small proportion of women at both clinics reported binge drinking during the periconceptional period (12.7% vs. 3.2%, p = 0.003) and beyond periconceptional period. Excluding employed women, no significant relationships were observed between alcohol use and demographic factors.

The study on pregnant women's acceptability of substance use screening and willingness to disclose use in prenatal pare conducted by Marie (2018), discovered that a substantial majority of pregnant women found screening acceptable for alcohol (97%), tobacco (98%), and drug use (97%) during prenatal care. Screening for alcohol use was more unacceptable among women who did not report risky alcohol use compared to women who did report risky alcohol use (P = 0.08). Tobacco use, drug use, and previous CPS involvement were

48

not associated with acceptability of screening. A substantial majority of pregnant women reported they are willing to honestly disclose alcohol (99%), tobacco (99%), and drug use (98%). Alcohol, tobacco, and drug use and prior CPS involvement were not associated with the willingness to disclose substance use.

The study conducted by Præstegaard et al. (2018) that explores whether TWEAK is a valid screening questionnaire to identify alcohol risk drinkers among pregnant women in Denmark, found that found that the sensitivity of TWEAK to identify periconceptional risk drinking was quite low, but its ability to identify risk drinkers during pregnancy was marginally higher. The study results suggested that older age (odds ratio 1.46, 95% confidence interval 0.95–2.23), current smoking (odds ratio 2.33, 95% confidence interval 1.63–3.33), being single (odds ratio 2.38, 95% confidence interval 1.38–4.11) and a TWEAK score with a cut-off score of ≥1 (odds ratio 2.75, 95% confidence interval 2.02–3.76) increased the risk of high-risk drinking during pregnancy.

The study by Sarkar et al. (2010) found that using traditional cut-off points, the TWEAK and T-ACE tools both performed similarly at identifying potential at-risk women. Upon comparison, there was no statistical difference in the effectiveness for one test performing better than next using either cut-off points of 2 (P = 0.66) or cut-off point of 3 (P = 0.38).

The study conducted by Williams et al. (2018) found that 'Coloured' women, having a marital status other than being married, experiencing violence or aggression in the past 12 months compared to more than 12 months ago, having a partner who drinks, and partner drug use are all independently associated with higher odds of self-reported alcohol use. In contrast, only partner tobacco use is independently associated with higher odds of biologically verified alcohol use.

49

Yonkers et al. (2011) discovered that the SURP-P consists of three informative questions that can be scored for high- or low-risk populations. The SURP-P algorithm for low-risk populations was mostly highly predictive of substance use in the validation subsample (Akaike's Information Criterion=579.75, Nagelkerke R^2=0.27) with high sensitivity (91%) and adequate specificity (67%). The high-risk algorithm had lower sensitivity (57%) but higher specificity (88%). This discovery was made during the study of Screening for Prenatal Substance Use: Development of the SURP-P.

The main findings of the study conducted by Chang et al. (2019) on the identification of substance use disorders among pregnant women: A comparison of screeners show that among the five questionnaires tested, only two (the CRAFFT and SURP-P) demonstrated modest overall accuracy. For example, while the NIDA Quick Screen had the highest accuracy for any substance use disorders (0.85), it was the most specific (0.98), the least sensitive (0.28) and had the smallest AUROC (0.62) compared to the other questionnaires. On the other hand, the CRAFFT had the largest AUROC (0.75), with reasonable sensitivity (0.68), specificity (0.83), and accuracy (0.81). The SURP-P had the most comparable sensitivity (0.78) and specificity (0.71), as reflected in its accuracy (0.72) and AUROC (0.74).

The study on conducted by Nikoo et al (2016) which focused on the effectiveness of prenatal screening for substance use critical consciousness, a promising curriculum for compassionate screening discovered that current prenatal screening for substance use is not effective despite being widely used. This is primarily due to patients not disclosing their use of the teratogenic substances. Insufficient physician education in substance use is a potential contributing factor for such low disclosure. The study conducted by Gerkin et al. (2010) found

50

that 19% of the sample had a positive toxicology screen but denied drug use on the DAST-10. The findings suggest that brief drug use screeners may have limited utility for pregnant women and that efforts to facilitate disclosure via reassurance and anonymity are unlikely to be sufficient in this population.

Finally, the study on prenatal substance use and perceptions of parent and partner use using the 4P's Plus Screener conducted by Oga et al. (2019) found that half (51%) of respondents reported either partner or parental problematic substance use. When partner or parent problematic substance use were considered as proxies for prenatal substance use, sensitivity was 65% and specificity was 55%. When used as intended, sensitivity was 94% and specificity was 29%. Including partner/parent questions increased sensitivity to 96% but lowered specificity (19%). Partner substance use and combined partner/parent use were associated with prenatal substance use [adjusted odds ratio (aOR): 2.0 (1.2, 2.4; p = 0.006); aOR= 1.6 (1.1, 2.5, p = 0.04)].

Table 3: *Characteristics of Included Articles*

Author	Title	Aim	Design	Sample	Location	Key Findings
1. Cowger et al. (2018)	Comparison and validation of screening tools for substance use in pregnancy: a cross-sectional study conducted in Maryland prenatal clinics	This study compares and validates three existing substance-use screeners for pregnancy—4 Ps Plus, National Institute on Drug Abuse (NIDA) Quick Screen, Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST) and the Substance Use Risk Profile-Pregnancy (SURP-P) scale.	Cross-Sectional Study	N = 500 pregnant women	Baltimore, Maryland, USA	The SURP-P and 4P's Plus had high sensitivity and negative predictive values, making them more ideal screening tests than the NIDA Quick Screen-ASSIST. A clear recommendation for a clinically useful screening tool for prenatal substance use is crucial to allow for prompt and appropriate follow-up and intervention.
2. Garg et al. (2016)	Validity of self-reported drug use information among pregnant women	This study assesses the validity of self-report for the use of major classes of illicit drugs and opioid-maintenance therapy among pregnant women at a substance abuse treatment programme.	Cohort Study	N = 102 pregnant women	Manchester Specialist Midwifery Service (MSMS), Manchester, England	This study found substantial underreporting for all classes of illicit drugs amongst pregnant women in a substance abuse treatment program. Further it was discovered that rates of underreporting are expected to be higher among the general population of pregnant women.
3. Chiodo et al. (2019)	Prenatal Alcohol Screening During Pregnancy by Midwives and Nurses	The purpose of this study was to describe, in a national sample of professional American midwives and nurse-midwives, their self-reported knowledge of the prevalence, levels and risks of drinking alcohol during pregnancy, attitudes toward drinking during pregnancy, knowledge of the prevalence and characteristics of FASD, awareness of and current practice in using standardised clinical screening tools during pregnancy, perceived barriers to in-pregnancy alcohol screening and intervention, and the relationships among midwives' knowledge in these areas and their clinical practice.	Descriptive Analysis	N = 6000	American College of Nurse-Midwives, USA	Screening for alcohol use during pregnancy by primary care providers, including nurse midwives, is a necessary first step in the primary prevention of FAS and other FASDs. Improving prenatal care provider knowledge about the effects of prenatal alcohol exposure and the availability of valid alcohol screening tools will improve detection of drinking during pregnancy, provide more opportunities for meaningful intervention, and ultimately reduce the incidence of FASD.
4. Ondersma et al. (2019)	Feasibility and Acceptability of e-Screening and Brief Intervention and Tailored	This study sought to evaluate the feasibility and acceptability of two high-reach technology-based interventions: electronic screening	Exploratory Study	N = 45	Detroit, Michigan	These two high-reach intervention elements showed strong feasibility and modest to high acceptability. Future efforts evaluating efficacy are warranted and should specifically examine the possibility that

52

	Text Messaging for Marijuana Use in Pregnancy	and brief intervention (e-SBI) and tailored text messaging, delivered either alone or in combination.				combining text messaging with a brief intervention may promote retention.
5. **Breanna (2020)**	Perinatal Substance Use Screenings in Marin County: a brief overview of screening protocols and identifying gaps in care	The purpose of this study was to determine the frequency of substance use screenings in care and assess how well the existing services in Marin County are serving the needs of pregnant and parenting women and identifying the gaps and/or weaknesses in current practice.	Quantitative Analysis	N= 5 healthcare professionals (perinatal case manager, substance abuse counsellor, obstetrician, and two paediatricians)	Southern California	Findings suggest more training and a cohesive approach to screening should be implemented for both hospital and clinic settings. More understanding is needed for other care physicians as well as a need to address the gaps in care for the younger adolescent population, African Americans, and changing the negative perception of healthcare maternal patients have toward them. Examining other social determinants of health are also future implications to consider in perinatal and postnatal care
6. **Moise (2019)**	Alcohol use, pregnancy and associated risk factors: a pilot cross-sectional study of pregnant women attending prenatal care in an urban city	This study determined individual correlates and the prevalence of alcohol use in pregnant women attending prenatal care at two health clinics in Lusaka, Zambia.	Cross-sectional Design	N = 79–120 pregnant women	Two public health clinics of Lusaka, Zambia, Kalingalinga urban clinic and Mtendere peri-urban clinic	Alcohol consumption is prevalent in the periconceptional period and during pregnancy in pregnant women attending prenatal care in Zambia. Findings underscore the need for targeted alcohol use screening and intervention for pregnant women
7. **Marie (2018)**	Pregnant Women's Acceptability of Substance Use Screening and Willingness to Disclose Use in Prenatal Care	To explore pregnant women's acceptability of alcohol, tobacco, and drug use screening and willingness to disclose their use in prenatal care. This research sought to centre the voices of pregnant women who are directly impacted by the adverse consequences of screening as central to informing the policies and practices that directly impact them, their health and wellbeing, the care they receive, and the formation of their families.	Cross-sectional Study	N = 589 pregnant women	Baltimore, Maryland and Southern Louisiana, USA	Pregnant women, including those who reported substance use or prior CPS involvement, found substance use screening in prenatal care acceptable and were willing to honestly disclose their use.
8. **Præstegaard et al. (2018)**	Is TWEAK a valid screening questionnaire to identify alcohol risk drinkers among pregnant women in Denmark?	To investigate the ability of TWEAK to identify alcohol risk drinkers among pregnant Danish women.	Cross-sectional Study	University Hospital (n = 1554),	Aarhus University Hospital and Fredericia	sensitivity of TWEAK to identify periconceptional risk drinking was quite low, but its ability to identify risk drinkers during pregnancy was marginally higher, which basically mean that in a Danish setting, TWEAK does not seem like an optimal screening tool to identify

53

				Fredericia Hospital (n = 499)	Hospital, Denmark	periconceptional risk drinkers, but it may be useful in identifying high-risk drinking during pregnancy.
9. Sarkar et al. (2010)	Assessment and detection: comparing the effectiveness of tweak and t-ace in determining problem drinkers in pregnancy	To compare the effectiveness of the TWEAK and T-ACE screening tools in identifying problem drinking using traditional cut-points (CP).	Cohort Study	N = 75 Problem Drinkers N = 100 Non-Problem Drinkers	Women calling the Alcohol Helpline for mothers at risk. Toronto, Canada	Despite the lack of difference in performance, improved specificity associated with TWEAK suggests that it may be better suited to screen at-risk populations seeking advice from a helpline.
10. Williams et al. (2018)	Predictors of Alcohol Use during Pregnancy among Women Attending Midwife Obstetric Units in the Cape Metropole, South Africa.	To determine predictors of alcohol use among pregnant women.	Cross-sectional Study	N= 5231 pregnant women	Greater Cape Town, South Africa	Findings highlight various demographic; social and partner substance use predictors for both self-reported and biologically verified alcohol use in two different models. Being Coloured, having a marital status other than being married, experiencing violence or aggression in the past 12 months compared to more than 12 months ago, having a partner who drinks, and partner drug use are all independently associated with higher odds of self-reported alcohol use. In contrast, only partner tobacco use is independently associated with higher odds of biologically verified alcohol use.
11. Yonkers et al. (2011)	Screening for Prenatal Substance Use: Development of the Substance Use Risk Profile-Pregnancy Scale	To report on the development of a questionnaire to screen for hazardous substance use in pregnant women and to compare the performance of the questionnaire with other drug and alcohol measures.	Cross-sectional Study	'training' (n=1,610) and 'validation' (n=1,074)	inner-city hospital clinic and was predominantly African American and Hispanic.	The Substance Use Risk Profile-Pregnancy scale can potentially detect a range of substances that may be abused. Clinicians need to further assess women with a positive screen to identify those who require treatment for alcohol or illicit substance use in pregnancy.
12. Chang et al. (2019)	Identification of substance use disorders among pregnant women: A comparison of screeners	To evaluate five self-report, non-proprietary questionnaires in the identification of substance use disorders [including alcohol, cannabis, opioids, and stimulants] among pregnant women.	Cross-sectional Study	N= 1220 pregnant women	Recruitment took place at three sites offering outpatient prenatal care (the Yale New Haven Health System in New Haven, CT; Massachusetts General	The main findings of this study show that among the five questionnaires tested, only two (the CRAFFT and SURP-P) demonstrated modest overall accuracy. In addition, the CRAFFT and SURP-P exhibited modest ability to identify substance use disorders among pregnant women. Future research is needed to develop an ideal questionnaire set in the complicated societal context which includes increasing rates of use and potential sanction

54

					Hospital in Boston, MA; and Henry Ford Health System in Detroit, MI), USA	
13. Nikoo et al. (2016)	Effectiveness of Prenatal Screening for Substance Use: Critical Consciousness, A Promising Curriculum for Compassionate Screening.	This analytical review examines the existing evidence for an achievable new direction for medical school curriculum teaching physicians the required skills to conduct a critically conscious, compassionate, and safe prenatal screening for substance use.	Non-Systematic Literature	N= Critical analysis of 47 relevant articles	Medline and Embase	Current prenatal screening for substance use is not effective despite being widely used. This is primarily due to patients not disclosing their use of the teratogenic substances. Insufficient physician education in substance use is a potential contributing factor for such low disclosure.
14. Grekin et al. (2010)	Drug Use During Pregnancy: Validating the Drug Abuse Screening Test Against Physiological Measures	Examined the ability of the Drug Abuse Screening Test (DAST-10) to identify prenatal drug use using hair and urine samples as criterion variables	Not Included	N = 300 pregnant women	Obstetric hospital in Detroit	Twenty-four percent of the sample had a positive drug screen. The sensitivity of the DAST-10 was only 47. Nineteen percent of the sample had a positive toxicology screen but denied drug use on the DAST-10. Findings suggest that brief drug use screeners may have limited utility for pregnant women and that efforts to facilitate disclosure via reassurance and anonymity are unlikely to be sufficient in this population.
15. Oga et al. (2019)	Prenatal Substance Use and Perceptions of Parent and Partner Use Using the 4P's Plus Screener	The study aimed to examine and evaluate the prevalence and utility of self-reported parental and partner substance use. This would then help in quantifying the degree to which sensitivity of the screener can be improved.	Quantitative Analysis	N= 500 pregnant women	Baltimore, Maryland, USA	Sensitivity of the 4P's Plus may improve with inclusion of self-reported problematic partner/parent substance use items in risk stratification

Chapter 4: Discussion

This study aimed to evaluate the empirical research on screening methods (biological or self-report) used to detect substance use in pregnant women. The keyword searches and reference mining yielded 487 articles, of which a total of 15 articles were included. The participants in these studies were pregnant women attending prenatal care, which was the key target population for the current review. The 15 articles included in this study also used a quantitative approach and were from the period from 2010 to 2021. The articles included in the review found that failure to detect appropriately and treat substance use during pregnancy might have a long-term detrimental effect for both the mother and the child.

Most of the articles included in the review were conducted in the global North (USA and Europe). There are only two articles that were conducted in the African context. This highlights the need for further research in other low-middle income contexts in Asia and the Global South to assess and address maternal substance use.

In terms of the research designs employed in the studies, seven used cross-sectional designs (Cowger et al., 2018; Moise, 2019; Marie, 2018; Praestegaard, 2018; Williams et al., 2018; Yonkers et al., 2011; Chang et al., 2019). Two studies made use of a cohort study (Garg et al., 2019; Sarkar et al., 2010), and the remaining studies included made use of the quantitative (descriptive) analysis design (Breanna, 2020; Chiodo et al., 2019; Oga et al., the final used a non-systematic literature design (Nikko et al., 2016).

The key findings on the studies reviewed highlight the importance of screening for substance use, as it assists clinicians to identify individuals who might be at high risk and then provide appropriate intervention or education, together with improving maternal and neonatal health outcomes through supportive approaches rather than punitive means (Oga et al., 2019).

56

Furthermore, the study of Moise (2019) conducted in two Zambian prenatal care clinics found that alcohol consumption is prevalent in the periconceptional period and during pregnancy. In the findings of Chang et al. (2019), universal voluntary self-report screening of pregnant women for alcohol and substance use is recommended to facilitate early identification and modification of a generally problematic behaviour. The study also suggested that it is important to identify not only women who are currently using substances while pregnant but also those who are at risk of using substances. Prior research has shown that the best predictor of continued prenatal alcohol use is past use, which is infrequently identified without systematic screening. This study suggests that to ensure the screening test has true positive results, the test must have high sensitivity. This has also been supported by the study that aimed at comparing three screening tools, namely the SURP-P, NIDA Quick Screen-ASSIST and the 4Ps Plus, which found that the SURP-P and the 4Ps Plus performed with high sensitivity and negative predictive values that were regarded as more effective, than the NIDA Quick Screen-ASSIST; however, the study also found that high sensitivity often comes at the expense of specificity, which was observed from the performance of those screening tools (Cowger et al., 2019).

In addition, Ondersma et al. (2019) found that among five questionnaires (SURP-P, CRAFFT, NIDA, WIDUS and 5Ps) tested in their study, only two screening tools (the CRAFFT and SURP-P) demonstrated modest overall accuracy. In addition, other screening tools include questions about the frequency of substance use or amount of alcohol consumed, as well as questions directly related to drinking behaviour or consequences (e.g., Ondersma et al., 2019). For example, the CAGE focuses on self-perception about the personal impact of drinking, or feeling annoyed when others are concerned about their drinking behaviour;

57

however, according to Chang et al. (2010), Floyd et al. (2010) and Volk et al. (1997), the CAGE has a limited sensitivity and specificity. The CAGE and the AUDIT were specifically designed for the general population; however, they were also validated for their use in pregnant women. The highly sensitive T-ACE and the TWEAK, both adapted from the CAGE, were validated for pregnant women in the previous studies (e.g., Russell, 1994; Russell et al., 1994; Sokol et al., 1989), but are used, respectively, by only 12.3% and 2.5% of the clinicians, and the ability of the TWEAK instrument to identify periconceptional risk drinking was fairly low, while the sensitivity of the TWEAK to identify risk drinking was marginally higher.

In the study of pregnant women presenting for prenatal care at one of the two urban sites in Zambia, self-reported perceptions of problematic partner and parent substance use were high. The study also found that the rate of substance use within the studied population was high, with positive urine drug tests for 27% of the sample. Additionally, they found that the 4Ps Plus performed moderately better in the population compared to the original validation in terms of sensitivity and false negative rates, but not in terms of negative predictive validity even though both were satisfactorily high (NTI Upstream, 2011). However, the 4Ps Plus is a screening rather than a diagnostic tool, and a positive screen is not necessarily indicative of maternal substance use. In a cross-sectional design of Moise (2019), in which pregnant women completed the T-ACE screening tool and validated alcohol-screening questionnaires on self-reported alcohol use at their regular prenatal visit, the study found that of all participants, about 40 (21.2%) were identified by the T-ACE as at-risk for problem drinking during pregnancy. The findings by Yonkers et al. (2011) showed that the 4Ps and 4Ps Plus had moderate to excellent sensitivity (81–95%) for detection of alcohol or illicit substance use in pregnancy but modest specificity (69–45%). The modified TWEAK had lower sensitivity (63%) and higher specificity (86%).

58

Another study suggested that providers should be competent in care for mothers with substance use disorders just as for any other patients. Competency training for substance use screening during pregnancy should continue to be visible and available for clinicians. The study also suggested that the participants in the study gave insight on the need for a universal screening protocol and expressed interest in more training and in understanding more patient care approaches to addressing the subject with patients. The study also found that there were only two screening tools that were identified by the participants (AUDIT-C and NIDA), however, those screening tools are only used as an incentive for insurance purposes and during training.
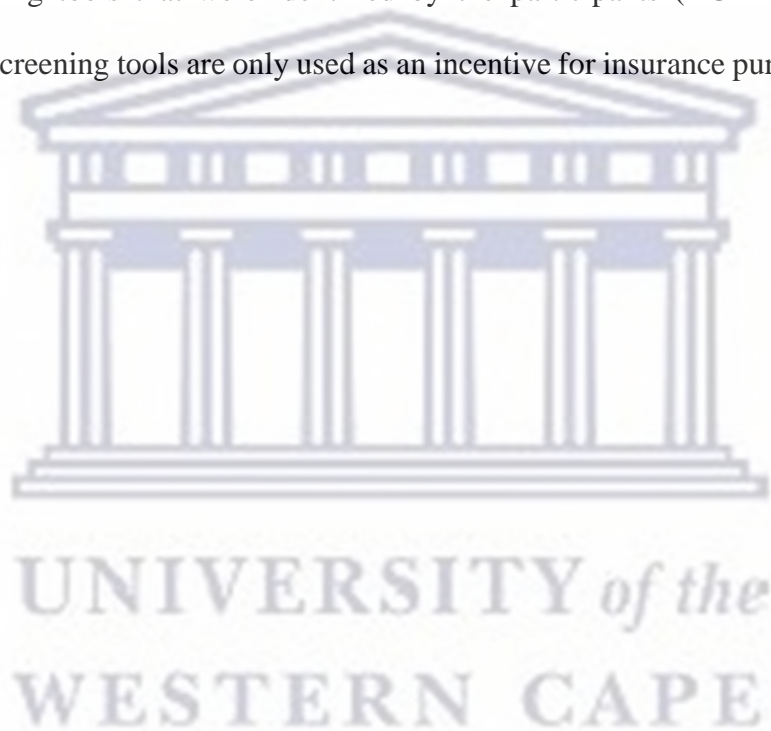
South African Context

The study findings are also important to consider with the South African context, which has among the highest levels of substance use in general, and maternal substance use. Unfortunately, there are a number weaknesses in the public health and monitoring system of the country. According to the Department of Social Development (DSD) Gazette for prevention of and treatment for substance use disorder policy (2023), South Africa lacks a system for monitoring substance use service quality compared to high-income countries that have made use of performance measurement systems to monitor the quality of all health services, including substance use services. Additionally, the lack of a system for monitoring substance use treatment is an additional hindrance particularly for pregnant women. South African women also encounter significant systemic, structural, social, cultural, and personal barriers in accessing treatment for drug use disorders, (DSD, 2023).

Further the DSD (2023) adopted the three key pillars introduced in the recommendations of the 2016 United Nations General Assembly Special Session (UNGASS) outcome document. Based on these key pillars, DSD has taken on the mandate of demand reduction and treatment of substance use disorders and related activities to reduce harm caused by substance use and abuse.

The current systematic review study can contribute towards the second pillar (early intervention) of the UNGASS recommendations adopted by DSD. The early intervention pillar highlights the criticality of determining the nature of substance use in individuals through screening. The department of social development Gazette for prevention of and treatment for substance disorder policy (2023) indicates that both DSD and other stakeholders should adapt internationally validated screening tools for use among this cohort.

Application of the tools identified in the current review has the potential to contribute to the initiative developed by the South African Medical Research Council, known as the

Service Quality Measures (SQM). It was designed to ensure that the performance measurement system for South Africa's Substance use treatment system is efficient. Myers et al. (2017) states that the Service Quality Measure (SQM) is the first performance measurement system for addiction treatment in low- and middle-income countries, which started in 2008. It is used to highlight the strengths and challenges related to effectiveness of treatment and accessibility of treatment programmes for people with substance use disorders.

The study findings contribute to Sustainable Development Goal 3 (Good Health and Well-being). This goal could be addressed in the context by ensuring free access to and assessment for testing maternal substance use.

As part of mitigation for substance use as a public health concern, this study could inform government around the issue of ensuring that all facilities for treatment of substance use are easily accessible, as well as the training of healthcare professionals in measuring substance use. This would have a great impact on the early stages of pregnancy and subsequently on the first 1000 days initiatives across South Africa. The study can contribute towards awareness raising through organisations within the South African context by encouraging testing for substance use and treatment of issues related to substance and alcohol disorder.

61

Conclusion, Limitations, and Recommendation

This review study contributes to the literature by establishing the different types of screening tools used to determine maternal substance, their utility, and limitations. The findings of this study have the potential to benefit the health institutions to identify the most widely used screening tools used to assess maternal substance use, and its efficacy across different contexts. The study findings highlight the importance of biological and self-report screening methods to gain a more holistic understanding of the mother's contextual realities impacting on her substance use. This is important to address mitigation and prevention efforts. It was also found that despite the major public health concern that maternal substance use poses, there are few studies that have examined the topics globally, and particularly in African and Asian contexts. This is important to determine the cross-cultural usability and comparability of the identified screening measures.

Limitations

It is important to acknowledge the studies' potential limitations, as this is inherent in all research. The studies used in this review were limited in number as it was a challenge to access freely available and published data on the measuring tools used in identifying substance use in pregnant women. In addition to this, the inclusion criterion of language whereby only English-language studies were included could have excluded studies published in other languages. Moreover, while grey literature and commissioned reports were excluded from the review and only peer-reviewed publications were considered, this did in turn enhance the rigour of the review. This means that the review may have missed one aspect of the literature that has not been formally published in journal articles or edited books.

Secondly, it is possible that other relevant studies were excluded from this review as it used a qualitative methodological approach to measure substance use in pregnant women. This is a potential area for a future systematic review in the field.

Recommendations

Based on the studies reviewed and included in this study it is clear that measuring substance use during pregnancy is important for all countries and for various reasons, (Forray, 2016). Foremost, prenatal care should be freely available and accessible to all pregnant women. In line with this, hospitals and clinics should consider the use of both self-report assessments and biological tests to determine maternal substance use, (Stone, 2013). Further the office of the surgeons (OSG) (2020) suggests that this should be advanced by policymakers and the necessary allocation from the national healthy budget be made available for maternal care. The implementation of relevant legislation and policies that provide support to the mother and child in cases where a mother is found to have been using substances would be essential. This could include in-and out-patient treatment programmes for the mother while she is still going through antenatal visits and assure protection of confidentiality. These tests would have to be administered on their first visits to clinics and monitored throughout their subsequent health visits. Another intervention could be psychoeducation focusing on knowledge-building among women regarding maternal substance use and the implications thereof for their own their child's health (Hser, 2015). This would also mitigate potential medication dependence in replacement or substituting for any drug use as they would know the long-term consequences.

References

American Psychological Association. (2022). Just-world hypothesis. In APA dictionary of

    psychology. https://www.dictionary.apa.org/just-world-hypothesis

Armstrong, M. A., Gonzales Osejo, V., Lieberman, L., Carpenter, D. M., Pantoja, P. M., &

    Escobar, G. J. (2003). Perinatal substance abuse intervention in obstetric clinics

    decreases adverse neonatal outcomes. *Journal of Perinatology*, 23(1), 3-9.

Chang, G., Orav, E. J., Jones, J. A., Buynitsky, T., Gonzalez, S., & Wilkins-Haug, L. (2011).

    Self-reported alcohol and drug use in pregnant young women: A pilot study of

    prevalence and associated factors. *Journal of addiction medicine*, 5(3), 221.

Chang, G., Ondersma, S. J., Blake-Lamb, T., Gilstad-Hayden, K., Orav, E. J., & Yonkers, K.

    A. (2019). Identification of substance use disorders among pregnant women: A

    comparison of screeners. Drug and Alcohol Dependence, 205, 107651.

Chang, G., Fisher, N. D., Hornstein, M. D., Jones, J. A., & Orav, E. J. (2010). Identification

    of risk drinking women: T-ACE screening tool or the medical record. *Journal of

    Women's Health*, 19(10), 1933-1939.

Chasnoff, I. J., Wells, A. M., McGourty, R. F., & Bailey, L. K. (2007). Validation of the 4P's

    Plus© screen for substance use in pregnancy validation of the 4P's Plus. *Journal of

    Perinatology*, 27(12), 744-748.

Chen, Y., & Vanderweele, T. J. (2018). Associations of Religious Upbringing with

    Subsequent Health and Well-Being from Adolescence to Young Adulthood: An

    Outcome-Wide Analysis. Department of Epidemiology

Coleman-Cowger, V. H., Oga, E. A., Peters, E. N., Trocin, K. E., Koszowski, B., & Mark, K.

    (2019). Accuracy of three screening tools for prenatal substance use. *Obstetrics and

    gynecology*, *133*(5), 952.

Clapp, M. A., James, K. E., & Friedman, A. M. (2020). Identification of delivery encounters using International Classification of Diseases, Tenth Revision, diagnosis and procedure codes. *Obstetrics & Gynecology*, 136(4), 765-767.

Choe, Y. M., Lee, B. C., Choi, I. G., Suh, G. H., Lee, D. Y., & Kim, J. W. (2019). Combination of the CAGE and serum gamma-glutamyl transferase: an effective screening tool for alcohol use disorder and alcohol dependence. Neuropsychiatric disease and treatment, 15, 1507–1515. https://doi.org/10.2147/NDT.S203855

Cook, J. L., Green, C. R., de la Ronde, S., Dell, C. A., Graves, L., Ordean, A., & Wong, S. (2017). Epidemiology and effects of substance use in pregnancy. *Journal of Obstetrics and Gynaecology Canada*, 39(10), 906-915.

Cooke, A., Smith, D., & Booth, A. (2012). Beyond PICO: the SPIDER tool for qualitative evidence synthesis. Qualitative health research, 22(10), 1435-1443.

Degenhardt, L., Saha, S., Lim, C. C., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., & WHO World Mental Health Survey Collaborators. (2018). The associations between psychotic experiences and substance use and substance use disorders: findings from the World Health Organization World Mental Health surveys. Addiction, 113(5), 924-934.

Department of Social Development. (2023). Government Notices: Prevention of and Treatment for Substance Use Disorders Policy. No 49658, 10 November 2023

De Meneses-Gaya, C., Zuardi, A. W., Loureiro, S. R., & Crippa, J. A. S. (2009). Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of psychometric properties. *Psychology & Neuroscience*, *2*(1), 83.

Eriksen, M. B., & Frandsen, T. F. (2018). The impact of patient, intervention, comparison, outcome (PICO) as a search strategy tool on literature search quality: a systematic review. *Journal of the Medical Library Association: JMLA*, *106*(4), 420.

Francis, J. M., Meyers, B., Nkosi, S., Williams, P. P., Carney, T., Lombard, C., Nel, E., & Morojele, N. (2019). The prevalence of religiosity and association between religiosity and alcohol use, other drug use, and risky sexual behaviours among grade 8-10 learners in Western Cape, South Africa. Published: February 13, 2019, https://doi.org/10.1371/journal.pone.0211322

Forray, A. (2016). Substance use during pregnancy [version 1; referees: 2 approved]: Department of Psychiatry, Yale School of Medicine, New Haven, CT, USA

Guttmann, A., Blackburn, R., Amartey, A., Zhou, L., Wijlaars, L., Saunders, N., & Gilbert, R. (2019). Long-term mortality in mothers of infants with neonatal abstinence syndrome: A population-based parallel-cohort study in England and Ontario, Canada. *PLoS medicine*, *16*(11), e1002974.

Han, B. H., Sherman, S. E., Link, A. R., Wang, B., & McNeely, J. (2017). Comparison of the Substance Use Brief Screen (SUBS) to the AUDIT-C and ASSIST for detecting unhealthy alcohol and drug use in a population of hospitalized smokers. *Journal of substance abuse treatment*, *79*, 67-74.

Hostage, J. C., Brock, J., Craig, W., & Sepulveda, D. (2020). Integrating screening, brief intervention and referral to treatment (SBIRT) for substance use into prenatal care. *Maternal and child health journal*, *24*(4), 412-418.

Jones, H. E., Browne, F. A., Myers, B. J., Carney, T., Ellerson, R. M., Kline, T. L., & Wechsberg, W. M. (2011). Pregnant and nonpregnant women in Cape Town, South Africa: drug use, sexual behavior, and the need for comprehensive services. *International Journal of Pediatrics*, *2011*.

Khan, K. S., Kunz, R., Kleijnen, J., & Antes, G. (2003). Five steps to conducting a systematic review. *Journal of the royal society of medicine*, *96*(3), 118-121.

Kim, H., Sefcik, J. S., & Bradway, C. (2017). Characteristics of qualitative descriptive studies: A systematic review. *Research in nursing & health*, *40*(1), 23-42.

Lander, L., Howsare, J., & Byrne, M. (2013). The impact of substance use disorders on families and children: from theory to practice. *Social work in public health*, *28*(3-4), 194-205.

Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: sensitivity and specificity. *Continuing education in anaesthesia critical care & pain*, *8*(6), 221-223.

Macleod, C. I., Young, C., & Molokoe, K. (2021). Alcohol use during pregnancy: prevalence and patterns in selected Buffalo City areas, South Africa. *African journal of reproductive health*, 25(1), 114–121. https://doi.org/10.29063/ajrh2021/v25i1.13

Meyers, B., Govender, R., Williams, P. P., Johnson, K., Kock, J. R., & Manderscheid, R. (2017). Readiness to adopt a performance measurement system for substance abuse treatment: Findings from the Service Quality Measurement initiative. DOI 10.7196

Methley, A. M., Campbell, S., Chew-Graham, C., McNally, R., & Cheraghi-Sohi, S. (2014). PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC health services research*, *14*(1), 1-10.

Moher, D., Shamseer, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic reviews*, *4*(1), 1-9.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine*, *151*(4), 264-269.

Myers, B., Koen, N., Donald, K. A., Nhapi, R. T., Workman, L., Barnett, W., & Stein, D. J. (2018). Effect of hazardous alcohol use during pregnancy on growth outcomes at

birth: Findings from a South African cohort study. *Alcoholism: Clinical and Experimental Research*, *42*(2), 369-377.

Ns for informing evidence-based healthcare decisions: a five-step approach (part 1/3). *Expert review of pharmacoeconomics & outcomes research*, *16*(6), 689–704. https://doi.org/10.1080/14737167.2016.1246960

Office of the Surgeon General (OSG). (2020). The Surgeon General's Call to Action to Improve Maternal Health [Internet]. Washington (DC): US Department of Health and Human Services; 2020 Dec. 4, STRATEGIES AND ACTIONS: IMPROVING MATERNAL HEALTH AND REDUCING MATERNAL MORTALITY AND MORBIDITY. Available from: https://www.ncbi.nlm.nih.gov/books/NBK568218/

Ondersma, S. J., Beatty, J. R., Puder, K. S., Janisse, J., & Svikis, D. S. (2019). Feasibility and acceptability of e-screening and brief intervention and tailored text messaging for marijuana use in pregnancy. *Journal of Women's Health*, *28*(9), 1295-1301.

Petersen Williams, P., Jordaan, E., Mathews, C., Lombard, C., & Parry, C. D. (2014). Alcohol and other drug use during pregnancy among women attending midwife obstetric units in the Cape Metropole, South Africa. *Advances in preventive medicine*, *2014*.

Peto, T., Affron, D., Afrough, B., Agasu, A., Ainsworth, M., Allanson, A., & Lumley, S. F. (2021). COVID-19: Rapid antigen detection for SARS-CoV-2 by lateral flow assay: A national systematic evaluation of sensitivity and specificity for mass-testing. *EClinicalMedicine*, *36*, 100924.

Reitan, T. (2019). Substance abuse during pregnancy: A 5-year follow-up of mothers and children. *Drugs: Education, Prevention and Policy*, *26*(3), 219-228.

Seib, C. A., Daglish, M., Heath, R., Booker, C., Reid, C., & Fraser, J. (2012). Screening for alcohol and drug use in pregnancy. *Midwifery*, *28*(6), 760-764.

Schauberger, C. W., Newbury, E. J., Colburn, J. M., & Al-Hamadani, M. (2014). Prevalence of illicit drug use in pregnant women in a Wisconsin private practice setting. American Journal of Obstetrics and Gynecology, 211(3), 255-e1.

Stone, R. (2015). Pregnant women and substance use: fear, stigma, and barriers to care. Stone Health and Justice (2015) 3:2 DOI 10.1186/s40352-015-0015-5

Swift, A., Heale, R., & Twycross, A. (2020). What are sensitivity and specificity? *Evidence-Based Nursing*, *23*(1), 2-4.

Taplin, S., Richmond, G., & McArthur, M. (2014). Identifying alcohol and other drug use during pregnancy: Outcomes for women, their partners and their children.

Tavella, R. A., De Abreu, V. O., Muccillo-Baisch, A. L., & Da Silva Junior, F. M. (2020). Prevalence of illicit drug use during pregnancy: A global perspective. *Anais da Academia Brasileira de Ciências*, *92*.

Van Mastrigt, G. A., Hiligsmann, M., Arts, J. J., Broos, P. H., Kleijnen, J., Evers, S. M., & Majoie, M. H. (2016). How to prepare a systematic review of economic evaluation

Vearrier, D., Curtis, J. A., & Greenberg, M. I. (2010). Biological testing for drugs of abuse. *Molecular, clinical and environmental toxicology*, 489-517.

Vythilingum, B., Stein, D. J., Roos, A., Faure, S. C., & Geerts, L. (2012). Risk factors for substance use in pregnant women in South Africa. *South African Medical Journal*, *102*(11), 851-854.

White Haefele, B. (2012). An explorative study on the impact of prenatal methamphetamine (tik) abuse on early child and school behaviour. *Acta Criminological: African Journal of Criminology & Victimology*, *2012*(sed-1), 56-69.

Wong, W., Lau, S., Choi, E., Tucker, J. D., Fairley, C. K., & Saunders, J. M. (2019). A Systematic Literature Review of Reviews on the Effectiveness of Chlamydia

69

Testing. *Epidemiologic reviews*, *41*(1), 168–175.

https://doi.org/10.1093/epirev/mxz007

Wong, S., Ordean, A., Kahan, M., Gagnon, R., Hudon, L., Basso, M., & de la Ronde, S.
(2011). Substance use in pregnancy. *Journal of Obstetrics and Gynaecology
Canada*, *33*(4), 367-384.

World Health Organization. (2014). Guidelines for the identification and management of
substance use and substance use disorders in pregnancy.

Yonkers, K. A., Howell, H. B., Gotman, N., & Rounsaville, B. J. (2011). Self-report of illicit
substance use versus urine toxicology results from at-risk pregnant women. *Journal of
substance use*, *16*(5), 372-380.

Yonkers, K. A., Gotman, N., Kershaw, T., Forray, A., Howell, H. B., & Rounsaville, B. J.
(2010). Screening for prenatal substance use: development of the Substance Use Risk
Profile-Pregnancy scale. *Obstetrics and gynecology*, *116*(4), 827.

Zou, Z., Wang, H., d'Oleire Uquillas, F., Wang, X., Ding, J., & Chen, H. (2017). Definition
of substance and non-substance addiction. *Substance and Non-substance Addiction*,
21-41.

70

71

Keyword search strings

	1	2	3	4	5
1	Substance use/abuse	Substance use/abuse	Substance use/abuse	Substance use/abuse	Substance use/abuse
2	Maternal Substance use	Maternal Substance use	Maternal Substance use	Maternal Substance use	Maternal Substance use
3	Prenatal drug use	Prenatal drug use	Prenatal drug use	Prenatal drug use	Prenatal drug use
4	Screening tools	Screening tools	Screening tools	Screening tools	Screening tools
5	Meconium	Meconium	Meconium	Meconium	Meconium
6	Self-report	Self-report	Self-report	Self-report	Self-report
7	Drug toxicology	Drug Toxicology	Drug Toxicology	Drug Toxicology	Drug toxicology
8		Prenatal Toxicology Screening	Prenatal Toxicology Screening	Prenatal Toxicology Screening	Prenatal Toxicology Screening
9		Systematic Review	Systematic Review	Systematic Review	Systematic Review
10			Biological assessment	Biological assessment	Biological assessment
11				pregnancy	Pregnancy
12					Utility of screening tools

72

APPENDIX B
Evaluation Tool for Quantitative Research Studies

Building on work within a project exploring the feasibility of undertaking systematic reviews of research literature on effectiveness and outcomes in social care, a set of evaluation tools have been developed to assist in the critical appraisal of research studies. The evaluation tool for quantitative studies contains six sub-sections: study evaluative overview; study, setting and sample; ethics; group comparability and outcome measurement; policy and practice implications; and other comments. It provides a template of key questions to assist in the critical appraisal of quantitative research studies, (Long et al., 2004).

Review Area	Key Questions
(1) STUDY OVERVIEW	
Bibliographic Details	Author, title, source (publisher and place of publication), year
Purpose	1. What are the aims of the study? 2. If the paper is part of a wider study, what are its aims?
Key Findings	3. What are the key findings of the study?
Evaluative Summary	4. What are the strengths and weaknesses of the study and theory, policy, and practice implications?
(2) STUDY, SETTING, SAMPLE AND ETHICS	
The Study	5. What type of study is this? 6. What was the intervention? 7. What was the comparison intervention? 8. Is there sufficient detail given of the nature of the intervention and the comparison intervention? 9. What is the relationship of the study to the area of the topic review?
Setting	10. Within what geographical and care setting was the study carried out?

Sample	11. What was the source population?
	12. What were the inclusion criteria?
	13. What were the exclusion criteria?
	14. How was the sample selected?
	15. If more than one group of subjects, how many groups were there, and how many people were in each group?
	16. How were subjects allocated to the groups?
	17. What was the size of the study sample, and of any separate groups?
	18. Is the achieved sample size sufficient for the study aims and to warrant the conclusions drawn?
	19. Is information provided on loss to follow up?
	20. Is the sample appropriate to the aims of the study?
	21. What are the key sample characteristics, in relation to the topic area being reviewed?

(3) ETHICS

Ethics	22. Was Ethical Committee approval obtained?
	23. Was informed consent obtained from participants of the study?
	24. Have ethical issues been adequately addressed?

(4) GROUP COMPARABILITY AND OUTCOME MEASUREMENT

Comparable Groups	25. If there was more than one group that was analysed, were the groups comparable before the intervention? In what respects were they comparable and in what were they not?
	26. How were important confounding variables controlled (e.g., matching, randomisation, in the analysis stage)?
	27. Was this control adequate to justify the author's conclusions?
	28. Were there other important confounding variables controlled for in the study design or analyses and what were they?
	29. Did the authors take these into account in their interpretation of the findings?

74

Outcome Measurement	30. What were the outcome criteria?
	31. What outcome measures were used?
	32. Are the measures appropriate, given the outcome criteria?
	33. What other (e.g., process, cost) measures are used?
	34. Are the measures well validated?
	35. Are the measures known responsive to change?
	36. Whose perspective do the outcome measures address (professional, service, user, carer)?
	37. Is there a sufficient breath of perspective?
	38. Are the outcome criteria useful/appropriate within routine practice?
	39. Are the outcome measures useful/appropriate within routine practice?
Time Scale of Measurement	40. What was the length of follow-up, and at what time points was the outcome measurement made?
	41. Is this period of follow-up sufficient to see the desired effects?
(5) POLICY AND PRACTICE IMPLICATIONS	
Implications	42. To what setting are the study findings generalisable? (For example, is the setting typical or representative of care settings and in what respects?)
	43. To what population are the study's findings generalisable?
	44. Is the conclusion justified given the conduct of the study (For example, sampling procedure; measures of outcome used, and results achieved?)
	45. What are the implications for policy?
	46. What are the implications for service practice?
(6) OTHER COMMENTS	
Other Comments	47. What was the total number of references used in the study?
	48. Are there any other noteworthy features of the study?
	49. List other study references
Reviewer	50. Name of reviewer
	51. Review date

Source: Long AF, Godfrey M, Randall T, Brettle AJ and Grant MJ (2002) *Developing Evidence Based Social Care Policy and Practice. Part 3: Feasibility of Undertaking Systematic Reviews in Social Care*. Leeds: Nuffield Institute for Health.

APPENDIX C: Data extraction table template

Author	Title	Aim	Design	Sample	Location	Key Findings