

## Original Article

# Olfaction in pregnancy: systematic review and meta-analysis

Shaley L. Albaugh BS<sup>1,\*</sup>, Lisa L. Wu BS<sup>1</sup>, Douglas Zhang BS<sup>1</sup>, Ashley Diaz BS<sup>1</sup>,  
Debra A. Werner MLIS<sup>1,2</sup>, Jayant M. Pinto MD<sup>3</sup>, E. Leslie Cameron PhD<sup>4</sup>

<sup>1</sup>Pritzker School of Medicine, Division of Biological Sciences, The University of Chicago, Chicago IL 60637, USA

<sup>2</sup>The John Crerar Library, The University of Chicago, Chicago, IL 60637, USA

<sup>3</sup>Section of Otolaryngology-Head and Neck Surgery, Department of Surgery, The University of Chicago, Chicago, IL 60637, USA

<sup>4</sup>Department of Psychological Science, Carthage College, Kenosha, WI 53140-1994, USA

\*Corresponding author: Pritzker School of Medicine, The University of Chicago, Chicago, IL 60637, USA. Email: [Shaley.Albaugh@bsd.uchicago.edu](mailto:Shaley.Albaugh@bsd.uchicago.edu)

Little attention has been paid to olfactory changes during pregnancy with contemporary studies limited in number and sample size. We examined whether pregnancy is associated with differences in olfactory performance and if there were any specific gestational ages at which these differences occur through a comprehensive systematic review and meta-analysis of the current literature. An initial electronic database search identified 234 citations, which were screened at the abstract level. Twenty-three citations were germane for full-text review, and 13 met criteria for inclusion. Our review assessed 5 olfactory measures of interest: odor identification ( $n = 11$  articles), threshold ( $n = 8$ ), discrimination ( $n = 5$ ), hedonics ( $n = 6$ ), and intensity ( $n = 5$ ). Nine of these 13 studies contained sufficient data for meta-analysis, and these studies included a total of 523 pregnant women and 365 non-pregnant controls. Despite previous subjective and objective reports of odor intolerances and odor hypersensitivity, we did not find any significant differences between pregnant and non-pregnant women in odor discrimination, thresholds, or hedonics. However, meta-analysis of 506 cases and 333 controls showed worse odor identification in pregnant women compared to controls in a random-effects model. Thus, we demonstrate worse performance at odor identification during pregnancy. In this review, we discuss the current evidence (and lack thereof) regarding olfaction in pregnancy as well as highlight current knowledge gaps in this field.

**Key words:** cognition, hyposmia, hyperosmia, odor identification, odor hedonics, odor thresholds

## Introduction

The sense of smell is vital for environmental hazard detection (Stevenson 2010) and along with taste and other senses influences eating behavior and nutrition (Beauchamp and Mennella 2011; Boesveldt and de Graaf 2017). Olfaction may help mothers avoid harmful substances, such as spoiled food containing pathogens or toxins, and focus on calorie dense nourishment needed for the increasing metabolic demands of pregnancy. Therefore, changes in olfactory perception during pregnancy may affect the health of the mother and fetus (Boesveldt and Parma 2021). However, despite these important implications, little attention has been paid to olfactory changes during pregnancy with contemporary studies limited in number and sample size.

A small number of studies have examined self-reported odor changes in pregnancy and have shown that as many as two-thirds of healthy women report enhanced sense of smell or hyperosmia (Cameron 2007), abnormal sensitivity (Nordin et al. 2004), or specific sensitivity to one or more odors (Nordin et al. 2007; Cameron 2014a). With self-evaluation of olfactory function considered unreliable (Landis et al. 2003; Lötsch and Hummel 2019), researchers have focused on objective olfactory assessments such as the University of Pennsylvania Smell Identification Test (UPSIT) (Doty et al.

1984) or “Sniffin’ Sticks” test (SST) (Hummel et al. 1997). To date, no studies have provided evidence for enhanced olfactory function in pregnancy with these objective tools, but instead report that healthy pregnant women (Savović et al. 2002; Ochsenbein-Kölblle et al. 2007; Nwankwo et al. 2017; Şimşek et al. 2021) and women with hyperemesis gravidarum (Yasar et al. 2016; Tan et al. 2020; Şimşek et al. 2021) experience olfactory impairment or hyposmia, while other studies report no differences in any olfactory measure between pregnant and non-pregnant women (Gilbert and Wysocki 1991; Laska et al. 1996; Kölblle et al. 2001; Cameron 2007, 2014b; Kyung-yeon 2014; Fornazieri et al. 2019). Overall, studies examining olfaction and pregnancy report conflicting results with little consensus.

To overcome this problem of conflicting information, we pooled data to investigate potentially novel insights. Through a comprehensive systematic review and meta-analysis of the current literature, we examined olfactory function in pregnancy across different types of olfactory measures, including odor threshold (OT), identification (OID), discrimination (OD), hedonics (OH) and intensity (OI). Our primary aim was to explore whether pregnancy is associated with differences in olfactory performances across these measures, while our secondary aim was to identify other potential moderating factors.

## Methods

### Literature search strategy

We performed a comprehensive literature search across PubMed, Scopus, Web of Science, and Google scholar databases of the English language. We examined human olfaction studies on several occasions from inception of the study until February 2022 following the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines (Liberati et al. 2009). Search terms included and related to “olfaction,” “smell,” and “pregnancy.” The full database strategy can be found in [Supplementary Table 1](#). Additional studies were identified by performing manual reference list searches in PubMed.

### Inclusion and exclusion criteria

Three authors (SA, LW, and DZ) independently screened all study titles and abstracts for eligibility using the Rayyan screening tool (Ouzzani et al. 2016). This screening phase sought to identify all studies providing information on olfaction and pregnancy using either objective olfaction instruments or quantitative measures. Outcomes of interests included different olfactory abilities, such as (i) OID, (ii) OT, (iii) OD, (iv) OH, and (v) OI. Inclusion criteria were: healthy, pregnant subjects  $\geq 18$  years of age with at least one outcome of interest measured numerically with an olfaction instrument. Articles were also included in the full-text screening phase if the method was unclear from the abstract. Articles were excluded if they reported only qualitative olfactory information or included patients with known recognized causes of olfactory dysfunction such as those with nasal or sinus disease, history of head injury, or neurocognitive disorders. There were no restrictions on trimester or parity. Reviews, abstracts, case reports, studies not available in the English language, unavailable studies, and studies with insufficient data were excluded. Disagreements among screeners were resolved through discussion to obtain consensus.

### Data extraction and quality assessment

Three researchers (SA, LW, DZ) independently extracted information, which included the following: author, publication year, study design, number of cases and controls, gestational trimester, olfaction test method(s), outcome means with standard deviations (SD) and main findings. Outcome data that were only presented graphically were extracted using WebPlotDigitizer (Rohatgi 2021). In addition, all three researchers evaluated included full-text documents using a modified Newcastle-Ottawa Scale (NOS) for cohort and case-control studies to assess methodological quality (Wells et al. 2000). The modified NOS has six criteria and the possibility of obtaining up to eleven stars if all criteria are met. The parameters assessed are in three categories: selection of the study groups, comparability of the groups, and the assessment of the outcome of interest. The scores range from 0 to 11, with “good” quality studies possessing scores from 7 to 11, “fair” quality studies possessing scores from 4 to 6, and “poor” quality studies possessing scores from 0 to 3. Further breakdown of the scoring for each study can found in [Supplementary Table 2](#).

### Statistical analysis

Meta-analyses were conducted using the *metafor* package in R (Viechtbauer 2010). The effect size in our analysis was the

standardized mean difference (SMD), which we calculated from the reported means and SDs. When unavailable, mean and SD were estimated from medians and ranges according to the proposed method by Hozo and colleagues (2005). Effect sizes were calculated for an olfactory task if two or more studies reported data. The heterogeneity (e.g. between-study variance) across analyses was calculated using  $I^2$  static, with heterogeneity quantified as a percentage. Low, moderate, and high heterogeneity correspond to 25%, 50%, and 75%, respectively.

For cross-sectional studies with trimester subgroups, these subgroups were aggregated into a single group (“collapsed trimester” group) by calculating a weighted average mean (i.e. mean regardless of trimester) and SD for the different types of olfactory data. We did this to account for the limited sample size for certain trimesters and to compare pregnant women to non-pregnant women regardless of trimester status. For longitudinal studies examining trimester subgroups, the first recorded measurement was used for collapsed analysis of a single pregnant group. Along with subgroups analyses comparing coefficients with the Wald-type test ( $Q_M$ ) (Borenstein et al. 2021), univariate multilevel meta-regressions were conducted to assess the influence of potentially relevant moderators on the effect size of olfactory tasks. Forest plots were used to summarize the results of included studies. Publication bias was assessed by funnel plot visualization as there were too few studies to use Egger’s regression test for funnel plot asymmetry (Page et al. 2022). All analyses were performed using R version 4.1.1 (R Core Team 2022). A  $P$ -value of  $< 0.05$  was considered statistically significant for all analyses.

## Results

### Study selection

[Figure 1](#) shows the identification and selection process for our included studies. The initial electronic database search identified 234 relevant abstracts, of which 23 manuscripts were included for full-text review. Of these, only 13 met inclusion criteria, with 10 records removed because they did not contain quantitative data of olfactory function ( $n = 7$ ), our population of interest ( $n = 1$ ), or our outcomes of interest ( $n = 2$ ; details in supplementary information). Sufficient data were available in 9 of these 13 included studies for meta-analysis. The 9 studies included a total of 523 pregnant women and 365 non-pregnant controls.

### Study characteristics

Full-text screening revealed 5 main areas of interest in the included nine studies: OID ( $n = 11$ ), OT ( $n = 8$ ), OD ( $n = 5$ ), OH ( $n = 6$ ), and OI ( $n = 5$ ). The studies that assessed OID, OT, OD, and OH yielded data amenable to meta-analysis. Studies that examined OI lacked comparable data. Most studies were fair quality ( $n = 8$ ), with five studies being good quality, and no studies being poor quality ([Supplementary Table 2](#)). Details of each of the included studies are summarized in [Table 1](#).

The tools used for quantitative assessment of olfactory function varied among the included studies, and only select instruments were utilized in multiple studies: 4 studies used the SST smell test battery of 16 odorants (Kölble et al. 2001; Ochsenein-Kölble et al. 2007; Kyung-yeon 2014; Nwankwo et al. 2017), 3 used the UPSIT (Dastur 2000; Cameron 2007;

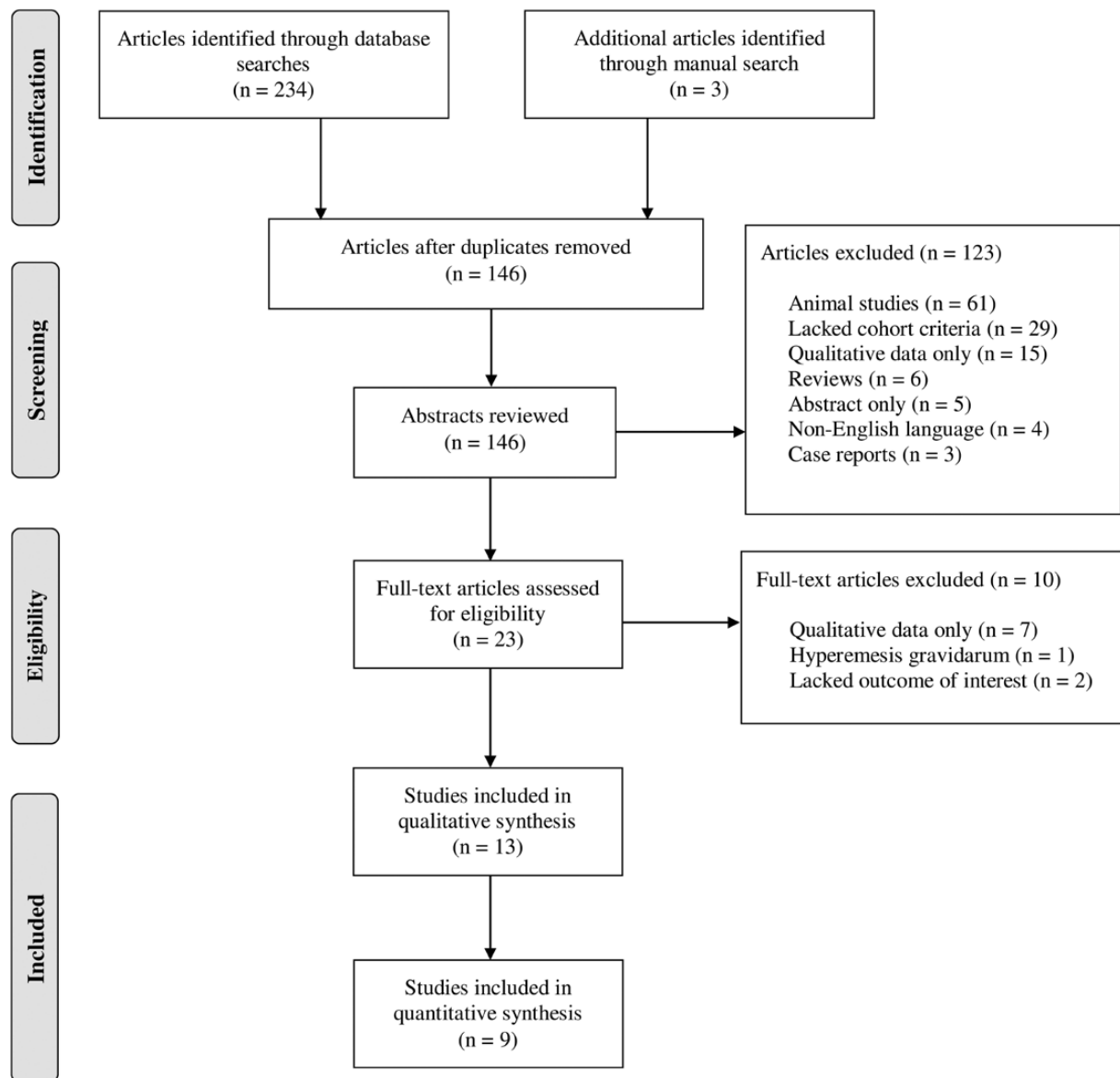


Fig. 1. PRISMA flow chart of olfaction in pregnancy study selection process.

Fornazieri et al. 2019) and 2 studies (Dastur 2000; Cameron 2014b) used the phenyl ethyl alcohol (PEA) smell threshold test with the same staircase protocol (Doty 2000).

### Sample characteristics

Among the 9 cross-sectional studies, only 2 did not specify gestational ages (Gilbert and Wysocki 1991; Cameron 2014b). All 5 longitudinal studies assessed women at each trimester and in the postpartum period, for a total of four assessments (Laska et al. 1996; Dastur 2000; Ochsenein-Kölble et al. 2007; Cameron 2014b; Kyung-yeon 2014). In addition to gestational information, the included samples of the 13 studies were from 9 different countries.

### Odor identification

Eight of the included studies provided data appropriate for meta-analysis investigating OID (Dastur 2000; Kölble et al. 2001; Cameron 2007; Ochsenein-Kölble et al. 2007;

Kyung-yeon 2014; Şimşek et al. 2015; Nwankwo et al. 2017; Fornazieri et al. 2019). These studies included a total of 506 cases and 333 controls.

Meta-analysis using these 8 studies showed worse OID in the collapsed trimester group compared to controls in a random-effects model ( $k = 8$ , SMD,  $-0.32$  [95% CI,  $-0.54$  to  $-0.09$ ];  $P = 0.05$ ;  $I^2 = 48.88\%$ ; see Fig. 2). Sensitivity analysis examining only cross-sectional studies separately showed similar results ( $k = 5$ , SMD,  $-0.45$  [95% CI,  $-0.70$  to  $-0.19$ ];  $P < 0.001$ ;  $I^2 = 41.52\%$ ; data not shown). On subgroup analyses with between-group comparison, trimester did not significantly influence effect size with respect to OID ( $Q_M = 0.25$ ,  $df = 2$ ,  $P = 0.88$ ). Also, while the effect size did not significantly depend on the type of olfactory test in comparisons of all 3 types ( $Q_M = 4.03$ ,  $df = 2$ ,  $P = 0.13$ ), use of the SST was associated with a larger effect size compared to use of the UPSIT ( $Q_M = 15.44$ ,  $df = 2$ ,  $P < 0.004$ ). There was no strong evidence of publication bias based on funnel plot visualization (Supplementary Fig. 1).



**Table 1.** Continued

Author, year	Study design	Country	Cases/controls <sup>a</sup>	Age (mean ± SD, years)		Gestational age studied (n)			Olfactory assessment tool(s)	Main finding(s)	Quality <sup>b</sup>	
				Cases	Controls	1T	2T	3T				PP
Kölbl et al. (2001) <sup>c</sup>	Cross-sectional	Switzerland	53/59	28.0 ± 5.6	32.0 ± 4.7	53	0	0	0	SST <sup>e</sup> (OI, OD, OID) VAS ratings of 10 “natural odors” (OI, OH)	Pooled sensitivity scores (i.e. sum of OT, OD, and OID scores) did not differ significantly between cases and controls. OID was significantly lower in cases than controls. OT, OD, and OI did not differ significantly between cases and controls. Odor-specific differences found with regards to OH (cigarette, rum and coffee)	Good
Savović et al. (2002)	Cross-sectional	Serbia	20/20	*	*	20	0	0	0	Olfactometer (OT, OID)	Cases had lower OT and OID, but difference non-significant	Fair
Olofsson et al. (2005)	Cross-sectional	Sweden	15/15	30.7 ± 4.7	31.1 ± 4.3	0	15	0	0	CSERP	Reported hypersensitivity in cases may be due to cognitive changes, not changes in sensory processing. Differences were non-significant ( <i>P</i> -value > 0.05)	Fair
Cameron (2007) <sup>c</sup>	Cross-sectional	United States	60/20	29.7	25.9	20	20	20	20	UPSIT (OID, OI, OH)	OID, OI, and OT did not differ between cases and controls, but odor-specific changes in OID, OI, and OH were reported. Cases in their 1T reported higher overall OI and lower OT, but these differences were found to be statistically non-significant in a 1-way ANOVA	Fair

Table 1. Continued

Author, year	Study design	Country	Cases/controls <sup>a</sup>	Age (mean ± SD, years)		Gestational age studied (n)			Olfactory assessment tool(s)	Main findings(s)	Quality <sup>b</sup>
				Cases	Controls	1T	2T	3T			
Ochsenbein-Kölbl et al. (2007) <sup>a</sup>	Longitudinal	Switzerland	38/46	29.0 ± 5.7	32.0 ± 4.7	38	38	38	38	Cases in their 3T and PP had significantly lower TDI scores compared to controls. OT was significantly lower (lower sensitivity per authors) in pregnant women in the 3T and PP compared to controls. OID did not change significantly across sessions and did not differ compared to controls. Pregnant individuals rated clove as significantly less pleasant than controls. Among pregnant women, acetic acid was rated as more pleasant during the 2T and 3T, but these ratings did not differ between pregnant women and controls when considering all four sessions. Coffee was rated significantly more unpleasant by pregnant women than controls during the 1T, but was not rated significantly more unpleasant in the 2T, 3T, or PP compared to controls	Good
Cameron (2014a)	Cross-sectional	United States	17/32	25	20	•	•	•	•	There was no significant difference in OT measurements between cases and controls	Fair
Cameron (2014b)	Longitudinal	United States	20/22	31	26	20	20	20	0	There were no significant differences in OT measurements between trimesters or between cases and controls. There were also no significant differences in sensitivity or response bias between trimesters or between cases and controls. However, there was a tendency for more response bias earlier in pregnancy	Fair

Table 1. Continued

Author, year	Study design	Country	Cases/controls <sup>a</sup>	Age (mean ± SD, years)		Gestational age studied (n)			Olfactory assessment tool(s)	Main finding(s)	Quality <sup>b</sup>	
				Cases	Controls	1T	2T	3T				PP
Kyung-yeon (2014) <sup>c</sup>	Longitudinal	Korea	50/40	30.3 ± 3.3	28.5 ± 6.6	50	*	29	SST <sup>d</sup> (OT, OD, OID)	No significant differences in OT, OD, OID or TDI	Good	
Şimşek et al. (2015) <sup>c</sup>	Cross-sectional	Turkey	92/30	27.4 ± 5.3	26.0 ± 4.2	31	30	31	0	B-SIT <sup>e</sup> (OID)	1T cases had significantly lower median identification scores compared to 2T, 3T, and controls. 1T women identified leather, pine, and soot significantly less well than controls. Both 2T and 3T OID scores did not differ significantly from those of controls	Good
Nwankwo et al. (2017) <sup>c</sup>	Cross-sectional	Nigeria	70/70	30.5 ± 3.9	28.5 ± 6.6	*	*	*	*	SST (OT, OD, OID)	OID and TDI were significantly lower in cases compared to controls	Good
Fornazieri et al. (2019) <sup>c</sup>	Cross-sectional	Brazil	124/50	26.4 ± 6.26	26.4 ± 5.0	47	33	44	32	UPSIT <sup>h</sup> (OID) VAS ratings of 4 common UPSIT odors (OH, OI)	No difference in OID between 1T, 2T, 3T, or and controls. There were several significant odor-specific differences ( <i>P</i> -value < 0.05) in OID, OH, and OI between various groups	Good

SD, standard deviation; 1T, 1st Trimester; 2T, 2nd Trimester; 3T, 3rd Trimester; PP, postpartum; OT, odor threshold; OD, odor discrimination; OID, odor identification; OI, odor intensity; OH, Odor Hedonics; TDI, sum of the OT, OD, and OID scores as measured using Sniffin' Sticks; VAS, Visual Analogue Scale; SST, Sniffin' Sticks; UPSIT, University of Pennsylvania Smell Identification Test; PEA, Phenyl Ethyl Alcohol; B-SIT, Brief Smell Identification Test; CSERP, chemosensory event-related potentials.

\*Not specified.

<sup>a</sup>All cases were pregnant women, and all controls listed were healthy, non-pregnant women.

<sup>b</sup>Quality was determined using a modified Newcastle-Ottawa Scale. Scores include good, fair, and poor.

<sup>c</sup>Included in meta-analysis.

<sup>d</sup>PEA threshold test performed as outlined in Doty et al. (2000).

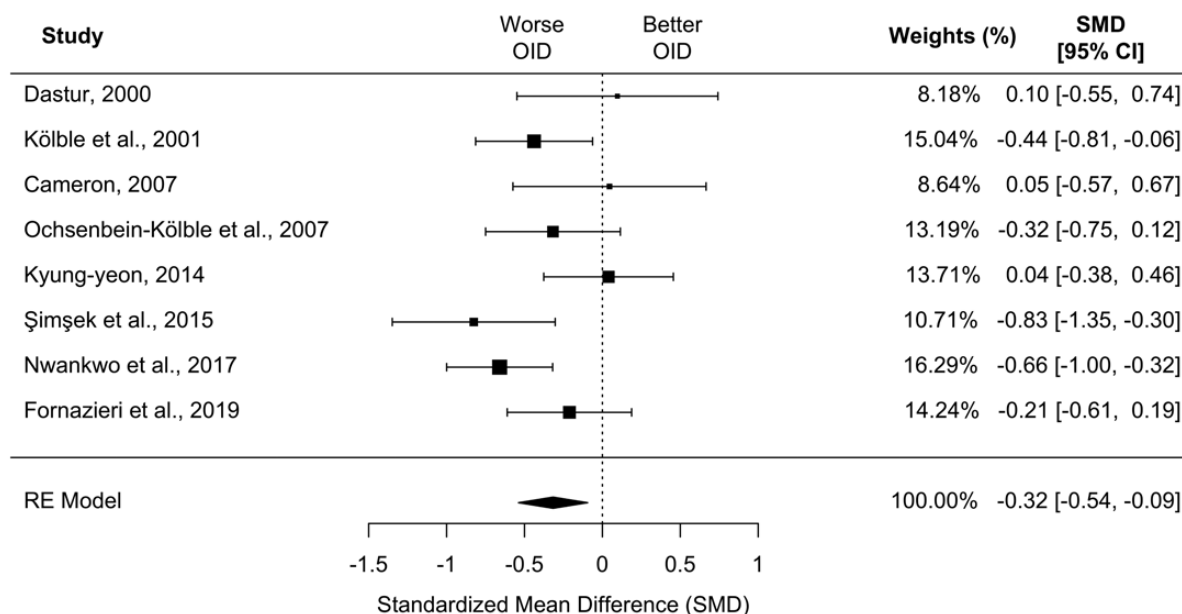
<sup>e</sup>All Sniffin's Sticks tests included the version with 16 odorants.

<sup>f</sup>Korean adaptation of the SST.

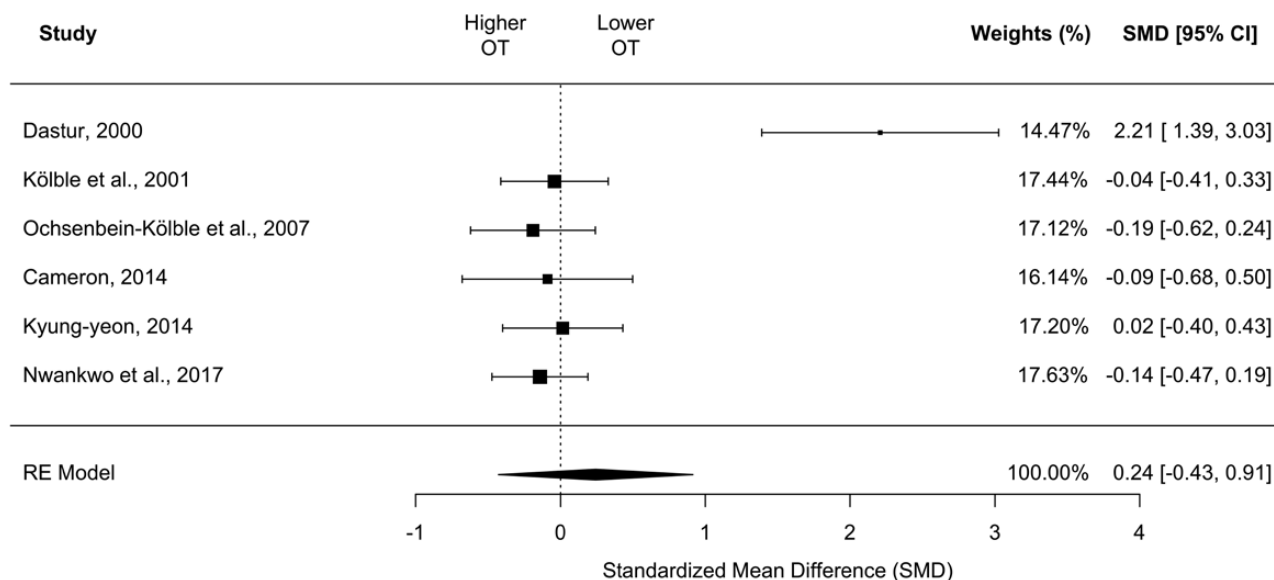
<sup>g</sup>Turkish adaptation of the B-SIT.

<sup>h</sup>Portuguese adaptation of the UPSIT.





**Fig. 2.** Forest plot depicting the standardized mean difference for odor identification in the collapsed, single pregnant group compared to controls. CI, confidence interval; OID, odor identification. Size of black square box represents relative sample size.



**Fig. 3.** Forest plot depicting the standardized mean difference for odor detection threshold in the collapsed, single pregnant group compared to controls. CI, confidence interval; OT, odor threshold. Size of black square box represents relative sample size.

### Odor detection threshold

Six of the included studies were appropriate for meta-analysis investigating OT (Dastur 2000; Kölbl et al. 2001; Ochsenbein-Kölbl et al. 2007; Cameron 2014b; Kyung-yeon 2014; Nwankwo et al. 2017). These studies included a total of 274 pregnant cases and 265 non-pregnant controls.

Meta-analysis using these six studies showed no differences in OT between the collapsed trimester group and the controls in a random-effects model ( $k = 6$ , SMD, 0.19 [95% CI, -0.23 to 0.61];  $P = 0.38$ ;  $I^2 = 81.38\%$ ; see Fig. 3). On subgroup moderator analyses, SMDs significantly differed with the type of olfactory test used ( $Q_M = 9.54$ ,  $df = 1$ ,  $P = 0.002$ ), with the SST (which measures threshold

to *n*-butanol) demonstrating a larger effect size compared to PEA threshold test. However, there were no differences by trimester when examining the two studies that contained data across gestational ages ( $Q_M = 3.44$ ,  $df = 4$ ,  $P = 0.49$ ). Whereas funnel plot asymmetry was observed, the number of studies is too small to infer publication bias (Supplementary Fig. 2) (Sterne et al. 2011).

### Odor discrimination

Four of the included studies were appropriate for meta-analysis investigating OD and all used SST (Kölbl et al. 2001; Ochsenbein-Kölbl et al. 2007; Kyung-yeon 2014; Nwankwo



et al. 2017). These studies included a total of 211 pregnant cases and 215 non-pregnant controls.

Meta-analysis using these 4 studies showed no differences in OD between the collapsed trimester subgroup and controls in a fixed-effect model ( $k = 4$ , SMD, 0.004 [95% CI, -0.19 to 0.19];  $P = 0.96$ ;  $I^2 = 0.00\%$ ; see Fig. 4). Sensitivity analyses examining only 1st trimester cases showed similar results ( $k = 3$ , SMD, 0.05 [95% CI, -0.19 to 0.28];  $P = 0.71$ ;  $I^2 = 0.00\%$ ). We were unable to compare other trimesters or perform moderator analyses. There was no strong evidence of publication bias based on funnel plot visualization (Supplementary Fig. 3).

### Odor Hedonics

The six studies that quantified OH varied in the instruments they used for assessment. Three studies utilized a similar instrument, a visual-analog scale (VAS) (Kölble et al. 2001; Ochsenbein-Kölble et al. 2007; Fornazieri et al. 2019). Two of the three studies examined the same 10 “natural odors” (Kölble et al. 2001; Ochsenbein-Kölble et al. 2007). The third study only studied 2 of the 10 “natural odors” resulting in 2 total overlapping odorants (coffee and clove) that were studied in all three studies that quantified OH using a VAS. However, only 2 studies contained sufficient data for one odorant, coffee, for analysis (Ochsenbein-Kölble et al. 2007; Fornazieri et al. 2019). These studies included a total of 85 pregnant cases and 96 non-pregnant controls.

Meta-analysis showed no differences in pleasantness of coffee between the single pregnant subgroup and controls in a random-effects model ( $k = 2$ , SMD, -0.70 [95% CI, -2.64 to 1.23];  $P = 0.48$ ;  $I^2 = 97.25\%$ ; see Fig. 5). Sensitivity analyses examining only 1st trimester cases showed similar results ( $k = 2$ , SMD, -1.06 [95% CI, -3.12 to 1.00];  $P = 0.31$ ;  $I^2 = 97.38\%$ ). Finally, on subgroup moderator analyses, the effect size was not influenced by gestational age ( $Q_M = 0.39$ ,  $df = 4$ ,  $P = 0.98$ ; data not shown). There was no strong

evidence of publication bias based on funnel plot visualization (Supplementary Fig. 4).

### Odor intensity

The five studies that examined OI varied in the instruments they used for measurement. Only 2 studies examined ratings of 10 “natural odors” using a VAS (Kölble et al. 2001; Ochsenbein-Kölble et al. 2007), and 2 other studies utilized UPSIT odors, but they employed different strategies and used a different number of odors (Cameron 2007; Fornazieri et al. 2019). Despite similar instruments, the studies were too heterogeneous for meaningful meta-analysis.

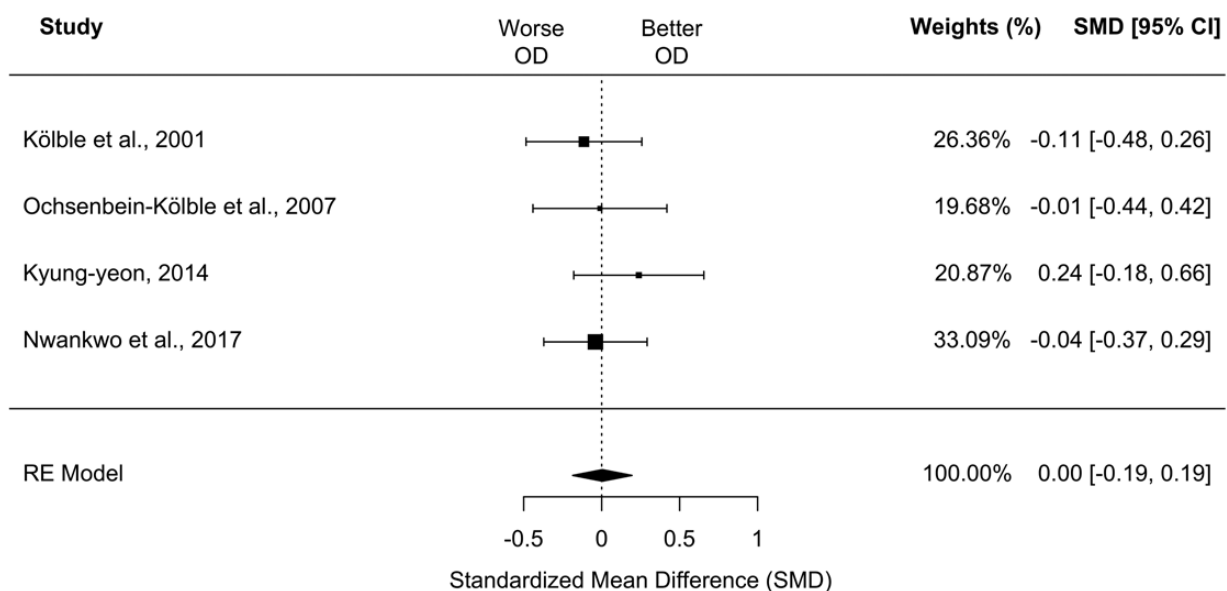
## Discussion

### General findings

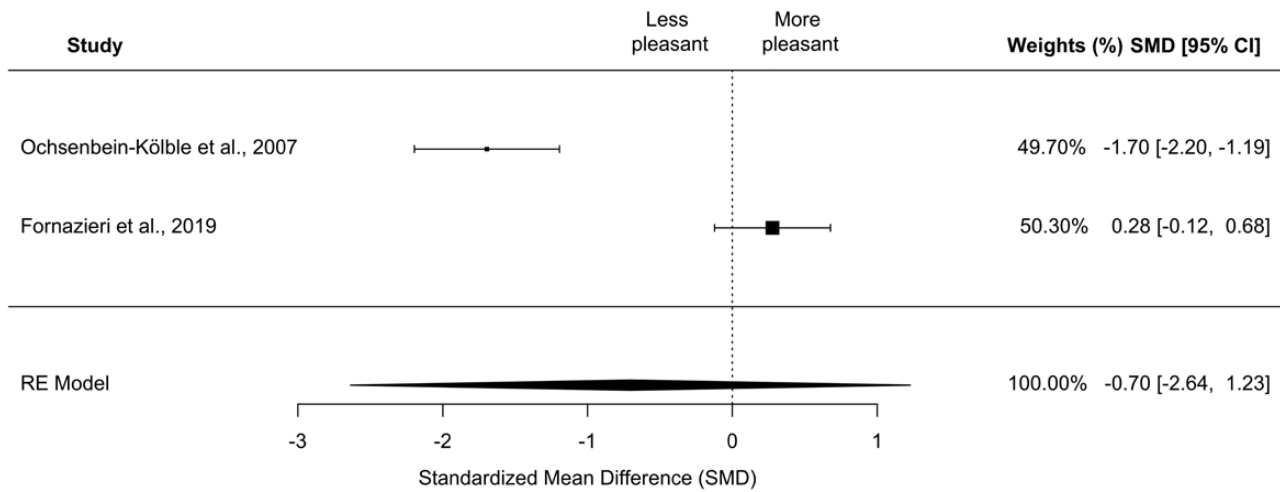
Our study is the first to explore olfaction in pregnancy with a systematic review and meta-analysis including data from different time points in pregnancy and from diverse populations. We primarily quantitatively examined four different types of olfactory measures: identification, detection, discrimination, and hedonics. Despite the subjective reports and previous studies that have reported odor intolerances and odor hypersensitivity in pregnancy, we did not find enhancement in sense of smell when measured objectively. Most of the studies included in our meta-analysis examined OID, and we found that pregnant women were worse at OID compared to controls. We did not find any trimester-specific effects on ID.

### Odor identification and cognition in pregnancy and the postpartum period

The exact mechanisms responsible for impaired odor identification in pregnancy remain unknown, but one hypothesis is that olfactory impairment may be related to the cognitive changes observed during pregnancy. Compared to discrimination and sensitivity, olfactory identification is thought to be



**Fig. 4.** Forest plot depicting the standardized mean difference for odor discrimination in the collapsed, single pregnant group compared to controls. CI, confidence interval; OD, odor discrimination. Size of black square box represents relative sample size.



**Fig. 5.** Forest plot depicting the standardized mean difference for odor hedonics of coffee in the collapsed, single pregnant group compared to controls. CI, confidence interval. Size of black square box represents relative sample size.

a higher order olfactory task that draws on semantic memory, verbal fluency, and executive function (Hedner et al. 2010), all of which have been shown in recent meta-analyses to be impaired in pregnancy (Henry and Rendell 2007; Davies et al. 2018). Furthermore, a converging body of evidence suggests that the decreased cognitive functioning seen during pregnancy may be a short-term effect related to the influence of unprecedented hormonal fluctuations on brain architecture (Hoekzema et al. 2017; Barth and de Lange 2020; Rehbein et al. 2022). Multiple studies have shown associations in reduced grey matter volumes, decreased cognition, and the hormonal variations during pregnancy (Hoekzema et al. 2017; Rehbein et al. 2022), with some of these grey matter changes lasting up to 2 years (Hoekzema et al. 2017). Additionally, some of the identified brain regions in these studies that are affected by pregnancy have considerable overlap with those involved in central olfactory processing, such as the amygdala, entorhinal cortex, and cingulate cortex (Soudry et al. 2011). These dynamic structural and functional changes may prepare women for maternal behavior, such as fostering maternal-infant attachment, recognizing the needs of the infant, and responding to threats to the infant's wellbeing (Barba-Müller et al. 2019). We hypothesize that higher-odor cognitive processing, which includes odor identification, may be inadvertently affected by this powerful neuroplastic remodeling. Whether there are any evolutionary benefits of these changes during pregnancy remains to be elucidated.

Furthermore, growing evidence also shows that changes in the maternal brain extend into the postpartum period. Researchers have observed that instead of the former decreased grey matter volume seen during active pregnancy, there is increased grey matter volume in large areas such as the prefrontal cortex, parietal lobe, and midbrain during the postpartum period (Kim et al. 2010). Some of the brain volume lost during pregnancy may be recovered after birth (Barba-Müller et al. 2019). Related to this, we did not find worse OID in the postpartum period, potentially indicating that the recovery of OID may be linked to these postpartum brain changes.

### Odor sensitivity

Over the past 3 decades, researchers have failed to consistently demonstrate objective sensitivity differences between

pregnant women and non-pregnant controls. The first study to examine olfactory detection thresholds longitudinally and quantitatively across pregnancy, Laska et al. (1996) used an ascending staircase procedure with *n*-butanol and found that thresholds in 20 pregnant women were higher (worse odor detection) in the 1st trimester but lower (better odor detection) in the 3rd trimester compared to 20 controls. Contrary to this early study, another more recent study examined 20 pregnant and 22 non-pregnant women longitudinally and 17 pregnant and 32 non-pregnant women cross-sectionally using a single staircase procedure with PEA and found no differences in detection thresholds (Cameron 2014b). Only 2 other studies have reported differences in olfactory detection thresholds: one reported decreased detection thresholds only from the 35th week of pregnancy until the postpartum period using a staircase procedure with *n*-butanol dilutions presented in Sniffin' Sticks in a geometric series (Ochsenbein-Kölbl et al. 2007), and the other, an unpublished report, found decreased detection thresholds throughout all 3 trimesters of pregnancy with the greatest effect seen in the 1st trimester using a single staircase procedure with PEA presented in glass jars (Dastur 2000).

In agreement with the lack of consistent OT-related findings in the current literature, our moderator analysis of trimester effects on olfactory detection thresholds did not show any difference by trimester or during the postpartum period with pooled data. However, we believe there are important explanations as to why we and other researchers have not detected differences. First and foremost, the monomolecular odor stimuli used to assess odor detection thresholds are likely not contextually relevant to pregnant women. Both non-food odors, the pleasant rose-like odor PEA and the non-familiar, "sweet" alcohol *n*-butanol, were used in the studies included in this meta-analysis, even though a majority of pregnant women report they are hypersensitive to unpleasant odors, particularly food odors (Nordin et al. 2004; Cameron 2007, 2014b). Since olfactory changes may theoretically prevent soon-to-be mothers from ingesting toxic substances (Profet 1988) and directing them to nutrient dense foods, odor sensitivity is likely selective and not generalizable across all odors. Given that neither past research with standardized and objective odor detection threshold testing nor our meta-analysis

demonstrate sensitivity differences, future research should focus on developing relevant monomolecular or multi-molecular odor stimuli to assess odor sensitivity in pregnancy.

### Odor intensity

Although several studies have explored pregnant women's ratings of the intensity of odors (Gilbert and Wysocki 1991; Laska et al. 1996; Kölblle et al. 2001; Cameron 2007; Ochsenein-Kölblle et al. 2007), data collection methods were too heterogeneous to support a meta-analysis. Given the heterogeneity of data collection and that only small differences were observed in the few studies that have examined intensity ratings of a range of odors, future research should explore ratings of odors that are at and above threshold. Such data may reveal ways in which pregnant women's experience with odors differs from non-pregnant women and may influence their behavior, particularly in food selection and enjoyment. Odor intensity and their hedonics, discussed below, may contribute to the nausea and vomiting that afflicts many pregnant women.

### Odor Hedonics

Odor hedonics, or pleasantness, might be the most important dimension of odors for pregnant women. This is the one dimension on which self-report and objective measures of olfaction are in accord. Pregnant women's self-report and their ratings of odor pleasantness both reflect an overall decrease in the pleasantness of most odors (Cameron 2014b). Multiple studies have found that there is at least a tendency for pregnant women to rate many odors as less pleasant than non-pregnant controls do (Gilbert and Wysocki 1991; Laska et al. 1996; Kölblle et al. 2001; Swallow et al. 2005; Cameron 2007; Ochsenein-Kölblle et al. 2007). However, these studies varied in both the method used to measure hedonics and, in the odors, presented.

Unfortunately, given the lack of a standardized test method for measuring odor hedonics, it was difficult for us to evaluate the quality of published data and to compare differences across studies using meta-analysis. Given this limitation, while previous studies have observed changes in pregnant individuals' tolerance of odorants such as coffee, cloves, cigarettes, rum, and acetic acid, our meta-analysis was limited to two samples and one odor (coffee). Our analysis did not reveal a significant difference in hedonic ratings between pregnant and non-pregnant women.

While historically a variety of subjective reporting methods have been used to evaluate odor hedonics, it would be beneficial for research to employ similar assessment tools to further explore odor hedonics in pregnancy. Moreover, the odors employed in most studies of olfaction in pregnancy have not been selected based on knowledge about pregnant women's experience with odors, but rather are based on standard odors used in testing human olfaction. Pregnancy may not affect the experience of all odors equally. For example, there is no readily apparent reason that odors such as phenyl ethyl alcohol (a rose odor), commonly used to test human OT, would be processed differently by pregnant women. Rather, pregnant women might be more sensitive to or be better able to identify odors such as coffee (as explored in our meta-analysis, but with only two studies from a single research group), which many women avoid consuming during pregnancy.

Furthermore, hedonic ratings could be used to select odors that many pregnant women rate as particularly pleasant or unpleasant that could then be used in objective olfactory tests, such as detection, discrimination, and identification.

### Limitations

The present study has several limitations. First, the number of included studies is small ( $n = 13$ ) as are their sample sizes, emphasizing that olfaction in pregnancy is an understudied topic. Furthermore, different olfactory tools were used. Both the UPSIT and SST were commonly employed, as well as different culturally adapted versions of these olfactory tools (Kyung-yeon 2014; Şimşek et al. 2015; Fornazieri et al. 2019). The use of the same type and version of olfaction tools would likely decrease the heterogeneity and enhance the reliability of our results. While the inclusion of ethnically and culturally diverse populations support high external validity for our findings and trends, cultural factors such as the familiarity or lack of familiarity with certain odors can influence the scores on these tests (Shu et al. 2007), which make the culturally adapted tools more in line with normative data (Doty 1995). One of our studies used a non-adapted tool (Nwankwo et al. 2017) that likely influenced the scores in that population.

Next, we noted limited and varying documentation of potential confounders important to the interpretation of olfactory performance results. For instance, three studies merely requested participants to refrain from smoking before testing (Kölblle et al. 2001; Ochsenein-Kölblle et al. 2007; Nwankwo et al. 2017), while only 6 of the 13 studies explicitly excluded women who smoked (Laska et al. 1996; Dastur 2000; Olofsson et al. 2005; Cameron 2007, 2014b; Şimşek et al. 2015), and no studies gathered data on depression. Smoking and depression have been reported to be associated with decreased olfactory performance (Kohli et al. 2016; Ajmani et al. 2017) and it should be noted by researchers whether participants were screened for these factors or not. Two other factors potentially important to olfaction in pregnancy that were not overtly mentioned in the studies were race/ethnicity (Pinto et al. 2014), and cognition (Yahiaoui-Doktor et al. 2019).

### Conclusions and future directions

Despite anecdotal evidence and survey data, our meta-analyses of studies that used objective olfactory tools revealed hyposmia, not hyperosmia, during pregnancy. This was only the case for odor identification; however, no differences were observed for other olfactory measures. Given the limited number of studies available for analysis and the range of olfactory tools and odors employed in those studies, these results should be interpreted with caution. There is a need for more studies at all stages of pregnancy and postpartum, with more diverse samples and with odors that are contextually relevant. Future studies should carefully select food or other odors that have established pleasantness and intensity ratings to better characterize odorant-specificity in pregnant women. Furthermore, studies could also focus on understanding the differences in sensory and cognitive changes in the olfactory system by administering cognitive test batteries along with olfactory testing. Finally, measuring olfactory performance and hormone levels concomitantly could help elucidate potential mechanisms by which the hormones present during pregnancy impact olfaction.

## Supplementary material

Supplementary material is available at *Chemical Senses* online.

## Funding

This work was supported by National Institute on Aging at the National Institutes of Health [AG030481, AG043538, AG048511, AG000243, AG029795 to JP].

## Conflict of Interests

The authors declare no competing conflict of interest.

## Author Contributions

The idea for this project originated with SA. EC and JP supervised project organization, analysis strategy, and review of results. DW assisted with search strategies. Data collection and analysis were performed by SA, LW, DZ. Writing and editing was performed by all authors.

## Data Availability

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Ajmani GS, Suh HH, Wroblewski KE, Pinto JM. Smoking and olfactory dysfunction: a systematic literature review and meta-analysis. *Laryngoscope* 2017;127:1753–1761.
- Barba-Müller E, Craddock S, Carmona S, Hoekzema E. Brain plasticity in pregnancy and the postpartum period: links to maternal caregiving and mental health. *Arch Womens Ment Health*. 2019;22:289–299.
- Barth C, de Lange AMG. Towards an understanding of women's brain aging: the immunology of pregnancy and menopause. *Front Neuroendocrinol* 2020;58:100850.
- Beauchamp GK, Mennella JA. Flavor perception in human infants: development and functional significance. *Digestion* 2011;83(Suppl 1):1–6.
- Boesveldt S, de Graaf K. The differential role of smell and taste for eating behavior. *Perception* 2017;46:3073–3319.
- Boesveldt S, Parma V. The importance of the olfactory system in human well-being, through nutrition and social behavior. *Cell Tissue Res*. 2021;383:559–567.
- Borenstein M, Hedges LV, Higgins JP, Rothstein HR. 2021. *Introduction to meta-analysis*. 2nd edition. Hoboken, NJ: John Wiley & Sons.
- Cameron EL. Measures of human olfactory perception during pregnancy. *Chem Senses*. 2007;32:775–782.
- Cameron EL. Pregnancy and olfaction: a review. *Front Psychol* 2014a;5:67.
- Cameron EL. Pregnancy does not affect human olfactory detection thresholds. *Chem Senses*. 2014b;39:143–150.
- Dastur FN. 2000. *A controlled, longitudinal study of olfactory perception and symptoms of pregnancy sickness*. Unpublished Doctoral Dissertation, Dalhousie University. ProQuest Dissertations & Theses Global.
- Davies SJ, Lum JA, Skouteris H, Byrne LK, Hayden MJ. Cognitive impairment during pregnancy: a meta-analysis. *Med J Aust*. 2018;208:35–40.
- Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. *Physiol Behav*. 1984;32:489–502.
- Doty RL. 1995. *The Smell Identification Test 9TM Administration Manual*. Philadelphia, USA: Sensonics, Inc.
- Doty RL. 2000. *The smell threshold test administration manual*. Books & Manuals [Internet]. Sensonics, Incorporated.
- Fornazieri MA, Prina DMC, Favoreto JPM, Rodrigues e Silva K, Ueda DM, de Rezende Pinna F, Voegels RL, Cameron EL, Doty RL. Olfaction during pregnancy and postpartum period. *Chemosens Percept*. 2019;12:125–134.
- Gilbert AN, Wysocki CJ. Quantitative assessment of olfactory experience during pregnancy. *Psychosom Med*. 1991;53:693–700.
- Hedner M, Larsson M, Arnold N, Zucco GM, Hummel T. Cognitive factors in odor detection, odor discrimination, and odor identification tasks. *J Clin Exp Neuropsychol*. 2010;32:1062–1067.
- Henry JD, Rendell PG. A review of the impact of pregnancy on memory function. *J Clin Exp Neuropsychol*. 2007;29:793–803.
- Hoekzema E, Barba-Müller E, Pozzobon C, Picado M, Lucco F, García-García D, Soliva JC, Tobeña A, Desco M, Crone EA, et al. Pregnancy leads to long-lasting changes in human brain structure. *Nat Neurosci*. 2017;20:287–296.
- Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5:13.
- Hummel T, Sekinger B, Wolf SR, Pauli E, Kobal G. “Sniffin” Sticks’: olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem Senses*. 1997;22:39–52.
- Kim P, Leckman JF, Mayes LC, Feldman R, Wang X, Swain JE. The plasticity of human maternal brain: longitudinal changes in brain anatomy during the early postpartum period. *Behav Neurosci*. 2010;124:695–700.
- Kohli P, Soler ZM, Nguyen SA, Muus JS, Schlosser RJ. The association between olfaction and depression: a systematic review. *Chem Senses*. 2016;41:479–486.
- Köblle N, Hummel T, von Mering R, Huch A, Huch R. Gustatory and olfactory function in the first trimester of pregnancy. *Eur J Obstet Gynecol Reprod Biol*. 2001;99:179–183.
- Kyung-yeon. 2014. *Olfactory function and smell perception during pregnancy and postpartum in Asian women*. Unpublished.
- Landis NB, Hummel T, Hugentobler M, Giger R, Lacroix JS. Ratings of overall olfactory function. *Chem Senses*. 2003;28:691–694.
- Laska M, Koch B, Heid B, Hudson R. Failure to demonstrate systematic changes in olfactory perception in the course of pregnancy: a longitudinal study. *Chem Senses*. 1996;21:567–571.
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gotzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher D. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med*. 2009;6:e1000100.
- Lötsch J, Hummel T. Clinical usefulness of self-rated olfactory performance—a data science-based assessment of 6000 patients. *Chem Senses*. 2019;44:357–364.
- Nordin S, Broman DA, Bringlöv E, Wulff M. A longitudinal descriptive study of self-reported abnormal smell and taste perception in pregnant women. *Chem Senses*. 2004;29:391–402.
- Nordin S, Broman DA, Bringlöv E, Wulff M. Intolerance to ambient odors at an early stage of pregnancy: health and disability. *Scand J Psychol*. 2007;48:339–343.
- Nwankwo U, Fasunla AJ, Oladokun A, Nwaorgu OG. Comparison between olfactory function of pregnant women and non-pregnant women in reproductive age group in Ibadan, Nigeria. *Niger J Clin Pract*. 2017;20:610–615.
- Ochsenbein-Köblle N, von Mering R, Zimmermann R, Hummel T. Changes in olfactory function in pregnancy and postpartum. *Int J Gynecol Obstet*. 2007;97:10–14.



- Olofsson JK, Broman DA, Wulff M, Martinkauppi M, Nordin S. Olfactory and chemo somato sensory function in pregnant women assessed with event-related potentials. *Physiol Behav.* 2005;86:252–257.
- Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. *Syst Rev.* 2016;5:210.
- Page MJ, Higgins JPT, Sterne JAC. 2022. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Thomas J, Higgins JPT, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. *Cochrane handbook for systematic reviews of interventions*. *cochrane*. Cochrane, 2022. Available from <http://www.training.cochrane.org/handbook>. Accessed 22 April 2022.
- Pinto JM, Schumm LP, Wroblewski KE, Kern DW, McClintock MK. Racial disparities in olfactory loss among older adults in the United States. *J Gerontol A Biol Sci Med Sci.* 2014;69:323–329.
- Profet M. The evolution of pregnancy sickness as protection to the embryo against Pleistocene teratogens. *Evol Theory.* 1988;8:177–190.
- R Core Team. 2022. R: a language and environment for statistical computing. In: R Foundation for Statistical Computing.
- Rehbein E, Kogler L, Kotikalapudi R, Sattler A, Krylova M, Kagan KO, Sundström-Poromaa I, Derntl B. Pregnancy and brain architecture: associations with hormones, cognition and affect. *J Neuroendocrinol.* 2022;34:e13066.
- Rohatgi A. 2021. Webplotdigitizer: Version 4.5'. <https://automeris.io/WebPlotDigitizer>. Accessed 10 February 2022.
- Savović SN, Ninčić DP, Lemajić SN, Pilića VI, Mandić A, Rajović J, Ivetić-Petrović VR. Olfactory perception in women with physiologically altered hormonal status (during pregnancy and postmenopause). *Med Pregl.* 2002;55:380–383.
- Shu CH, Yuan BC, Lin SH, Lin CZ. Cross-cultural application of the “Sniffin’Sticks” odor identification test. *Am J Rhinol.* 2007;21:570–573.
- Şimşek G, Muluk NB, Arıkan OK, Dag ZO, Şimşek Y, Dag E. Marked changes in olfactory perception during early pregnancy: a prospective case-control study. *Eur Arch Otorhinolaryngol.* 2015;272:627–630.
- Şimşek Y, Şimşek G, Bayar Muluk N, Arıkan OK. Olfactory dysfunction and oxidative stress in pregnant women with hyperemesis gravidarum. *Arch Gynecol Obstet.* 2021;304:657–661.
- Soudry Y, Lemogne C, Malinvaud D, Consoli SM, Bonfils P. Olfactory system and emotion: common substrates. *Eur Ann Otorhinolaryngol Head Neck Dis.* 2011;128:18–23.
- Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, Carpenter J, Rücker G, Harbord RM, Schmid CH, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *Br Med J.* 2011;343:d4002.
- Stevenson RJ. An initial evaluation of the functions of human olfaction. *Chem Senses.* 2010;35:3–20.
- Swallow BL, Lindow SW, Aye M, Masson EA, Alasalvar C, Quantick P, Hanna J. Smell perception during early pregnancy: no evidence of an adaptive mechanism. *BJOG Int J Obstet Gynaecol.* 2005;112:57–62.
- Tan PC, Kartik B, Thanendran P, Zakaria R, Win ST, Omar SZ. Taste, smell and food-related nausea and vomiting responses in hyperemesis gravidarum: a case-controlled study. *Sci Rep.* 2020;10:4445.
- Viechtbauer W. Conducting meta-analyses in R with the metafor Package. *J Stat Softw.* 2010;36:1–48.
- Wells GA, Shea B, O’Connell D, Peterson J, Welch V, Losos M, Tugwell P. 2000. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. In.: Oxford. Available from [https://www.ohri.ca/programs/clinical\\_epidemiology/oxford.asp](https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp). Accessed 20 June 2022.
- Yahiaoui-Doktor M, Luck T, Riedel-Heller SG, Loeffler M, Wirkner K, Engel C. Olfactory function is associated with cognitive performance: results from the population-based LIFE-adult-study. *Alzheimer’s Res Ther* 2019;11:43.
- Yasar M, Sagit M, Zeki Uludag S, Ozcan I. Does odor and taste identification change during hyperemesis gravidarum? *Med Glas (Zenica).* 2016;13:50–55.