Transnational University Limburg is a unique collaboration of two universities in two countries: the University of Hasselt and Maastricht University.

UHASSELT
KNOWLEDGE IN ACTION

www.uhasselt.be
Universiteit Hasselt
Campus Hasselt:
Martelarenlaan 42 | 3500 Hasselt
Campus Diepenbeek:
Agoralaan Gebouw D | 3590 Diepenbeek

UHASSELT
KNOWLEDGE IN ACTION

Maastricht University

Faculty of Sciences

School for Information Technology

Master of Statistics and Data Science

Master's thesis

Prenatal exposure to maternal distress: effects on health and autonomous nervous system function in adulthood.

Abdelrahman Adel Ahmed Elreweny

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics and Data Science, specialization Data Science

SUPERVISOR :
dr. Cecile KREMER

SUPERVISOR :
Prof. Dr. Bea VAN DEN BERGH

2025
2026

Faculty of Sciences
School for Information Technology
Master of Statistics and Data Science

Master's thesis

Prenatal exposure to maternal distress: effects on health and autonomous nervous system function in adulthood.

Abdelrahman Adel Ahmed Elreweny

Thesis presented in fulfillment of the requirements for the degree of Master of Statistics and Data Science, specialization Data Science

SUPERVISOR :

dr. Cecile KREMER

SUPERVISOR :

Prof. Dr. Bea VAN DEN BERGH

Acknowledgements

"And Allah brought you out of the wombs of your mothers while you knew nothing, and gave you hearing, sight, and intellect, so that you might be thankful." (Qur'an, 16:78)

All praise, gratitude, and thanks are due to Allah, now and forever. Every step of this journey, every ease after difficulty, and every ability to persist were by His will and mercy. As the Prophet Muhammad PBUH said: *"Whoever is not grateful to the people is not grateful to Allah."* In this spirit, I would like to express my sincere gratitude to those who supported me throughout this journey.

I am deeply grateful to my internal Dr. Cécile KREMER and my external supervisor Prof. Dr. Bea Van den Bergh for their guidance, patience, and understanding, and for creating a supportive and respectful academic environment. I also thank Hasselt University, along with all professors, advisors, and staff, for the opportunity to pursue this master's program and for an enriching academic experience.

My heartfelt gratitude goes to my brother, Dr. Ahmed. Nearly two and a half years ago, you believed in me when I needed it most, and you stood by me in ways that only Allah knows fully. I will never forget what you did for me, and I thank Allah for blessing me with you as a brother.

To my mother and my sisters: although we have been unable to gather for years, and despite the many hardships and stresses you have endured, your unwavering support, love, and prayers have always accompanied me across the distance. Your strength has been a constant source of motivation, and I carry you with me in every achievement.

To my friends in Genk, Hasselt, Leuven, and beyond: thank you for your advice, encouragement, and presence when I needed support. I am truly grateful to have you in my life.

To my beloved wife, I am deeply thankful for your patience, care, and strength during a challenging period of this journey. As we now await the arrival of our daughter, I am continually reminded of the responsibility, patience, and perseverance that this new chapter entails.

Finally, while conducting this study on maternal stress during pregnancy and its long-term effects on offspring, I could not help but reflect on the immense suffering endured by people in general, and particularly by pregnant women, in regions affected by war and deprivation, such as Gaza and Sudan. The anxiety of hunger, cold, displacement, loss of loved ones, and constant fear faced by mothers and their unborn children in such circumstances goes beyond measure. I pray that Allah grants them relief, protection, and justice.

All thanks and praise belong to Allah alone.

Abstract

Prenatal maternal anxiety has been proposed as a potential risk factor for long-term alterations in offspring stress regulation through developmental programming of the autonomic nervous system (ANS). However, evidence regarding its enduring effects on adult stress physiology remains mixed. The present study examined the association between prenatal maternal anxiety and autonomic and endocrine stress markers in adulthood, assessed during a laboratory stress paradigm, including heart rate variability (HRV), skin conductance level (SCL), salivary cortisol, salivary alpha-amylase, and their ratio.

Data were drawn from a longitudinal cohort study in which offspring were assessed at ages 27–29 years during a standardized laboratory stress protocol. Longitudinal linear mixed-effects models were applied to evaluate baseline levels, stress-related changes, and recovery dynamics, complemented by summary metrics capturing change from baseline, recovery slopes, and peak responses. Missing data were handled via likelihood-based methods and multiple imputation under missing-at-random assumptions, with sensitivity analyses conducted to assess robustness under missing-not-at-random scenarios.

Across most outcomes, including HRV indices and salivary cortisol, no evidence was found for systematic differences in stress reactivity or recovery associated with prenatal maternal anxiety. Selected HRV measures indicated differences in tonic autonomic activity, and salivary alpha-amylase showed gender-specific associations with prenatal maternal anxiety at baseline and peak levels. These effects were not reflected in altered temporal stress trajectories and were not accompanied by consistent changes across physiological systems. Sensitivity analyses supported the robustness of the main conclusions.

Overall, the findings suggest that prenatal maternal anxiety is not associated with widespread or persistent dysregulation of adult stress responses, but may be linked to subtle, context-dependent differences in baseline autonomic functioning.

Keywords: *prenatal maternal anxiety, autonomic nervous system, heart rate variability, salivary biomarkers, stress*

Contents

1 Introduction **8**
1.1 Background . . . 8
1.2 Research Questions . . . 9
1.3 Hypotheses . . . 10

2 Description of the Datasets **10**
2.1 Original Cohort and Later Follow-up . . . 10
2.2 HRV and Skin Conductance Dataset (RQs 1 and 3) . . . 11
2.2.1 Study sample and design . . . 11
2.2.2 Physiological variables in the HRV/SCL dataset . . . 11
2.2.3 Missingness in the HRV/SCL dataset . . . 12
2.3 Saliva Biomarker Dataset: Cortisol and Alpha-Amylase (RQ2) . . . 13
2.3.1 Study sample and design . . . 13
2.3.2 Physiological variables in the cortisol and alpha-amylase dataset . . . 13
2.3.3 Missingness in the cortisol and alpha-amylase dataset . . . 14

3 Methods **14**
3.1 Procedure . . . 14
3.2 Methodology for HRV and SCL Data . . . 15
3.2.1 Participants and design . . . 15
3.2.2 Data preprocessing and transformations . . . 16
3.2.3 Missing data . . . 19
3.2.4 Exploratory data analysis . . . 21
3.2.5 Model-building strategy for HRV and SCL outcomes (LMM) . . . 21
3.2.6 MNAR sensitivity analysis (delta adjustment) . . . 23
3.3 Methodology for Cortisol and sAA Data . . . 25
3.3.1 Data preprocessing and transformations . . . 25
3.3.2 Missing data . . . 25
3.3.3 Exploratory analysis of salivary cortisol and alpha-amylase . . . 27
3.3.4 Model-building strategy for salivary cortisol and alpha-amylase . . . 27
3.3.5 Sensitivity analysis under MNAR assumptions . . . 30

4 Results **30**
4.1 HRV and SCL Results . . . 30
4.1.1 MAR results . . . 30
4.1.2 MNAR sensitivity analysis results . . . 32
4.2 Cortisol and sAA Results . . . 32
4.2.1 Time-since-CPT mixed models . . . 32
4.2.2 Summary response metrics . . . 36
4.2.3 MNAR sensitivity analysis . . . 36

5 Discussion **36**

6 Possible drawbacks **37**

7 Ethical thinking, societal relevance and stakeholder awareness **38**

8 Conclusion **39**

9 Future research **39**

List of Tables

1 Structural variables in the HRV/SCL dataset. 12
2 Physiological variables in the HRV/SCL dataset 12
3 Physiological and timing variables in the saliva biomarkers dataset 14
4 Observed ranges and recommended transformations for HRV/SCL variables . 17
5 Descriptive statistics of transformed HRV/SCL variables 19
6 Missingness in HRV/SCL dataset 20
7 Final models for HRV/SCL analysis 23
8 Original-scale salivary variables and applied transformations 25
9 Summary statistics of log-transformed salivary variables 25
10 Missingness in salivary data 27
11 Final LMM models for salivary data 29
12 Participants distribution in HRV/SCL dataset 31
13 HRV/SCL results: interaction model 32
14 HRV/SCL results: interaction dropped final models 33
15 Salivary Cortisol results: LMM-CPT with interaction 34
16 Salivary Cortisol results: LMM-CPT main effects model 35
17 Salivary Alpha-Amylase results: LMM-CPT final model 35
18 Salivary Alpha-Amylase-to-Cortisol ratio results: LMM-CPT final model . . 35

List of Figures

1 Structure and timeline of the laboratory visit 16
2 Histogram of raw-RMSSD 18
3 Q–Q plot of raw-RMSSD 18
4 Histogram of log-transformed RMSSD 18
5 Q–Q plot of log-transformed RMSSD 18
6 Histogram of raw Salivary Cortisol 26
7 Q–Q plot of raw Salivary Cortisol 26
8 Histogram of log-transformed Salivary Cortisol 26
9 Q–Q plot of log-transformed Salivary Cortisol 26
10 Time since CPT distribution 29
11 * 41
12 * 41
13 * 41
14 * 41
15 Distribution of SDNN before and after logarithmic transformation. 41
16 * 42
17 * 42
18 * 42
19 * 42
20 Distribution of RSA0 before and after logarithmic transformation. 42
21 * 42

22 * 42
23 * 43
24 * 43
25 * 43
26 * 43
27 Distribution of HF before and after logarithmic transformation. 43
28 * 44
29 * 44
30 * 44
31 * 44
32 Distribution of LF before and after logarithmic transformation. 44
33 * 45
34 * 45
35 * 45
36 * 45
37 Distribution of LFHF before and after logarithmic transformation. 45
38 * 46
39 * 46
40 * 46
41 * 46
42 Distribution of Average_SCL_uS before and after logarithmic transformation. 46
43 * 47
44 * 47
45 * 47
46 * 47
47 Distribution of Salivary Alpha Amylase before and after logarithmic transformation. 47
48 * 48
49 * 48
50 * 48
51 * 48
52 Distribution of Salivary Alpha Amylase to Cortisol ratio before and after logarithmic transformation. 48
53 lnRMSSD and lnCortisol residual diagnostics plots 49
54 lnRMSSD model diagnostics 49
55 lnCortisol model diagnostics 49
56 Participant-level Cook's distance for the lnCortisol model 50
57 Participant-level leverage for lnCortisol model 50

List of Abbreviations

SDNN	Standard Deviation of Normal-to-Normal Intervals
RMSSD	Root Mean Square of Successive Differences
RSA	Respiratory Sinus Arrhythmia
HF	High-frequency HRV power
LF HRV	Low-frequency HRV power
LF/HF	The Ratio of Low-frequency to High-frequency HRV power
PEP	Pre-ejection Period
SCL	Skin Conductance Level
Cortisol	Salivary cortisol
sAA	Salivary alpha-amylase
AACR	Alpha-amylase to cortisol ratio
REML	Restricted Maximum Likelihood
ML	Maximum Likelihood
DL	Direct Likelihood
ln	Natural Logarithm
MI	Multiple Imputation
MAR	Missing at Random
MNAR	Missing Not at Random

1 Introduction

1.1 Background

The universal burden of mental illness has been continually increasing over recent decades, exerting a substantial impact on both individual well-being and public health systems. Recent epidemiological studies have reported rising prevalence rates of mental and psychological disorders worldwide (Vigo et al., 2016; Turk et al., 2023). In this context, there is growing recognition that identifying risk factors early in life and implementing preventive strategies may be more effective and cost-efficient than interventions applied later in adulthood (Bauer et al., 2016). Accordingly, prenatal exposure to maternal stress and anxiety during pregnancy has received increasing attention as a potential determinant of long-term offspring health outcomes.

The Developmental Origins of Health and Disease (DOHaD) framework proposes that adverse conditions during sensitive periods of prenatal development can influence neuroendocrine and physiological systems in ways that persist into adulthood (Barker, 2004; Van den Bergh et al., 2025). Within this paradigm, prenatal maternal distress is considered a particularly important exposure, given the heightened plasticity of fetal biological systems during gestation.

However, researchers have not yet reached consensus on the precise biological mechanisms underlying these long-term associations. Several studies suggest that multiple interacting biological systems may mediate the effects of prenatal maternal distress. This complexity is reflected in the inconsistent associations reported between maternal distress and cortisol levels during pregnancy, indicating that no single pathway fully accounts for observed offspring outcomes (van den Heuvel et al., 2018).

A growing body of longitudinal and experimental research has examined the consequences of prenatal stress for offspring health across the lifespan. These studies have demonstrated associations between prenatal stress exposure and later disruptions in cognitive functioning, emotional regulation, autonomic nervous system (ANS) activity, and stress-related symptom perception in offspring (Van den Bergh and Marcoen, 2004; von Leupoldt et al., 2017; Turk et al., 2023). In parallel, numerous studies have linked gestational maternal anxiety to alterations in both the structure and function of the offspring's brain. Longitudinal prospective evidence indicates that increased maternal stress during pregnancy is associated with reduced gray matter volume in childhood (Buss et al., 2010), alterations in emotional and cognitive processing during adolescence (Van den Bergh et al., 2008), differences in white matter microstructure in adulthood (Van den Bergh et al., 2025), and modified resting-state functional connectivity patterns (Turk et al., 2023).

More recently, advances in neuroimaging techniques such as diffusion kurtosis imaging (DKI) and neurite orientation dispersion and density imaging (NODDI) have enabled more refined characterization of white matter architecture. Using these approaches, studies have identified associations between elevated prenatal maternal stress and widespread, albeit subtle, alterations in white matter microstructure in adulthood (Van den Bergh et al., 2025). These findings are consistent with theoretical models proposing that prenatal anxiety may disrupt axonal development and myelination, particularly during periods of accelerated neural growth such as mid-pregnancy.

Beyond structural brain alterations, prenatal maternal stress has also been associated with long-term changes in physiological stress regulation systems. In particular, evidence suggests involvement of autonomic and endocrine pathways, including altered functioning of the hypothalamic–pituitary–adrenal (HPA) axis and changes in heart rate variability (HRV), a widely used indicator of parasympathetic nervous system activity (Van den Bergh et al., 2008; Braeken et al., 2013).

Despite substantial progress in this field, several important gaps remain. Although prenatal anxiety has been extensively studied, long-term effects on adult brain architecture, autonomic nervous system functioning, and mental health outcomes remain comparatively underexplored relative to research focusing on childhood and adolescence. While early-life studies have documented associations between prenatal anxiety and neurodevelopmental and behavioral outcomes in younger populations (Van den Bergh and Marcoen, 2004), extending follow-up into adulthood is essential for evaluating the persistence and continuity of these effects.

Second, many previous studies have adopted hypothesis-driven approaches targeting specific neural regions, such as the amygdala or hippocampus. Although informative, such region-specific analyses may overlook large-scale or network-level alterations. Data-driven whole-brain approaches have therefore been advocated to more comprehensively assess neurodevelopmental outcomes following prenatal stress exposure (Turk et al., 2023).

Third, while dysregulation of the hypothalamic–pituitary–adrenal (HPA) axis—most commonly assessed via cortisol levels—has been proposed as a key mechanism linking prenatal maternal anxiety to offspring outcomes, findings remain inconsistent. Several studies have reported weak or absent associations between prenatal maternal anxiety and maternal or infant salivary cortisol levels (van den Heuvel et al., 2018). These inconsistencies suggest that complementary biological pathways, including autonomic nervous system functioning and large-scale neural network connectivity, may contribute to long-term outcomes.

This study utilizes a longitudinal cohort in which 86 pregnant women were recruited between 1986 and 1987 and assessed for emotional well-being using the State–Trait Anxiety Inventory (STAI). Prenatal maternal anxiety was measured during gestational weeks 12–22, and offspring were followed into adulthood. The primary objective of this study is to investigate associations between prenatal maternal anxiety and autonomic, psychological, and endocrine outcomes in offspring aged 27–29 years.

Specifically, this study compares resting and experimentally induced stress conditions to examine the main and interactive effects of prenatal maternal anxiety and stress condition on: (i) autonomic nervous system activity, indexed by heart rate variability (HRV) parameters including SDNN, RMSSD, RSA, HF, LF, LF/HF, and pre-ejection period (PEP); (ii) stress-related physiological responses, indexed by salivary cortisol (HPA-axis activity), salivary alpha-amylase (sAA; sympathetic activity), and the sAA/cortisol ratio; and (iii) skin conductance levels, reflecting sympathetic arousal.

1.2 Research Questions

1. **HRV**: *a)* Is there a significant effect of induced stress on HRV parameters compared to rest? *b)* Does prenatal maternal anxiety affect these parameters? *c)* Is there an interaction effect between condition and prenatal anxiety?

2. **Salivary biomarkers**: *a)* Do cortisol, alpha-amylase, and the sAA/Cor ratio differ between baseline and stress? *b)* Is there an effect of prenatal anxiety on these biomarkers? *c)* Is there an interaction effect?

3. **Skin conductance**: *a)* Does skin conductance differ between rest and stress? *b)* Does prenatal anxiety influence SCL values? *c)* Is there an interaction effect?

1.3 Hypotheses

We hypothesize that prenatal maternal anxiety has long-term consequences for stress physiology. Specifically, we expect that:

- Offspring of mothers with high prenatal anxiety will show reduced parasympathetic activity, indicated by lower SDNN, RMSSD, RSA, and HF HRV.
- These individuals will also show altered stress-system activity, reflected by lower pre-ejection period (PEP) values, higher skin conductance levels, increased salivary alpha-amylase (sAA; sympathetic activity), and elevated salivary cortisol concentrations (HPA-axis activity).
- The physiological differences will be more pronounced under stress, suggesting increased reactivity and/or reduced recovery capacity in the high-anxiety group.

2 Description of the Datasets

2.1 Original Cohort and Later Follow-up

Between 1986 and 1987, eighty-six pregnant women were recruited at the University Hospital Gasthuisberg in Leuven, Belgium, to participate in a longitudinal study examining the long-term association between maternal psychological well-being during pregnancy and offspring health outcomes. During pregnancy, participants completed mental health questionnaires at several gestational time points (12–22, 23–31, and 32–40 weeks of pregnancy).

In the present follow-up study, the focus was on their offspring in adulthood, when they were between 27 and 29 years old. Of the 86 originally recruited pregnancies, 52 offspring (27 females, 25 males) agreed to participate. Offspring were classified into two groups based on the prenatal state anxiety scores of their mothers measured during gestational weeks 12–22, using the State Anxiety subscale of the State–Trait Anxiety Inventory (STAI):

- Low-to-medium prenatal maternal anxiety (Low PMA)
- High prenatal maternal anxiety (High PMA)

This grouping was predefined in the original dataset and based on a percentile cut-off of maternal state anxiety scores. Participants with scores below the 75th percentile (< 43) were classified as Low PMA, whereas those with scores at or above the 75th percentile were classified as High PMA (≥ 43).

Two primary datasets were used in the present thesis:

- A heart rate variability and skin conductance level dataset (HRV/SCL)
- A salivary cortisol and alpha-amylase dataset (AAS)

Both datasets were derived from psychophysiological laboratory assessments conducted in the offspring participants around the age of 28 years (see subsequent sections for details).

2.2 HRV and Skin Conductance Dataset (RQs 1 and 3)

2.2.1 Study sample and design

Out of the 52 offspring participants, the HRV/SCL dataset used in the present analyses included 48 individuals (27 females and 21 males), of whom 35 belonged to the low prenatal maternal anxiety (PMA) group and 13 to the high PMA group. The sources of missingness are described in subsequent sections.

In this dataset, each participant contributed repeated measurements as a result of multiple tasks performed during the laboratory visit. Specifically, six predefined segments were used to extract outcome variables for HRV analyses. In contrast, skin conductance level (SCL) data were available for only two segments per participant.

The repeated-measures structure was represented by the variable `Subsequent_observations` (hereafter referred to as `Time`), which indexed the experimental segments as follows:

- **T1 and T6**: Two non-stress, rest-like periods corresponding to baseline and post-stress recovery.
- **T2–T5**: Four stress periods corresponding to different stress-inducing tasks.

For SCL, the dataset contained only two observations per participant, neither of which corresponded to a pure rest condition. Consequently, both SCL observations were considered stress-related, and it was not possible to assess stress-versus-rest effects for this outcome. Analyses of SCL were therefore restricted to comparisons between the low and high PMA groups during stress.

2.2.2 Physiological variables in the HRV/SCL dataset

The dataset contained both structural and physiological outcome variables. Table 1 summarizes the structural variables in the dataset, while Table 2 presents the physiological variables included in the HRV/SCL dataset. As shown in Table 1, the variable `REST_VS_STRESS` distinguishes between rest-like (non-stress) and stress conditions. The rest-like condition combines the baseline period (T1) and the post-stress recovery period (T6). This aggregation was used to improve the stability of rest estimates by increasing the number of observations (48 participants $\times$ 2 rest periods = 96 observations), compared to 192 observations during stress periods.

The repeated-measures structure of the data arises from multiple measurement segments per participant, allowing for both within-subject and between-subject comparisons of physiological outcomes.

As indicated in Table 2, all physiological variables (except PEP) were transformed prior to analysis, since they exhibited skewed distributions. Details regarding variable distributions and transformation choices are provided in the Methods section.

Table 1: Structural variables in the HRV/SCL dataset.

Variable	Type	Description
Subject_ID	Categorical (subject)	Unique identifier for each participant
Time	Categorical (6 levels)	Measurement segment (T1–T6) within the laboratory visit
REST_VS_STRESS	Binary (0/1)	0 = non-stress rest-like condition (baseline and recovery), 1 = stress condition
Group	Binary (0/1)	Prenatal maternal anxiety group (low vs high)
Gender	Binary (m/f)	Sex of the offspring participant (male vs female)

Table 2: Physiological variables in the HRV/SCL dataset. Physiological interpretations are based on standard HRV and psychophysiology guidelines (Shaffer and Ginsberg, 2017; Laborde et al., 2017; Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, 1996).

Variable	Unit	Mainly reflects	Form used in analyses
SDNN	ms	Global HRV (mixed sympathetic/parasympathetic)	Natural logarithm
RMSSD	ms	Short-term vagal (parasympathetic) HRV	Natural logarithm
RSA_0	ms	Vagal (parasympathetic) cardiac control	$\log(RSA + 1)$
LF power	ms^2	Mixed baroreflex-related HRV	Natural logarithm
HF power	ms^2	Vagal (parasympathetic) HRV	Natural logarithm
LF/HF ratio	–	Crude sympathovagal index[a]	Natural logarithm
PEP	ms	Cardiac sympathetic drive	Raw values
SCL	μS	Sympathetic electrodermal activity	$\log(SCL + 0.01)$

[a] Physiological interpretation of LF and LF/HF as pure markers of sympathetic activity or sympathovagal balance is debated; results are interpreted with caution.

2.2.3 Missingness in the HRV/SCL dataset

Several sources of missingness were present in the HRV/SCL dataset as received for the present thesis. First, some participants did not participate in specific parts of the laboratory protocol. Second, technical issues occurred during physiological data acquisition and processing using the VU University Ambulatory Monitoring System (VU-AMS), which was employed to record HRV and SCL signals during the laboratory visit.

During the original data extraction and quality control procedures, summary statistics (e.g., mean RMSSD, SDNN, RSA, HF, LF, PEP, and SCL) were automatically derived by the VU-AMS system for predefined task segments. As part of this preprocessing pipeline, all measurements were checked to ensure values fell within physiologically plausible ranges. Observations exceeding three standard deviations from the mean of the corresponding variable were flagged as outliers and coded as missing. In addition, for several HRV indices, reliable estimation was not possible for signal fragments shorter than 30 seconds. Consequently,

measurements based on fragments shorter than this minimum duration were also coded as missing.

This criterion particularly affected the T5 segment, corresponding to the cold pressor task, which accounted for a substantial proportion of missing values in the dataset. All of these preprocessing steps were implemented prior to the dataset being provided for the present analyses.

As a result of these factors, the HRV/SCL dataset included data for 48 offspring participants rather than the original 52, and not all participants contributed valid measurements for all six task segments. The resulting dataset is therefore unbalanced, with varying numbers of observations per participant.

2.3 Saliva Biomarker Dataset: Cortisol and Alpha-Amylase (RQ2)

2.3.1 Study sample and design

Out of the 52 participants included in the original cohort, 51 participants were retained for the salivary cortisol and alpha-amylase analyses. One participant was excluded because no salivary measurements were available in the provided datasets. The final sample consisted of 27 females and 24 males, with 38 participants classified in the low prenatal maternal anxiety (PMA) group and 13 in the high PMA group.

During the laboratory visit, each participant contributed repeated salivary measurements as a result of multiple predefined sampling moments surrounding the experimental tasks. In total, up to eight saliva samples were collected per participant, corresponding to eight predefined measurement occasions. This repeated-measures structure is represented by the variable `Sample`, taking integer values from 1 to 8, and applies to salivary cortisol, salivary alpha-amylase, and their ratio.

2.3.2 Physiological variables in the cortisol and alpha-amylase dataset

The salivary dataset comprised repeated measurements of salivary cortisol, salivary alpha-amylase and the ratio of salivary alpha-amylase to cortisol (AACR) collected at up to eight predefined sampling occasions per participant during the laboratory visit. In addition to the physiological outcomes, timing-related variables were incorporated into the salivary dataset to support longitudinal modeling of the responses. These variables were derived from recorded sample timestamps and included the exact date and time of saliva collection, as well as a variable representing the time elapsed since the cold pressor task (CPT), expressed in minutes. These timing variables were used to index and model temporal dynamics in cortisol and alpha-amylase responses.

The original laboratory data were provided in wide format, with separate variables corresponding to each of the eight saliva samples. For analysis, the data were reshaped and augmented to construct an analysis-ready dataset retaining the repeated-measures structure indexed by the variable `Sample` (values 1–8). All physiological outcomes were analyzed on the natural logarithmic scale to address right-skewness and improve distributional properties.

Table 3 provides an overview of the physiological variables included in the cortisol and alpha-amylase analyses.

Table 3: Physiological and timing variables in the cortisol and alpha-amylase dataset

Variable	Description	Transformation
Cortisol	Salivary cortisol concentration	Natural log
Alpha-amylase	Salivary alpha-amylase activity	Natural log
AACR	Alpha-amylase to cortisol ratio (laboratory-provided)	Natural log
Sample	Predefined saliva sampling occasion (1–8)	–
Datetime	Recorded date and time of saliva collection	–
Minutes from CPT	Time elapsed since CPT (minutes)	–

2.3.3 Missingness in the cortisol and alpha-amylase dataset

One participant from the original cohort did not have any available salivary measurements and was therefore excluded from the cortisol and alpha-amylase analyses, resulting in a final sample size of 51 participants. Within this sample, missingness was present in both the salivary outcomes and the associated timing information, reflecting incomplete saliva sampling and incomplete recording of sampling times.

The handling of missing data and the implications for statistical modeling are described in detail in the Methodology section.

3 Methods

3.1 Procedure

The laboratory visit consisted of three parts. Across these parts, salivary cortisol and alpha-amylase samples were collected eight times. Figure 1 illustrates the structure of the laboratory visit and the timing of saliva sample collection.

During the first part, two salivary cortisol samples were collected. Participants underwent an fMRI scan, with the first sample taken immediately before the scan and the second sample collected 50–75 minutes after the scan.

Participants then proceeded to the second part, during which heart rate variability (HRV) recordings were initiated using the VU University Ambulatory Monitoring System (VU-AMS) (de Geus et al., 1995; Willemsen et al., 1996). This part comprised three sessions:

- **Session 1**: Mini-International Neuropsychiatric Interview (MINI) (Sheehan et al., 1998).
- **Session 2**:
  - Test of Visual Attention (TVA) (Bundesen, 1990).
  - Block Design and Vocabulary subtests of the Wechsler Adult Intelligence Scale (Wechsler, 2008).
  - Trail Making Test (Part A) (Reitan and Wolfson, 1985).
- **Session 3**: The first stress task, consisting of mental arithmetic calculations performed under a three-minute time pressure.

During this part, an additional module was attached to the VU-AMS device to record galvanic skin response (skin conductance level).

The third and final part of the laboratory visit comprised the remaining stress tasks followed by a recovery period. This part included: (1) a dyspnea task, (2) a socially evaluated cold pressor task, and (3) the completion of several questionnaires assessing psychological distress as well as family- and job-related variables.

Before the dyspnea task, participants completed questionnaires assessing symptom perception, including the Dyspnea Catastrophising Scale (Solomon et al., 2015; De Peuter et al., 2006), the Pain Catastrophising Scale (Sullivan et al., 1995), and the Patient Health Questionnaire for somatic symptoms (PHQ-15) (Kroenke et al., 2002). Participants then completed two stress tasks to evaluate dyspnea perception: a magnitude estimation task and a sustained breathing task (von Leupoldt et al., 2017).

The final stress-inducing task was the socially evaluated cold pressor task (CPT) (Schwabe and Schächinger, 2018; Schwabe et al., 2008). During this task, participants were instructed to immerse their hand in cold water for as long as they could tolerate, up to a maximum duration of three minutes. To enhance social-evaluative stress, participants were required to maintain eye contact with a camera recording them throughout the task. If a participant withdrew their hand earlier, the task was terminated.

The remaining six salivary cortisol and alpha-amylase samples were collected during this third part as follows:

- **Sample 3**: before the magnitude estimation task (i.e., at the start of the third part).
- **Samples 4–8**: after the CPT, at intervals of 10–15 minutes.

Following completion of the CPT, participants entered a recovery period during which they completed additional questionnaires.

3.2 Methodology for HRV and SCL Data

3.2.1 Participants and design

As described previously in Section 2.2.1, HRV and SCL data were collected from 48 offspring participants during a laboratory visit conducted when participants were approximately 28 years old. For HRV, six measurement segments were recorded per participant. This repeated-measures structure captures within-subject variability across rest-like and stress conditions.

The six HRV segments corresponded to different tasks performed during the laboratory visit:

- **T1**: Mini-International Neuropsychiatric Interview (rest period prior to stress induction)
- **T2**: Mental arithmetic calculations (stress task 1)
- **T3**: Magnitude estimation task (stress task 2)
- **T4**: Sustained breathing task (stress task 3)

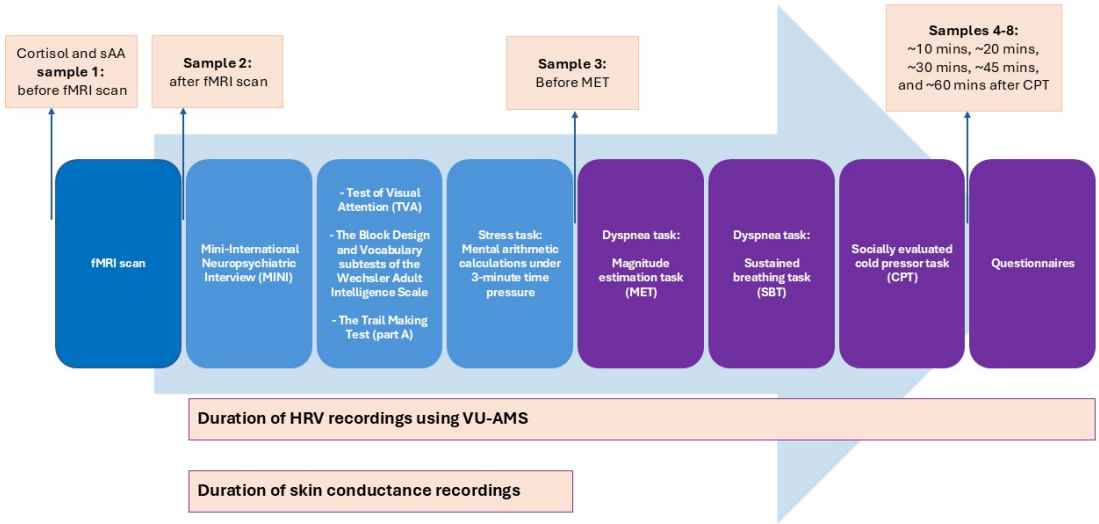

Figure 1: Structure of the laboratory visit and timing of salivary cortisol and alpha-amylase sampling.

- **T5**: Socially evaluated cold pressor task (stress task 4)
- **T6**: Recovery period (rest period following stress induction)

To model stress-related effects, the binary variable `REST_VS_STRESS` was used, distinguishing between rest-like segments (T1 and T6) and stress segments (T2–T5).

For SCL, only two measurement segments were available per participant. As neither segment corresponded to a pure rest condition, both SCL observations were classified as stress-related. Consequently, analyses of SCL focused on differences between prenatal maternal anxiety (PMA) groups during stress rather than on stress-versus-rest comparisons.

Participants were classified into low and high PMA groups based on maternal state anxiety scores measured during mid-gestation.

3.2.2 Data preprocessing and transformations

Section 3.2.1 describes the selected rest and stress segments used in the present study. Accordingly, all HRV and SCL summary statistics (e.g., mean SDNN, RMSSD, RSA, HF, LF, PEP, and mean SCL) were calculated per segment by the VU University Ambulatory Monitoring System (VU-AMS) during the original data extraction. As described earlier in Section 2.2.3, observations identified as outliers (exceeding three standard deviations from the variable-specific mean) and statistics derived from signal fragments shorter than 30 seconds—particularly during the cold pressor task—were coded as missing. All of these preprocessing steps were implemented prior to the data being provided for the present thesis; thus, the preprocessed dataset served as the starting point for all subsequent analyses. All further steps described below were performed by the author.

Table 2 in Section 2.2.2 summarizes the physiological outcome variables of interest, including time-domain and frequency-domain HRV indices, pre-ejection period (PEP), and skin conductance level (SCL), and indicates the transformations applied to each variable. Because the HRV/SCL dataset contained missing values, exploratory inspection of the observed data was conducted with caution and without drawing inferential conclusions. The purpose of this inspection was limited to assessing distributional properties in order to inform appropriate transformation choices.

Table 4: Observed ranges and recommended transformations for physiological variables

Variable	Range	Contains 0 / negatives	Recommended transform
PEP (ms)	46 – 162	No	None
HF (ms^2)	17 – 4443	No	$\log(y)$
LF (ms^2)	100 – 7377	No	$\log(y)$
LF/HF ratio	0.04 – 14.0	No	$\log(y)$
Average SCL (μS)	0.005 – 17.25	No (near-zero)	$\log(y + 0.01)$
RMSSD (ms)	6.49 – 172	No	$\log(y)$
SDNN (ms)	19.37 – 210	No	$\log(y)$
RSA_0 (ms)	0 – 161	Yes	$\log(y + 1)$

To this end, the observed ranges of the physiological variables on their raw scales were examined, as summarized in Table 4. Inspection of these ranges suggested substantial asymmetry for several variables, motivating further evaluation of their distributions using graphical methods. In line with established recommendations for HRV and psychophysiological data, logarithmic transformations were applied to most variables to reduce right skewness and stabilize variance. For SCL, which included near-zero values, a small constant was added prior to log transformation ($\log(y + 0.01)$). For RSA, a shifted logarithmic transformation ($\log(y+1)$) was used to avoid undefined values. PEP was not transformed, as its distribution was approximately symmetric and it is interpretable on the original scale (with shorter PEP indicating stronger sympathetic activation).

To illustrate the effect of transformation, RMSSD is presented as a representative example. Figures 2 and 3 show the distribution of RMSSD on the raw scale, while Figures 4 and 5 show the distribution after logarithmic transformation. Corresponding histograms and Q–Q plots for all remaining physiological variables are provided in the Appendix A.

After transformation, the distributions of the physiological variables were re-inspected using histograms and Q–Q plots. Additionally, descriptive statistics of the transformed variables, collapsed across time points and based on available observations only, are summarized in Table 5. These summaries are provided to give an overview of the scale and variability of the transformed measures and are not used for inferential purposes. Most transformed variables exhibited approximately symmetric distributions with minor tail deviations, consistent with known physiological floor effects for some measures.

All subsequent analyses were conducted using the transformed variables (with the exception of PEP). Final model estimates were back-transformed to the original scale to facilitate interpretation.

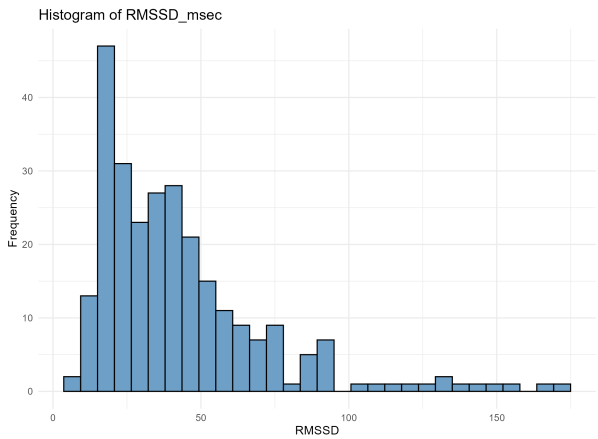

Figure 2: Histogram of RMSSD on the raw scale (pre-transformation).

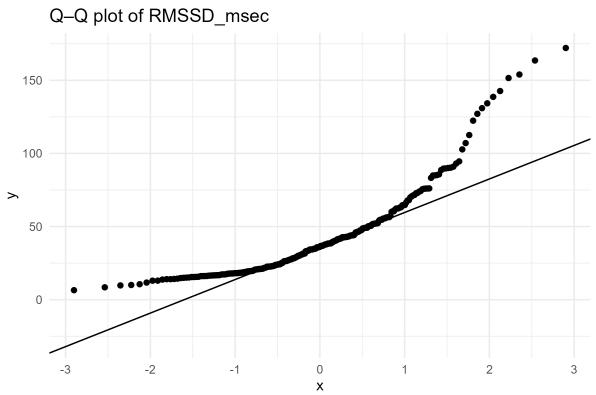

Figure 3: Q–Q plot of RMSSD on the raw scale (pre-transformation).

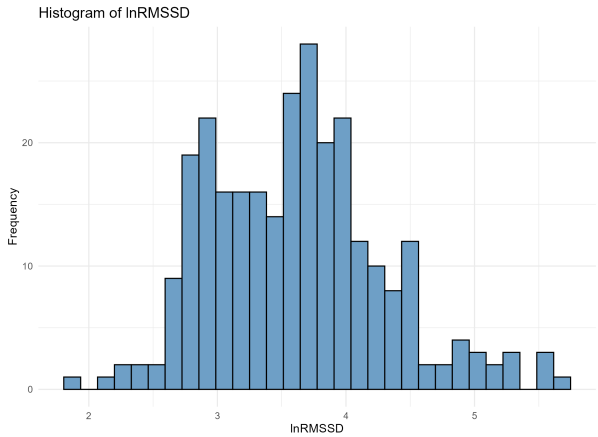

Figure 4: Histogram of lnRMSSD after logarithmic transformation.

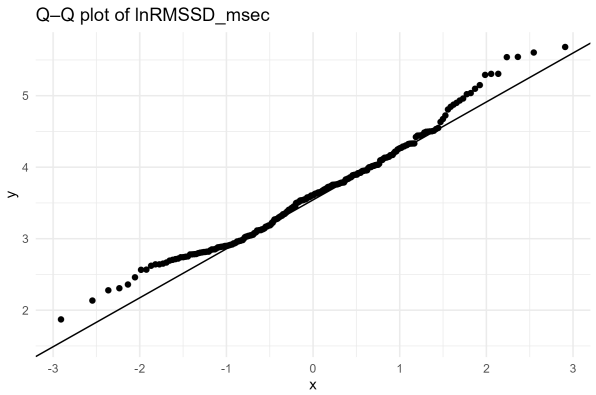

Figure 5: Q–Q plot of lnRMSSD after logarithmic transformation.

Table 5: Descriptive statistics of transformed HRV/SCL variables collapsed across time points (available observations only).

Variable	Min	Mean	SD	Max
$\ln RSA_0$ (ms)	0.00	4.01	0.62	5.69
lnAverage SCL (μS)	-4.20	0.28	2.43	2.85
lnSDNN (ms)	2.96	4.13	0.49	5.49
lnRMSSD (ms)	1.87	3.61	0.70	5.68
lnHF (ms^2)	2.89	6.14	1.20	9.84
lnLF (ms^2)	4.60	6.97	0.86	9.45
ln(LF/HF)	-0.95	1.20	0.71	2.82
PEP (ms)	46.0	95.8	24.8	162.0

3.2.3 Missing data

In longitudinal studies, missing data are a common and important issue. Missingness can affect both the structure and interpretation of longitudinal datasets. In the present study, if no data were missing, the dataset would be balanced, with each participant contributing the same number of measurements (six HRV segments and two SCL segments). However, the observed HRV/SCL data were unbalanced due to missing observations.

Specifically, missingness resulted in:

- Unequal numbers of measurement segments per participant;
- Substantial segment-specific missingness for some variables, particularly for HF, LF, and LF/HF, where very few observations were available at T2 (three participants out of 48) and no observations were available at T5.

Thus, despite a balanced study design, the realized dataset was unbalanced as a result of missing data.

Different patterns of missingness can arise in longitudinal data. These include: (1) complete data, where no observations are missing; (2) monotone missingness, where missing values follow a sequential pattern such that once an observation is missing, all subsequent observations are also missing (often due to dropout); and (3) non-monotone missingness, where observations are intermittently missing across time points, such that missing values may occur between observed values (Diggle et al., 2002; Verbeke and Molenberghs, 2000; Molenberghs and Verbeke, 2005). Non-monotone missingness represents a more complex pattern, as missingness does not follow a single directional sequence.

Table 6 summarizes the extent and structure of missingness across time points for the physiological variables. As shown in the table, observations may be present at later time points even when earlier measurements are missing, and vice versa. This indicates that the missing data pattern in the HRV/SCL dataset is predominantly non-monotone.

Missing data typically arise for reasons beyond the control of investigators, and in some cases may be related to the outcome of interest. Consequently, it is important to consider the underlying mechanism that governs missingness. While certain analytical approaches allow missing data to be ignored under specific assumptions, explicit consideration of the

missingness mechanism is generally required, as the true process generating missing data is rarely known with certainty.

For the primary analyses, missing outcome data were assumed to be missing at random (MAR). Under the MAR assumption, the probability that an observation is missing may depend on observed data, but does not depend on the unobserved value itself, conditional on the observed information (Diggle et al., 2002).

Table 6: Missingness by time point and missing-data patterns for physiological variables

Variable	Struct. T	Miss T1	Miss T2	Miss T3	Miss T4	Miss T5	Miss T6	Non-monotone	Representative patterns (count)
lnRSA0_msec		0	0	0	0	12	1	Yes	OOOOOO (36); OOOOMO (11); OOOOMM (1)
lnAverage_SCL_uS		0	0	–	–	–	–	No	OO—- (48)
lnSDNN_msec		0	1	0	0	11	0	Yes	OOOOOO (36); OOOOMO (11); OMOOOO (1)
lnRMSSD_msec		0	1	0	0	11	0	Yes	OOOOOO (36); OOOOMO (11); OMOOOO (1)
lnHF_ms	T5	1	45	2	0	–	1	Yes	OMOO-O (42); OOOO-O (3); OMMO-O (2); MMOO-M (1)
lnLF_ms	T5	0	45	2	0	–	0	Yes	OMOO-O (43); OOOO-O (3); OMMO-O (2)
lnLFHF	T5	0	45	2	0	–	1	Yes	OMOO-O (42); OOOO-O (3); OMMO-O (2); OMOO-M (1)
PEP_msec		1	0	0	0	11	0	Yes	OOOOOO (36); OOOOMO (11); MOOOOO (1)

Note. Miss T1–T6 indicate the number of subjects missing at each time point (wide format). Pattern legend: O = observed, M = missing, – = structural (time point not collected or 100% missing). Non-monotone indicates the presence of at least one observed value following a missing value within a subject.

Under the missing at random (MAR) assumption, several valid approaches exist for the analysis of incomplete longitudinal data. These include direct likelihood-based methods, such as linear mixed-effects models, as well as multiple imputation (MI). In likelihood-based mixed-effects models, missing outcome data are handled implicitly through maximum likelihood or restricted maximum likelihood estimation, yielding unbiased parameter estimates under MAR (Verbeke and Molenberghs, 2000; Molenberghs and Verbeke, 2005). This approach constitutes the primary analytical framework used in the present study and is described in detail in the Statistical Analysis section.

Multiple imputation –first introduced by Donald Rubin in the late 1970s– represents an alternative strategy for handling incomplete data under MAR and provides a flexible framework for generating completed datasets. Rather than replacing each missing value with a single estimate, MI replaces missing values with multiple plausible values drawn from a predictive distribution, thereby accounting for uncertainty associated with missingness. This process results in M completed datasets, each of which can be analyzed using standard complete-data methods, with results subsequently combined using established pooling rules.

Several imputation strategies are available, and the choice of method depends on the missingness pattern and the types of variables involved. Fully Conditional Specification (FCS) is a commonly used approach for datasets with non-monotone missingness. When combined with Predictive Mean Matching (PMM), FCS imputes missing values by drawing observed values from cases with similar predicted means, preserving the empirical distribution of the data and ensuring plausible imputations for continuous variables.

In the present study, multiple imputation under MAR was not used as the primary estimation method. Instead, MI served as a preparatory step to facilitate sensitivity analyses under missing not at random (MNAR) assumptions using pattern-mixture models. Details of the imputation procedure and subsequent MNAR analyses are provided in later sections.

3.2.4 Exploratory data analysis

Following data preprocessing and transformation, exploratory data analysis was conducted on the observed HRV/SCL data to examine distributional properties, variability across time points, and patterns across experimental conditions and prenatal maternal anxiety (PMA) groups. These exploratory analyses were descriptive in nature and were used to inform subsequent modeling decisions, without drawing inferential conclusions.

In line with the exploratory framework described by Verbeke and Molenberghs (2000), the following aspects were examined separately for each outcome variable:

- Mean outcome values stratified by (1) time segment, (2) condition (rest-like vs. stress), (3) PMA group (low vs. high), and (4) the interaction between condition and PMA group;
- Variance patterns across the same stratifications;
- Correlation structure across repeated measurements, explored using semi-variograms. Semi-variograms decompose sources of variability by illustrating how variance changes as a function of the time lag between observations, thereby providing insight into the relative contributions of random effects, measurement error, and serial correlation;
- Individual trajectories (profiles), examined across time segments, condition, PMA group, and their interaction.

Exploratory visualizations were generated primarily using the `ggplot2` package in `RStudio`, while semi-variograms were constructed using SAS.

3.2.5 Model-building strategy for HRV and SCL outcomes (LMM)

Statistical analyses of HRV and SCL outcomes were conducted using linear mixed-effects models (LMMs) estimated via Direct Likelihood (DL). Direct Likelihood (DL) refers here to likelihood-based estimation via maximum likelihood or restricted maximum likelihood in linear mixed-effects models. Under the missing at random (MAR) assumption, likelihood-based estimation yields valid and unbiased inference without requiring explicit imputation of missing outcome values, provided that the missingness mechanism is ignorable (Verbeke and Molenberghs, 2000; Molenberghs and Verbeke, 2005). This likelihood-based framework constituted the primary analytical approach in the present study.

Because linear mixed-effects models require specification of three main components—the mean structure, the between-subject variability (random effects), and the within-subject covariance structure—a structured and strategic model-building procedure was followed. The general approach described by Verbeke and Molenberghs (2000) was used to guide model development for each outcome variable. The following steps were implemented.

1. **Initial mean structure and OLS residuals.** For each outcome variable, the most elaborate possible ordinary least squares (OLS) regression model was first fitted. This model included Condition (rest vs. stress), prenatal maternal anxiety (PMA) group, Gender, and all two-way interactions. The purpose of this step was to remove systematic trends in the data and to obtain OLS residual profiles, which were subsequently

used to inform decisions regarding the random-effects and residual covariance structures.

2. **Random-effects structure.** A random-intercept model was selected as the final random-effects structure for all outcomes. Random-slope models for Condition were also evaluated. Although these models converged, the estimated random-slope variance was consistently close to zero across outcomes and did not meaningfully improve model fit. This conclusion was further supported by inspection of residual profiles stratified by the condition variable (`REST_VS_STRESS`), which showed comparable dispersion across conditions, indicating no substantial heterogeneity in variance between rest-like and stress periods. Given the sample size (N = 48) and the repeated-measures design, the random-intercept specification provided a parsimonious and stable representation of between-subject variability.

3. **Residual covariance structure.** The residual covariance structure was chosen by fitting mixed models to the OLS residuals with no fixed-effects structure (i.e., $E(\varepsilon_i) = 0$), a random intercept, and candidate residual covariance structures. Two primary candidates were considered: independent homogeneous residuals (SIMPLE) and a first-order autoregressive structure [AR(1)]. Model selection was guided by a combination of Akaike Information Criterion (AIC), convergence behavior, and residual diagnostics. Semi-variograms were used at this stage as a diagnostic tool to assess the presence and strength of serial correlation, rather than for inferential purposes. In most cases, semi-variograms suggested weak but systematic positive serial correlation, and AR(1) models showed slightly improved fit relative to SIMPLE models without convergence issues; accordingly, AR(1) was selected in these cases. Alternative structures, including compound symmetry and heterogeneous covariance structures (e.g., ARH(1), CSH), were also evaluated but did not provide additional improvement.

4. **Fixed-effects reduction.** After selecting the random-effects and residual covariance structures, the fixed-effects structure was simplified. Because likelihood ratio tests for fixed effects are not valid under restricted maximum likelihood (REML), comparisons of fixed-effects structures were conducted using maximum likelihood estimation. Interaction terms with uninformative, non-significant effects were removed sequentially, except for the key interaction between Condition and PMA group, which was retained at this stage due to its theoretical relevance.

   After refitting the selected model under REML, the Condition × PMA group interaction was consistently non-significant across outcomes. In such cases, a second, more parsimonious final model was fitted in which this interaction term was removed. Both model specifications are reported, allowing assessment of the robustness of results to inclusion of the theoretically motivated interaction term.

Special considerations applied to the analysis of skin conductance level (SCL). Because only two SCL segments were available per participant and neither corresponded to a pure rest condition, analyses focused on PMA group differences during stress periods. Accordingly, the initial mean structure for SCL included Group, Gender, and their interaction, and a

simple residual covariance structure was used. The final SCL model assessed main effects of PMA group and Gender.

Table 7 summarizes the final model specification for each outcome variable. All models included a random intercept for Subject_ID. Model assumptions were evaluated using residual histograms and Q–Q plots.

Table 7: Summary of final mixed models for HRV and SCL outcomes. All models include a random intercept for Subject_ID.

Outcome	Final model (1)	Final -main- model (2)	Residual covariance
lnRMSSD, lnSDNN, $\ln RSA_0$, lnLF	Group, Condition, Gender, Group×Condition	Group, Condition, Gender	AR(1)
PEP, lnHF	Group, Condition, Gender, Group×Condition	Group, Condition, Gender	SIMPLE
ln(LF/HF)	Group, Condition, Gender, Group×Condition, Condition×Gender	Group, Condition, Gender, Condition×Gender[a]	SIMPLE
lnSCL	Group×Gender	Group, Gender	SIMPLE

[a] For lnLFHF, the Group×Condition interaction was dropped in the final model; the Condition×Gender interaction and main effect of Group were retained.

3.2.6 MNAR sensitivity analysis (delta adjustment)

As indicated earlier, the missing at random (MAR) assumption was adopted for the primary analyses. Under MAR, missingness is assumed to depend only on observed data and not on the unobserved response itself, conditional on the observed variables and covariates. However, in psychophysiological stress paradigms, missingness may plausibly depend on unobserved outcomes. For example, stress tasks often induce discomfort and movement, which may increase the likelihood of missing data among participants with stronger physiological stress responses (e.g., lower HRV). It is also plausible that such mechanisms may differ by prenatal maternal anxiety (PMA) group.

Because the true missingness mechanism cannot be verified empirically, a pattern-mixture sensitivity analysis using δ-adjustment was conducted to assess the robustness of the results under plausible departures from MAR in both directions. This approach evaluates how inferences change when values that were originally missing are systematically shifted relative to their values imputed under MAR.

As a preparatory step, missing outcome values were first imputed under the MAR assumption, providing completed datasets that served as the reference point for subsequent MNAR adjustments. In the long-format imputed dataset, values that were originally miss-

ing were explicitly flagged (e.g., $Miss_lnRMSSD = 1$) to distinguish them from originally observed values, which were left unchanged throughout the sensitivity analysis.

Three MNAR patterns were considered:

- **Stress pattern**: Stress segments (T2–T5) were flagged with an indicator $I_{stress} = 1$ for both PMA groups. This pattern represents departures from MAR where missingness is associated with being in a stress condition, irrespective of PMA group.
- **High PMA group**: Across rest and stress segments, observations from participants in the high PMA group were flagged with an indicator $I_{high} = 1$. This pattern represents departures from MAR where missingness is associated with PMA group membership, irrespective of condition.
- **High PMA under stress**: For stress segments only ($REST_VS_STRESS = 1$), observations from participants in the high PMA group ($Group = 1$) were flagged with an indicator $I_{highstress} = 1$. This pattern represents departures from MAR where missingness is increased specifically among high-PMA participants during stress.

For each pattern, a symmetric grid of δ values was specified:

$$\delta \in \{-0.3, -0.2, -0.1, 0, 0.1, 0.2, 0.3\},$$

where $\delta = 0$ corresponds to the MAR baseline. Negative values correspond to systematically lower values than those imputed under MAR (e.g., $\delta = -0.1$, -0.2, -0.3 correspond to approximately 10%, 18%, and 26% lower HRV values, respectively), while positive values correspond to systematically higher values (approximately 10%, 22%, and 35% higher, respectively). Including both negative and positive δ values provides a conservative assessment by allowing for departures from MAR in either direction.

For a given MNAR pattern and δ value, δ-adjustments were applied only to values that were originally missing. Specifically, for a given outcome (e.g., lnSDNN), the adjusted value was defined as:

$$lnSDNN_{MNAR} = lnSDNN_{MAR} + \delta \times patternindicator,$$

where the pattern indicator corresponds to I_{stress}, I_{high}, or $I_{highstress}$. Adjustments were applied on the log scale for all log-transformed outcomes; for PEP, adjustments were applied on the original scale. Skin conductance level (SCL) was excluded from the MNAR sensitivity analysis because it was fully observed.

This procedure resulted in 21 MNAR-adjusted datasets (three patterns $\times$ seven δ values), each based on the MAR-imputed data (with 10 imputations). For each adjusted dataset and each outcome, both final mixed-effects model specifications were fitted: (i) the model including the key interaction between Condition and PMA group, and (ii) the reduced model in which this interaction was removed. This allowed direct evaluation of the sensitivity of both the primary interaction of interest and the corresponding main-effects model to departures from the MAR assumption.

Model estimates, standard errors, confidence intervals, and p-values were obtained for each adjusted dataset and combined using standard pooling rules. This was done in SAS using PROC MIANALYZE. Sensitivity was evaluated according to three scenarios:

- Results remain consistent across all MNAR patterns and δ values, indicating robustness to a wide range of departures from MAR;
- Results differ from MAR-based conclusions only for extreme δ values, suggesting sensitivity only under strong MNAR assumptions;
- Results differ from MAR-based conclusions even for small δ values, indicating sensitivity to modest departures from MAR.

3.3 Methodology for Cortisol and sAA Data

3.3.1 Data preprocessing and transformations

Salivary cortisol and salivary alpha-amylase data were prepared for analysis by reshaping the original wide-format laboratory output into a long-format dataset reflecting the repeated-measures structure across predefined sampling occasions. Each observation corresponds to a specific participant and saliva sample indexed by the variable `Sample` (1–8).

Due to pronounced right-skewness in the distributions of cortisol, alpha-amylase, and their ratio, all salivary outcomes were analyzed on the natural logarithmic scale. Log-transformation was applied prior to statistical modeling to improve distributional properties and stabilize variance across measurement occasions.

Descriptive summaries and distributional characteristics of the salivary variables on the original and log-transformed scales are reported in Tables 8 and 9, respectively. Histograms and normal quantile–quantile plots illustrating the distributions before and after transformation are shown in Figures 6-9. Corresponding histograms and Q–Q plots for sAA and AACR are provided in the Appendix A.

Table 8: Original-scale salivary variables and applied transformations

Variable	Unit	Range	Transformation
Cortisol	nmol/L	0.26 – 9.60	Natural log
Alpha-amylase	U/mL	11.03 – 9535.49	Natural log
AACR	–	4.19 – 31784.97	Natural log

Table 9: Summary statistics of log-transformed salivary variables

Variable (unit)	Min	Mean	SD	Max
lnCortisol (nmol/L)	-1.35	0.32	0.69	2.26
lnAmy (U/mL)	2.40	6.82	1.12	9.16
lnAACR (–)	1.43	6.48	1.35	10.37

3.3.2 Missing data

Missingness was present in both the salivary outcomes and the time-related covariates. Table 10 summarizes the number of observed measurements per saliva sampling occasion for the

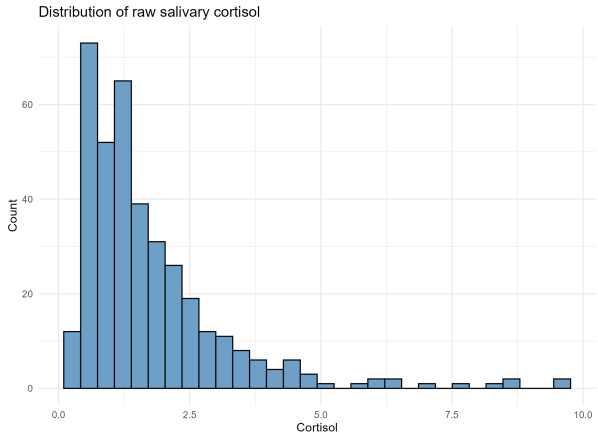

Figure 6: Histogram of Salivary Cortisol on the raw scale (pre-transformation).

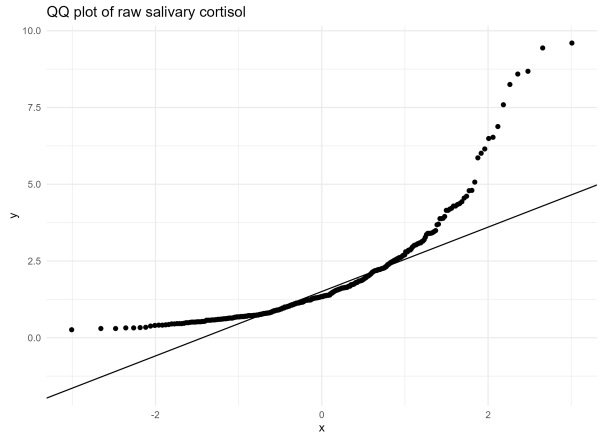

Figure 7: Q–Q plot of Salivary Cortisol on the raw scale (pre-transformation).

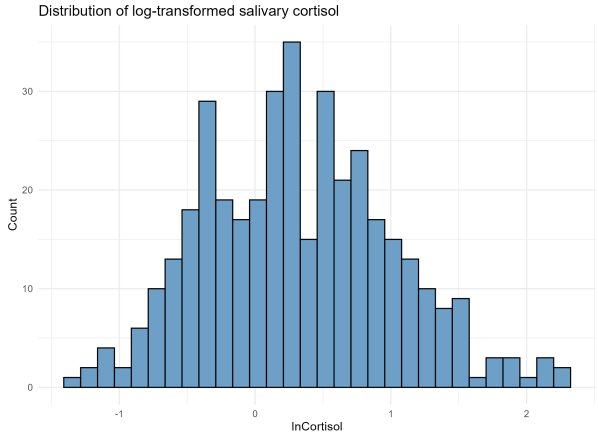

Figure 8: Histogram of lnCortisol after logarithmic transformation.

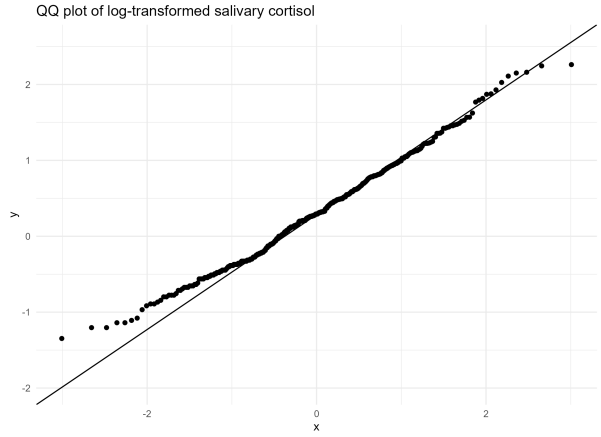

Figure 9: Q–Q plot of lnCortisol after logarithmic transformation.

log-transformed salivary outcomes, indicating a non-monotone missing-data pattern across participants. The time-from-CPT variable was fully observed for samples S1–S4 (N = 51), with observed sample sizes decreasing to 50 at S5, 48 at S6, and 45 at both S7 and S8. Missingness in this variable resulted from incomplete availability of sampling-time information.

The handling of missing data and the assumptions underlying the statistical analyses are addressed in the following sections.

Table 10: Observed sample sizes by saliva sampling occasion for log-transformed salivary outcomes

Variable	S1	S2	S3	S4	S5	S6	S7	S8
lnCortisol	46	50	50	49	48	48	44	45
lnAlpha-amylase	46	49	48	45	46	47	43	43
lnAACR	43	48	47	44	46	47	43	43

Note. Values indicate the number of observed measurements at each predefined saliva sampling occasion. Missingness followed a non-monotone pattern across participants.

3.3.3 Exploratory analysis of salivary cortisol and alpha-amylase

Similar to what was mentioned in Section 3.2.4, exploratory analyses were conducted to examine the distributional characteristics, temporal patterns, and associations between the salivary outcomes prior to statistical modeling. Mean levels and variability of the log-transformed salivary measures were inspected as a function of time from the cold pressor task (CPT), stratified by prenatal maternal anxiety (PMA) group and sex.

In addition, correlations between log-transformed cortisol, alpha-amylase, and their ratio were examined, and individual-level profiles were visually inspected to assess within-subject variability and heterogeneity in response patterns over time from CPT. These exploratory analyses informed the subsequent modeling strategy.

3.3.4 Model-building strategy for salivary cortisol and alpha-amylase

In studies of salivary stress biomarkers, analyses often rely on comparisons with a pre-stressor baseline collected a fixed period before task onset. In the present study, analyses focused on modeling salivary outcomes as a function of time since the cold pressor task (CPT). Earlier sampling occasions were not used as a baseline because the physiological state of participants prior to CPT could not be reliably determined. In particular, sample S3 was collected during a stress-related task, and the elapsed time between this task and the CPT varied substantially across participants, making it unclear whether individuals had returned to a comparable physiological state before CPT onset. Moreover, salivary cortisol and alpha-amylase exhibit distinct temporal response profiles to acute stress, with cortisol peaking later than alpha-amylase (about 20-30 minutes after a stressor for cortisol, much more quickly for sAA), which further complicates assumptions about uniform stress or recovery phases across sampling occasions. As a result, modeling trajectories relative to CPT onset provided a consistent temporal reference point across participants. Modeling time since CPT as a continuous variable allowed these differential response dynamics to be captured without imposing assumptions about discrete stress phases.

Salivary cortisol, alpha-amylase, and AACR trajectories were analyzed using linear mixed-effects models to account for the repeated-measures structure of the data. Model building followed the same strategy as described in Section 3.2.5 for the HRV analyses, using direct likelihood estimation under the MAR assumption. Fixed effects included centered time since CPT, prenatal maternal anxiety group, gender, and interactions with centered time since CPT, with outcome-specific extensions to allow for nonlinear time effects where supported by the data. The random effects structure was initially specified for the intercept as well as linear and quadratic time terms and were subsequently reduced based on model fit criteria, resulting in a random-intercept specification. An autoregressive AR(1) structure was selected to model serial correlation between repeated measurements. Table 11 summarizes the final model specification for each outcome variable. Model assumptions were evaluated using residual histograms and Q–Q plots.

During exploratory data analysis, two late post-CPT observations with unusually large time-since-CPT values were identified. Figure 10 illustrates the distribution of time since CPT, highlighting these observations by participant ID. To assess their potential influence on model estimates, the final mixed-effects models were refitted after excluding these two observations. The resulting estimates led to the same substantive conclusions across outcomes; however, for cortisol, the direction of the interaction term differed while remaining statistically non-significant.

In addition to trajectory-based mixed-effects models, secondary analyses were conducted using summary metrics derived from the salivary time series. Specifically, delta (defined as the difference between the peak response and the concentration measured at Sample 2, used as a predefined reference point), peak level, and recovery slope were considered, as these metrics are commonly used to summarize endocrine response magnitude and recovery. Summary metrics were constructed from the multiply imputed salivary data, thereby avoiding bias due to incomplete sampling and enabling subsequent sensitivity analyses under MNAR assumptions. Associations between prenatal maternal anxiety group, gender, and the derived metrics were examined using linear regression models, with parameter estimates and standard errors pooled across imputations.

Table 11: Final linear mixed-effects models for salivary cortisol, alpha-amylase, and AACR as a function of time since CPT

Fixed effect	Cortisol M1	Cortisol M2	Alpha-amylase	AACR
Intercept	×	×	×	×
Centered time (t_c)	×	×	×	×
Centered time^2 (t_c^2)			×	×
Group (high vs low/medium)	×	×	×	×
Gender (female vs male)	×	×	×	×
Group × time (t_c)	×			
Gender × time (t_c)				
Group × gender			×	×
Group × time^2 (t_c^2)			×	×
Gender × time^2 (t_c^2)			×	×
Random effects	Random intercept (subject)			
Residual structure	AR(1)			
Estimation	REML			

Note. × indicates that the term is included in the model. Time since CPT was centered. Cortisol Model 1 includes a group × time interaction; Cortisol Model 2 represents the final reduced cortisol model. Alpha-amylase and AACR models include quadratic time effects and quadratic time interactions.

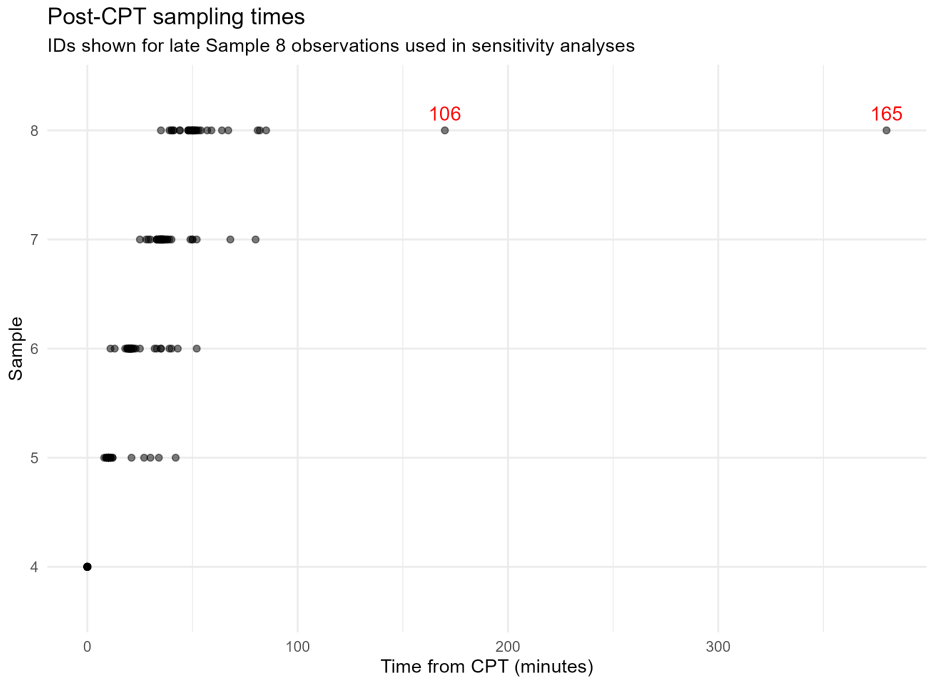

Figure 10: Distribution of time since CPT with identification of two late post-CPT observations by participant ID.

3.3.5 Sensitivity analysis under MNAR assumptions

To assess the robustness of the findings to potential departures from the MAR assumption, sensitivity analyses were conducted under a missing-not-at-random (MNAR) framework using a delta-adjustment approach. This approach assumes that unobserved salivary values may systematically differ from observed values by a fixed amount on the log scale.

Delta adjustments were applied separately for earlier post-CPT samples (Samples 4–6) and later recovery samples (Samples 7–8), reflecting the possibility that missingness mechanisms may differ across phases of the recovery period. For early post-CPT samples, delta values of 0, ± 0.1, and ± 0.2 were considered, while for later recovery samples, a wider grid including 0, ± 0.1, ± 0.2, and ± 0.3 was examined. Delta shifts were applied only to observations that were missing in the original data.

For each combination of early and recovery delta values, the final analysis models were refitted and results were pooled across imputations. The direction and statistical significance of the main effects were compared with the corresponding MAR-based results to evaluate sensitivity to MNAR assumptions.

4 Results

4.1 HRV and SCL Results

4.1.1 MAR results

Participant characteristics. The HRV/SCL dataset included 48 offspring participants, of whom 35 belonged to the low-to-medium prenatal maternal anxiety (PMA) group and 13 to the high PMA group. The distribution of participants by PMA group and gender is shown in Table 12.

Interaction models (Group $\times$ Condition). The first set of analyses (mentioned in Table 7) examined whether the effect of stress differed between PMA groups by including an interaction between PMA group and condition (rest vs. stress). Results from these interaction models are presented in Table 13. Across all outcomes, the Group $\times$ Condition interaction was not statistically significant. For all variables, confidence intervals included the null value (original scale ratio = 1 for log-transformed outcomes after back transformation; estimate = 0 for PEP), and all corresponding p-values exceeded the 0.05 significance level. These findings indicate no evidence that stress-related changes in HRV, SCL, or PEP differed between the low-to-medium and high PMA groups.

Main-effects models. Given the absence of significant interaction effects, reduced mixed-effects models including only main effects of condition, PMA group, and gender were fitted (except LFHF, which included Gender $\times$ Condition interaction). Results from these final models are presented in Table 14.

For RMSSD, there was no evidence of a statistically significant effect of stress, PMA group, or gender. On average, RMSSD during stress was estimated to be approximately 1% lower than during rest (ratio = 0.99, 95% CI: 0.92–1.07, $p = 0.82$). After accounting

for condition and gender, no significant difference in RMSSD was observed between PMA groups (ratio = 1.14, 95% CI: 0.76–1.71, $p = 0.52$).

In contrast, SDNN was significantly higher during stress compared with rest after accounting for gender and PMA group. Specifically, SDNN during stress was estimated to be approximately 16% higher than during non-stress periods (ratio = 1.16, 95% CI: 1.08–1.24, $p < 0.001$). There was no evidence of a PMA group difference in SDNN after accounting for condition and gender (ratio = 1.10, 95% CI: 0.85–1.43, $p = 0.47$). Gender differences in SDNN did not reach statistical significance.

For RSA, no statistically significant effects of stress, PMA group, or gender were observed. RSA during stress did not differ significantly from rest (ratio = 1.11, 95% CI: 0.97–1.21, $p = 0.18$), and no evidence of a PMA group difference was found (ratio = 1.22, 95% CI: 0.89–1.70, $p = 0.21$).

Analysis of frequency-domain HRV indices showed that stress was associated with a marked increase in LF power. After accounting for gender and PMA group, the average LF during stress was approximately 66% higher than during rest (ratio = 1.66, 95% CI: 1.42–1.93, $p < 0.0001$). After adjusting for condition and PMA group, females exhibited significantly lower LF power compared with males (ratio = 0.64, 95% CI: 0.41–1.00, $p = 0.044$). No statistically significant difference in LF power was observed between PMA groups.

For HF power, no statistically significant main effects of stress, PMA group, or gender were observed in the final model. Although HF was higher during stress, this effect did not reach statistical significance (ratio = 1.18, 95% CI: 0.92–1.25, $p = 0.43$).

For the LF/HF ratio, a statistically significant Gender × Stress interaction was observed, indicating that the effect of stress on LF/HF differed by gender after accounting for PMA group. Among males (the reference group), stress was associated with a significant increase in LF/HF (ratio = 1.33, 95% CI: 1.11–1.60, $p = 0.003$). The significant interaction term (ratio = 1.30, 95% CI: 1.02–1.71, $p = 0.034$) indicates that the stress-related increase in LF/HF was approximately 30% larger in females compared with males. After accounting for condition and gender, no statistically significant difference in LF/HF was observed between PMA groups.

Skin conductance level (SCL) was analyzed only during stress periods. After adjusting for gender, participants in the high PMA group exhibited significantly lower SCL compared with those in the low-to-medium PMA group (ratio = 0.12, 95% CI: 0.03–0.51, $p = 0.005$), corresponding to an approximately 88% lower mean SCL in the high PMA group. No evidence of a gender difference in SCL during stress was observed.

Finally, analysis of PEP showed a strong effect of stress. On average, PEP during stress was approximately 10–11 ms shorter than during rest, indicating increased cardiac sympathetic activation during stress (estimate = −10.55 ms, 95% CI: −15.51 to −5.66, $p < 0.0001$). No statistically significant effects of PMA group or gender on PEP were observed.

Table 12: Distribution of participants by prenatal anxiety group and gender (HRV/SCL dataset).

Group	Females	Males
Low/medium prenatal anxiety (0)	20	15
High prenatal anxiety (1)	7	6

Table 13: Interaction model results (Group × Condition)

Outcome	Ratio / Estimate[a]	SE [b]	95% CI (Lower)[c]	95% CI (Upper)[c]	p-value
lnRMSSD	1.10	0.090	0.922	1.310	0.294
lnSDNN	1.000	0.081	0.855	1.170	0.996
$\ln RSA_0$[d]	1.070	0.107	0.867	1.320	0.530
lnHF	1.236	0.186	0.850	1.769	0.275
lnLF	1.034	0.174	0.733	1.460	0.851
ln(LF/HF)	0.951	0.141	0.723	1.241	0.693
PEP	1.810	5.671	-9.378	12.980	0.750

Ratios and confidence intervals are reported on the original scale for log-transformed outcomes; for PEP, estimates and confidence intervals are reported on the raw scale.

Standard errors and p-values are reported on the model scale.

RSA and SCL models were fitted on ln(RSA + 1) and ln(SCL + 0.01), respectively; reported ratios and confidence intervals therefore pertain to the shifted scale.

Main effects are not interpreted marginally in the presence of a statistically significant interaction.

4.1.2 MNAR sensitivity analysis results

The MNAR sensitivity analyses described in Section 3.2.6 were conducted to assess the robustness of the results obtained under the Missing At Random (MAR) assumption. Across all MNAR patterns and δ values considered, the direction, magnitude, and statistical significance of almost all the parameter estimates (including confidence intervals, standard errors, and p-values) remained highly consistent with the corresponding results obtained under MAR using direct likelihood estimation.

In particular, no substantive changes were observed for the key effects of interest, including the Condition × PMA group interaction and other retained fixed effects in the final models. This consistency was observed across all examined departures from MAR, including stress-related, group-related, and stress-by-group–specific missingness patterns.

Overall, these findings indicate that the main conclusions drawn from the MAR-based analyses are robust to plausible deviations from the MAR assumption.

4.2 Cortisol and sAA Results

4.2.1 Time-since-CPT mixed models

Salivary cortisol (lnCort). As shown in Tables 15, the interaction between time (centered) and PMA group was not statistically significant. The estimated ratio was close to unity, and the confidence interval included 1, indicating no evidence that temporal changes in cortisol differed between prenatal anxiety groups. Table 16 shows the results for the main effects linear mixed model for salivary cortisol, none of the fixed effects showed statistically significant associations with cortisol levels. The estimated group effect indicated that individuals in the high prenatal maternal anxiety group had, on average, cortisol levels that

Table 14: Final mixed-effects models (main effects)

Outcome	Effect	Ratio[a]	SE	95% CI (Lower)[b]	95% CI (Upper)[c]	p-value
lnRMSSD	Stress vs non-stress	0.991	0.04	0.917	1.071	0.820
	High vs low/med PMA	1.142	0.201	0.761	1.714	0.520
	Female vs male	0.771	0.179	0.540	1.110	0.155
lnSDNN	**Stress vs non-stress**	**1.157**	**0.035**	**1.080**	**1.240**	**<0.001**
	High vs low/med PMA	1.101	0.131	0.846	1.433	0.467
	Female vs male	0.803	0.117	0.634	1.016	0.067
lnRSA[e]	Stress vs non-stress	1.111	0.060	0.971	1.210	0.178
	High vs low/med PMA	1.224	0.160	0.887	1.70	0.213
	Female vs male	0.894	0.143	0.670	1.193	0.440
lnLF	**Stress vs non-stress**	**1.661**	**0.077**	**1.423**	**1.930**	**<0.0001**
	High vs low/med PMA	1.141	0.242	0.703	1.809	0.5884
	Female vs male	**0.640**	**0.221**	**0.413**	**0.998**	**0.044**
lnHF	Stress vs non-stress	1.177	0.082	0.919	1.252	0.427
	High vs low/med PMA	1.610	0.377	0.751	3.416	0.220
	Female vs male	0.722	0.338	0.366	1.431	0.3548
ln(LF/HF)	**Stress vs non-stress[d]**	**1.331**	**0.092**	**1.111**	**1.601**	**0.003**
	High vs low/med PMA	0.683	0.197	0.460	1.020	0.061
	Female vs male[d]	0.822	0.187	0.565	1.211	0.301
	Gender × Stress	**1.300**	**0.122**	**1.021**	**1.710**	**0.034**
lnSCL[e]	**High vs low/med PMA**	**0.120**	**0.725**	**0.028**	**0.513**	**0.005**
	Female vs male	1.088	0.649	0.295	4.014	0.879
PEP	**Stress vs non-stress**	**-10.551**	**2.515**	**-15.510**	**-5.660**	**<0.0001**
	High vs low/med PMA	-0.761	5.580	-11.995	10.473	0.892
	Female vs male	-0.393	4.997	-10.458	9.671	0.948

Back-transformed to the original scale; SE and p-values are reported on the log scale, except for PEP, which is reported on the raw scale.

RSA and SCL models were fitted on ln(RSA + 1) and ln(SCL + 0.01), respectively; ratios and confidence intervals therefore pertain to the shifted scale.

Not marginally interpreted since there is an interaction.

were approximately 9.5% higher than those in the low/medium group; however, the confidence interval was wide and included unity, indicating substantial uncertainty. The effect of centered time was close to null, suggesting no evidence of systematic change in cortisol levels over time. Gender was also not significantly associated with cortisol, with females showing slightly lower levels than males on average, though this difference was not statistically significant. Overall, the model did not provide evidence for differences in cortisol levels associated with prenatal maternal anxiety group, time, or gender.

Salivary alpha-amylase (lnAmy). Results from the linear mixed model for the alpha-amylase are presented in Table 17. In the linear mixed model for salivary alpha-amylase, a significant association with linear time was observed, with alpha-amylase levels decreasing by approximately 0.6% per unit increase in centered time. The quadratic time term was not statistically significant, indicating no evidence for curvature in the time trajectory. Prenatal maternal anxiety group was not significantly associated with overall alpha-amylase levels. A significant gender effect was observed, with females showing substantially lower alpha-amylase levels compared to males at the reference time point in the low/med PMA group, corresponding to an estimated reduction of approximately 67%. In addition, a statistically significant interaction between prenatal maternal anxiety group and gender was observed, indicating that female participants in the high prenatal anxiety group exhibited approximately fourfold higher alpha-amylase levels at the reference time point compared to females in the low/medium group. Quadratic time and its interactions with prenatal maternal anxiety group and gender were not statistically significant, suggesting no evidence for group- or gender-specific differences in the curvature of the response.

Alpha-amylase-to-cortisol ratio (lnAACR). Results from the linear mixed model for the alpha-amylase-to-cortisol ratio are presented in Table 18. Neither linear nor quadratic time effects were statistically significant, indicating no evidence for systematic changes in lnAACR over the measurement period. Similarly, no significant interactions between quadratic time and prenatal maternal anxiety group or gender were observed, suggesting that the temporal profile of lnAACR did not differ by prenatal anxiety exposure or gender.

A significant baseline gender effect was observed, with females showing lower lnAACR values compared to males at the reference time point in the low/medium prenatal maternal anxiety group, corresponding to an estimated reduction of approximately 65%. Prenatal maternal anxiety group and the group-by-gender interaction were not statistically significant, indicating no evidence for overall group differences in lnAACR.

Table 15: Interaction between time (centered) and prenatal maternal anxiety group on salivary cortisol (lnCort). Estimates and confidence intervals are back-transformed and presented as ratios.

Effect	Ratio	95% CI	SE (log)	p-value
Time × Group	1.001	[0.998, 1.004]	0.00164	0.523

Table 16: Main effects linear mixed model for salivary cortisol (lnCort). Estimates and confidence intervals are back-transformed and presented as ratios.

Effect	Ratio	95% CI	SE (log)	p-value
Intercept	1.215	[0.924, 1.597]	0.1348	0.1548
t_c	1.000	[0.999, 1.000]	0.000777	0.7052
Group (PMA high vs low/med)	1.095	[0.740, 1.619]	0.1980	0.6479
Gender (female vs male)	0.886	[0.628, 1.250]	0.1727	0.4883

Table 17: Linear mixed model for salivary alpha-amylase (lnAmy). Estimates and confidence intervals are back-transformed and presented as ratios.

Effect	Ratio	95% CI	SE (log)	p-value
Intercept (t=0, PMA=low/med, gender=male)	1229.0	[759.0, 1950.9]	0.2430	<0.0001
t_c (linear time)	**0.994**	**[0.9898, 0.9988]**	**0.002263**	**0.0140**
t_{c2} (quadratic time; PMA=low/med, gender=male)	1.000	[0.9996, 1.0004]	0.000179	0.9670
Group (PMA high vs low/med, t=0, gender=male)	0.683	[0.264, 1.771]	0.4729	0.4246
Gender (female vs male, t=0, PMA=low/med)	**0.331**	**[0.169, 0.648]**	**0.3341**	**0.0017**
Group × Gender (female vs male, t=0)	**3.918**	**[1.072, 14.39]**	**0.6437**	**0.0393**
$t_{c2}\times$ Group (PMA high vs low/med, gender=male)	1.000	[0.9997, 1.0004]	0.000179	0.9579
$t_{c2}\times$ Gender (female vs male, PMA=low/med)	1.000	[0.9997, 1.0004]	0.000180	0.6128

Table 18: Linear mixed model for the alpha-amylase-to-cortisol ratio (lnAACR). Estimates and confidence intervals are back-transformed and presented as ratios.

Effect	Ratio	95% CI	SE (log)	p-value
Intercept (t=0, PMA=low/med, gender=male)	975.1	[537.3, 1773.8]	0.3082	<0.0001
t_c (linear time)	0.997	[0.9915, 1.0016]	0.002516	0.1737
t_{c2} (quadratic time; PMA=low/med, gender=male)	1.000	[0.9997, 1.0005]	0.000198	0.6873
Group (PMA high vs low/med, t=0, gender=male)	0.623	[0.185, 2.096]	0.6038	0.4360
Gender (female vs male, t=0, PMA=low/med)	**0.350**	**[0.145, 0.824]**	**0.4253**	**0.0172**
Group × Gender (female vs male, t=0)	4.611	[0.878, 22.39]	0.8233	0.0699
$t_{c2}\times$ Group (PMA high vs low/med, gender=male)	1.000	[0.9995, 1.0003]	0.000198	0.7243
$t_{c2}\times$ Gender (female vs male, PMA=low/med)	1.000	[0.9996, 1.0004]	0.000199	0.9552

4.2.2 Summary response metrics

Delta (peak - S2) Delta models assessing change from baseline yielded consistent conclusions across outcomes with respect to temporal change. For salivary cortisol (lnCort) and the alpha-amylase-to-cortisol ratio (lnAACR), no evidence was found for group- or gender-specific differences in change from baseline, in line with the longitudinal mixed-effects analyses. For salivary alpha-amylase (lnAmy), delta models likewise showed no group- or gender-specific differences in change from baseline; however, this does not preclude baseline subgroup differences observed in the longitudinal model, as delta analyses do not capture effects confined to the reference time point.

Slope recovery Slope recovery models assessing post-stress recovery trajectories yielded consistent null findings across outcomes. For salivary cortisol (lnCort), salivary alpha-amylase (lnAmy), and the alpha-amylase-to-cortisol ratio (lnAACR), no statistically significant interactions between prenatal maternal anxiety group and gender were observed, indicating no evidence for differential recovery slopes between subgroups. Furthermore, main effects of prenatal maternal anxiety group and gender were not statistically significant for any outcome, suggesting comparable recovery rates across groups and genders.

Peak response Peak response analyses revealed a statistically significant interaction between prenatal maternal anxiety group and gender for salivary alpha-amylase (lnAmy). Specifically, female participants in the high prenatal anxiety group exhibited higher peak alpha-amylase levels compared to females in the low/medium group, corresponding to an approximately 3.7-fold increase (ratio 3.7, 95% CI 1.18–11.4, $p = 0.025$). A significant main effect of gender was also observed, with females showing lower peak alpha-amylase levels compared to males. No significant main effect of prenatal maternal anxiety group alone was detected. In contrast, peak cortisol (lnCort) and the alpha-amylase-to-cortisol ratio (lnAACR) showed no significant interaction or main effects.

4.2.3 MNAR sensitivity analysis

Sensitivity analyses under missing not at random (MNAR) assumptions were conducted to assess the robustness of the results obtained under the missing at random (MAR) framework. Across all salivary outcomes (lnCort, lnAmy, and lnAACR), the direction and statistical significance of the main findings were largely unchanged under the MNAR models. Specifically, conclusions regarding time effects, group and gender effects, and interaction terms remained consistent with those obtained from the primary longitudinal mixed-effects analyses.

No MNAR specification led to a meaningful change in the interpretation of the results, indicating that the findings are robust to plausible departures from the MAR assumption.

5 Discussion

This study examined the long-term association between prenatal maternal anxiety and adult offspring stress physiology using complementary autonomic and endocrine markers, including heart rate variability (HRV), salivary cortisol, salivary alpha-amylase, and their ratio. Taken

together, the findings suggest limited evidence for pervasive or enduring dysregulation of stress systems in adulthood, while highlighting selective and system-specific associations.

Across HRV measures, stress-related modulation of autonomic activity was evident, supporting the validity of the experimental paradigm. In particular, pre-ejection period (PEP) showed a robust stress effect, indicating sympathetic activation during stress exposure. The increase in SDNN during stress likely reflects the inclusion of both sympathetic and parasympathetic influences and increased overall variability under task engagement, rather than enhanced vagal control. In addition, contrary to our initial hypothesis of higher sympathetic arousal, skin conductance levels were lower in the high prenatal anxiety group. Importantly, however, HRV analyses did not reveal consistent group-by-condition interactions, suggesting that while baseline or tonic autonomic differences may be present, prenatal maternal anxiety was not associated with altered autonomic reactivity or recovery dynamics during stress.

Findings from the salivary biomarkers largely converged with this pattern. For salivary cortisol, longitudinal trajectories, baseline-to-change metrics, recovery slopes, and peak responses provided no evidence for group- or gender-specific differences in hypothalamic-pituitary adrenal (HPA) axis activity. These null findings were consistent across multiple analytical approaches and were robust under MNAR sensitivity analyses, indicating no systematic alteration of adult cortisol stress responses associated with prenatal maternal anxiety.

Salivary alpha-amylase revealed a more nuanced pattern. While no group- or gender-specific differences were observed in time-dependent trajectories or recovery slopes, alpha-amylase levels at the reference time point in the CPT model and at peak response differed by prenatal maternal anxiety group in a gender-specific manner. These findings suggest that prenatal maternal anxiety may be associated with differences in sympathetic tone or responsivity in specific subgroups, particularly among females, rather than with generalized alterations in reactivity or recovery. The absence of corresponding effects in cortisol and most HRV reactivity measures indicates that these associations are outcome-specific and context-dependent.

The alpha-amylase-to-cortisol ratio further supported this interpretation, as no evidence was found for long-term alterations in the balance between sympathetic and HPA axis activity across longitudinal, summary, or sensitivity analyses. Together with the HRV findings, this suggests that prenatal maternal anxiety may be linked to subtle differences in tonic sympathetic arousal, as reflected in SCL and at the reference time point or peak alpha-amylase measures, rather than to broad dysregulation of stress system dynamics.

Overall, the convergence of HRV and salivary results indicates that prenatal maternal anxiety is not associated with widespread or persistent changes in adult stress reactivity or recovery. Where associations were observed, they were limited in scope, system-specific, and primarily evident at baseline or peak levels rather than in dynamic responses to stress. The consistency of conclusions across physiological systems, analytical strategies, and MNAR sensitivity analyses strengthens confidence in the robustness of these findings.

6 Possible drawbacks

Several methodological considerations should be taken into account when interpreting the results of this study.

First, although linear mixed-effects models offer flexibility for analyzing repeated measures, the inclusion of interaction terms and non-linear time effects renders some fixed effects conditional on reference levels of time and covariates. Accordingly, baseline contrasts were interpreted cautiously, with emphasis placed on time-dependent effects and interactions of substantive interest.

Second, stress responses were characterized using both longitudinal trajectories and summary metrics (delta change, recovery slopes, and peak responses). While this multi-metric approach improves robustness, summary measures inevitably simplify complex temporal dynamics and may be less sensitive to subtle differences in response shape or timing.

Third, not all available salivary samples could be reliably incorporated into the analyses. Substantial variability in the timing of saliva collection and in the temporal spacing between stress exposure and sampling meant that the full set of eight samples was not consistently comparable across participants. As a result, analyses focused on time windows with the most reliable temporal alignment, particularly those capturing recovery dynamics, which may have reduced sensitivity to detect effects occurring at earlier or more transient phases of the stress response.

Fourth, the sample size, while adequate for primary longitudinal analyses, may have limited power to detect small or subgroup-specific effects, especially in models including interactions between prenatal maternal anxiety and gender.

Finally, skin conductance level (SCL) was not assessed during resting baseline periods, limiting interpretation to tonic levels outside of explicit rest conditions. In addition, diagnostic checks for the lnSCL models indicated mild departures from model assumptions, which may have affected the precision of estimates involving this outcome.

7 Ethical thinking, societal relevance and stakeholder awareness

Ethics This study was conducted in accordance with ethical principles for research involving human participants. Ethical approval was obtained from the relevant institutional review boards, and all participants provided informed consent. Data were handled confidentially and analyzed in pseudonymized form to ensure participant privacy, particularly given the sensitive nature of prenatal mental health and physiological measurements.

Care was taken to interpret and communicate the findings responsibly. Prenatal maternal anxiety was examined as a potential risk factor rather than a causal determinant, and results were framed to avoid deterministic or stigmatizing interpretations. Emphasizing the absence of widespread long-term physiological dysregulation helps prevent maternal blame and supports a balanced understanding of developmental vulnerability and resilience.

Relevance The societal relevance of this study lies in its contribution to the evidence base on the long-term effects of prenatal maternal anxiety. By integrating autonomic and endocrine markers and applying robust analytical strategies, the findings suggest that prenatal anxiety is not associated with pervasive alterations in adult stress physiology, but may be linked to subtle, context-dependent differences. This perspective is relevant for public health

communication and clinical practice, where reassurance and prevention must be balanced.

Stakeholder Key stakeholders include pregnant individuals, healthcare and mental health professionals, researchers, and policymakers. The results underline the importance of supporting maternal mental health during pregnancy while also indicating that prenatal anxiety does not inevitably lead to long-term physiological impairment in offspring. Overall, ethical considerations and stakeholder perspectives were integrated throughout the study to ensure responsible conduct and communication of the research.

8 Conclusion

This thesis examined whether prenatal exposure to maternal anxiety is associated with long-term alterations in autonomic and endocrine stress regulation in adulthood. Using a combination of heart rate variability measures and salivary stress biomarkers, and applying robust longitudinal and sensitivity analyses, the study aimed to evaluate both stress reactivity and recovery processes.

Overall, the findings indicate that prenatal maternal anxiety is not associated with widespread or persistent dysregulation of adult stress physiology. While some outcome- and subgroup-specific differences were observed, particularly in tonic or baseline sympathetic markers, there was no consistent evidence for altered stress reactivity or recovery dynamics. These results suggest substantial resilience of adult stress regulatory systems following prenatal exposure to maternal anxiety.

By integrating multiple physiological systems and analytical approaches, this study contributes to a more nuanced understanding of prenatal programming effects, emphasizing variability and context-dependence rather than deterministic long-term impairment. The findings underscore the importance of cautious interpretation of prenatal risk factors and support a balanced perspective on developmental vulnerability and resilience.

9 Future research

Future research could build on the present findings in several ways. First, studies with larger sample sizes may be better powered to detect small or subgroup-specific effects, particularly interactions involving gender or other moderating factors. Replication in independent cohorts would further strengthen conclusions regarding the robustness and generalizability of the observed patterns.

Second, future studies may benefit from more tightly controlled and precisely timed stress paradigms, particularly for salivary biomarkers. Improved temporal alignment of sampling relative to stress exposure could enhance sensitivity to detect early reactivity and peak responses and allow more detailed modeling of non-linear recovery dynamics.

Third, integrating additional contextual and psychological variables, such as current stress, coping styles, or resilience factors, may help clarify why associations with prenatal maternal anxiety appear to be outcome- or subgroup-specific. Longitudinal designs that

track stress physiology across multiple life stages could further elucidate whether observed effects remain stable, emerge later, or attenuate over time.

Finally, combining physiological measures with neuroimaging, genetic, or epigenetic data may provide deeper insight into the biological pathways underlying potential prenatal programming effects. Such multimodal approaches could help distinguish between transient adaptations and long-term alterations in stress regulation.

Declaration on the use of generative AI tools

I acknowledge the use of a generative artificial intelligence tool (ChatGPT, OpenAI) during the preparation of this master's thesis. The tool was used in a supportive manner to improve the clarity, grammar, and academic style of the manuscript, as well as to assist with the refinement, readability, and debugging of statistical code written in `R` and `SAS`. All scientific content, data analyses, model specifications, interpretations, and conclusions were developed independently by the author. All AI-generated suggestions were critically reviewed, verified, and adapted where necessary. The author remains fully responsible for the content, analyses, and code presented in this thesis.

Code availability. All data processing, statistical analyses, model fitting, sensitivity analyses, and figure generation were performed using `SAS` and `R`. The complete analysis code, including exploratory analyses, model diagnostics for all outcomes, and MNAR sensitivity analyses, is available in a version-controlled Git repository: `https://github.com/Abdelrahman2799/Thesis_Project`.

Appendices

Appendix A: Exploratory distribution plots of physiological and salivary biomarkers

This appendix presents histograms and Q–Q plots for all physiological variables as well as salivary biomarkers (sAA and AACR) included in both the HRV/SCL and salivary cortisol datasets, shown before and after transformation. These plots were used to assess distributional properties and to evaluate the impact of the applied transformations. In the main text, RMSSD and cortisol are presented as representative examples; corresponding distributional plots for all remaining variables are provided in this appendix.

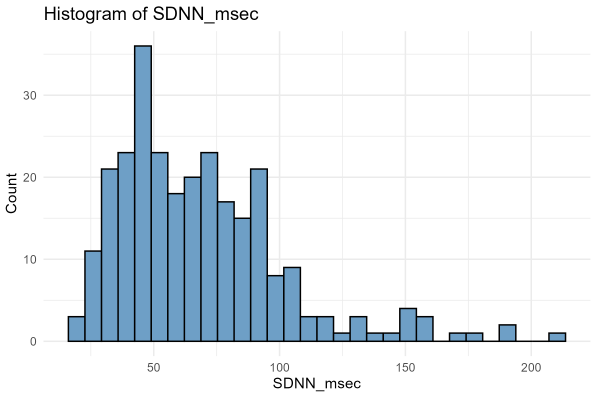

Figure 11: *
(a) Histogram of SDNN (raw scale)

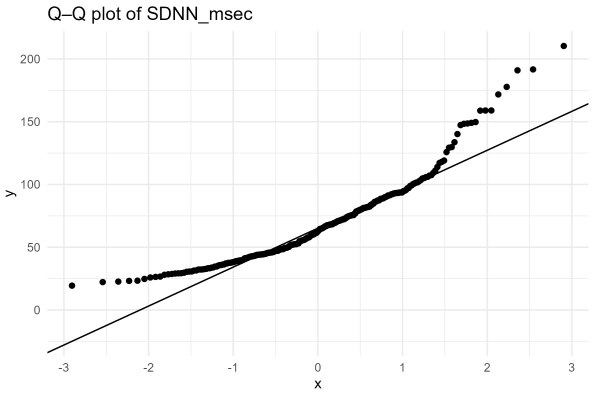

Figure 12: *
(b) Q–Q plot of SDNN (raw scale)

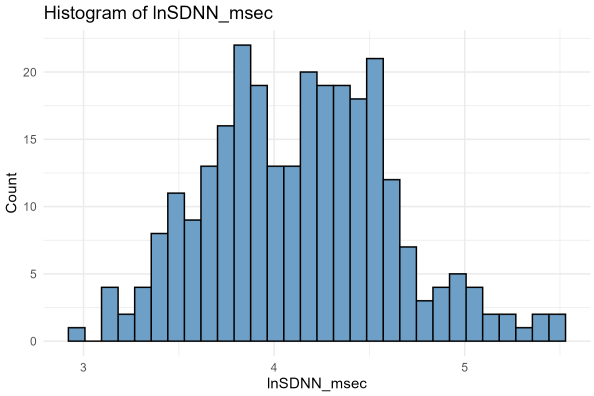

Figure 13: *
(c) Histogram of lnSDNN

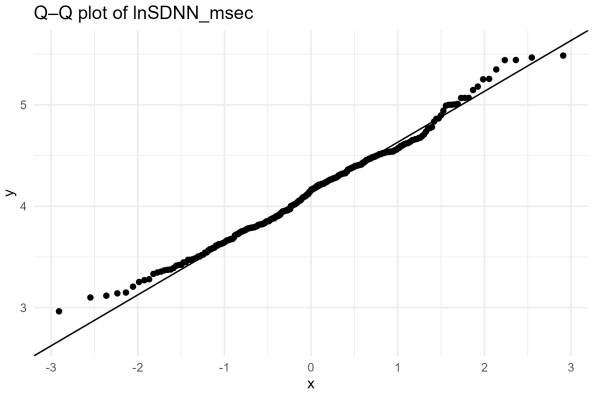

Figure 14: *
(d) Q–Q plot of lnSDNN

Figure 15: Distribution of SDNN before and after logarithmic transformation.

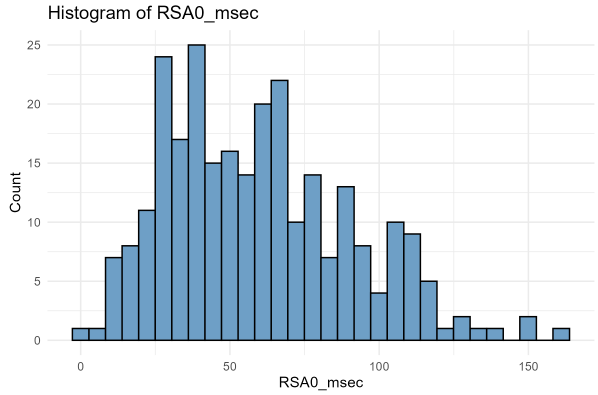

Figure 16: *
(a) Histogram of RSA0 (raw scale)

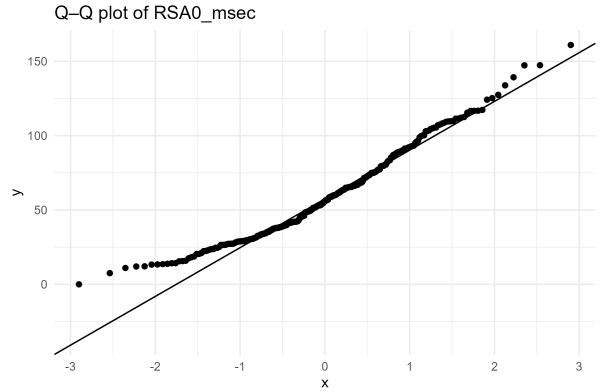

Figure 17: *
(b) Q–Q plot of RSA0 (raw scale)

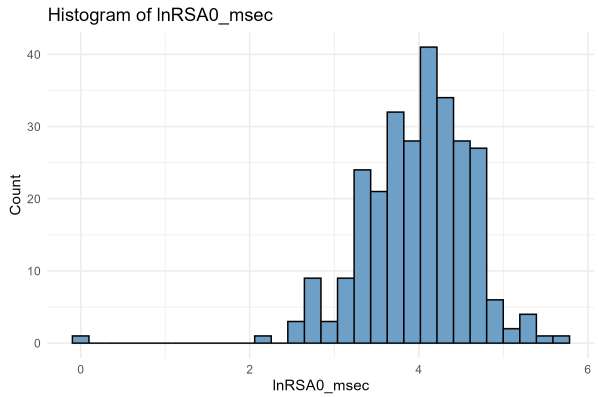

Figure 18: *
(c) Histogram of lnRSA0

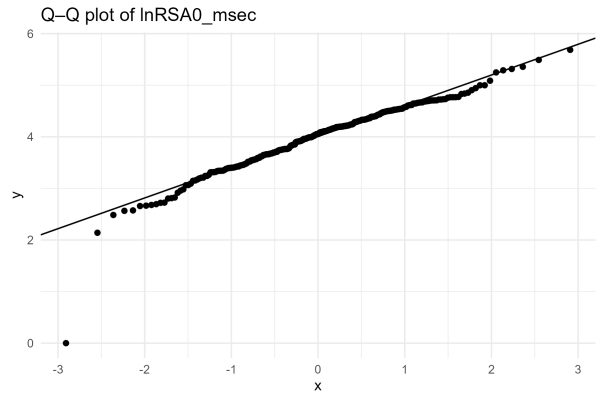

Figure 19: *
(d) Q–Q plot of lnRSA0

Figure 20: Distribution of RSA0 before and after logarithmic transformation.

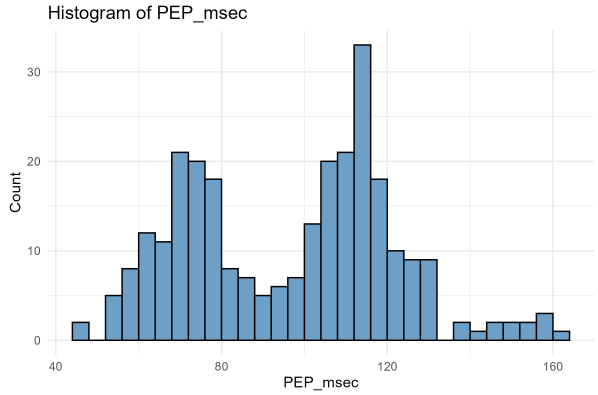

Figure 21: *
(a) Histogram of PEP (raw scale)

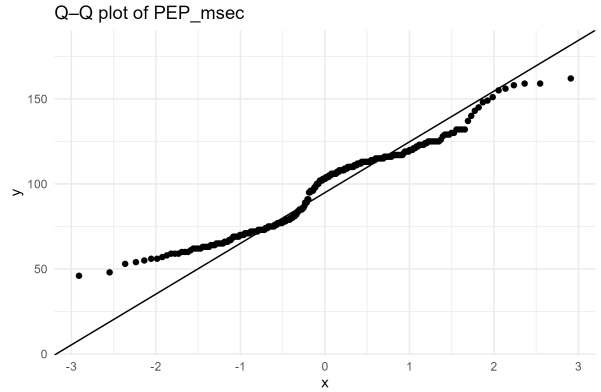

Figure 22: *
(b) Q–Q plot of PEP (raw scale)

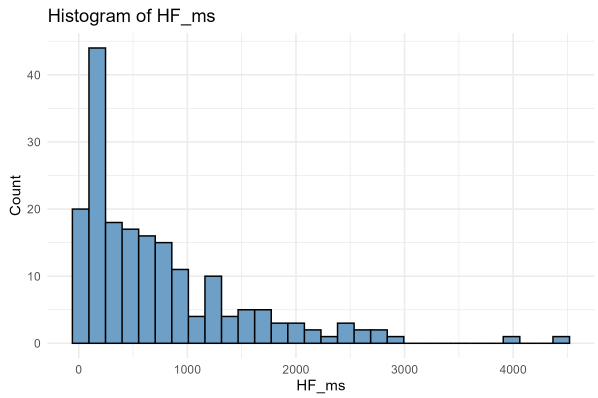

Figure 23: *
(a) Histogram of HF (raw scale)

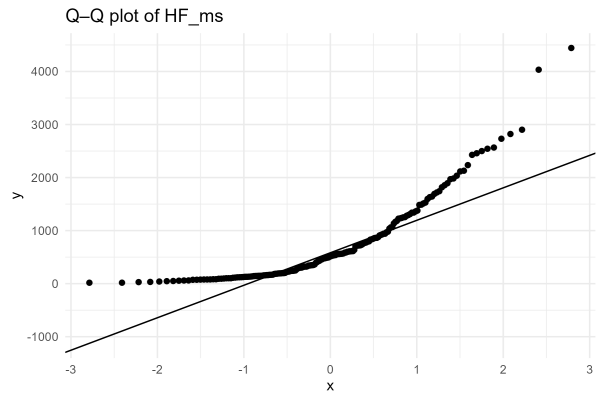

Figure 24: *
(b) Q–Q plot of HF (raw scale)

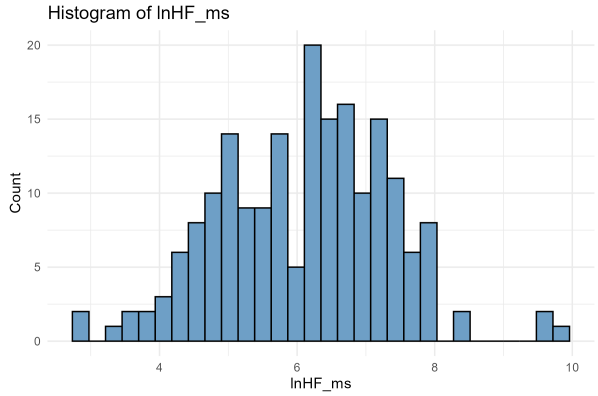

Figure 25: *
(c) Histogram of lnHF

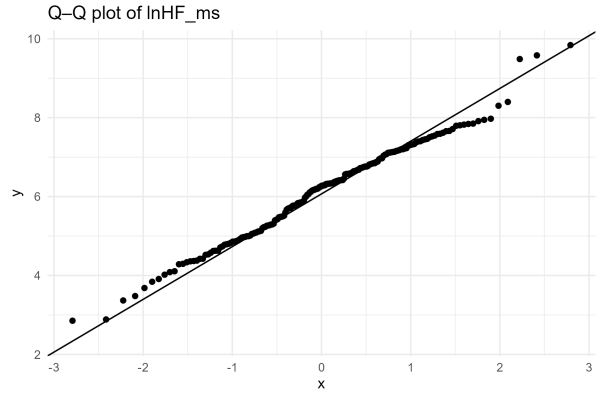

Figure 26: *
(d) Q–Q plot of lnHF

Figure 27: Distribution of HF before and after logarithmic transformation.

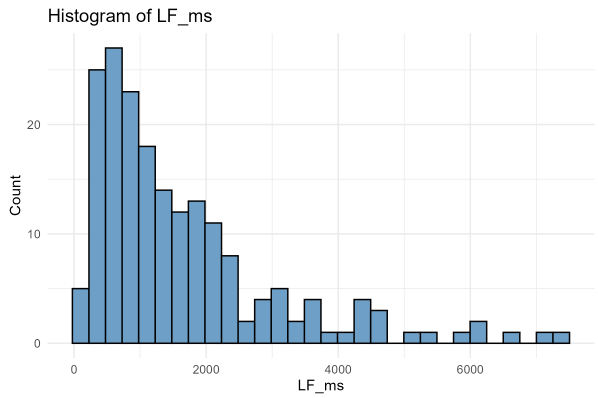

Figure 28: *
(a) Histogram of LF (raw scale)

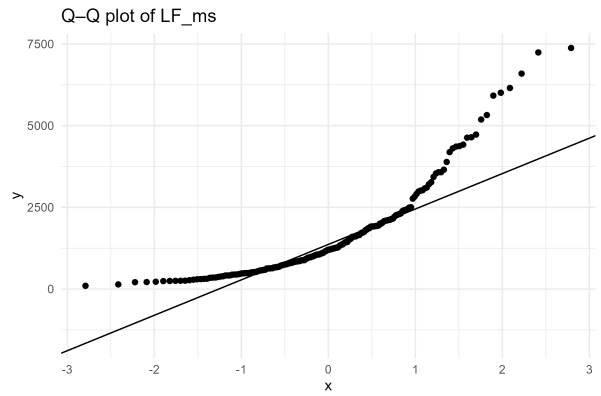

Figure 29: *
(b) Q–Q plot of LF (raw scale)

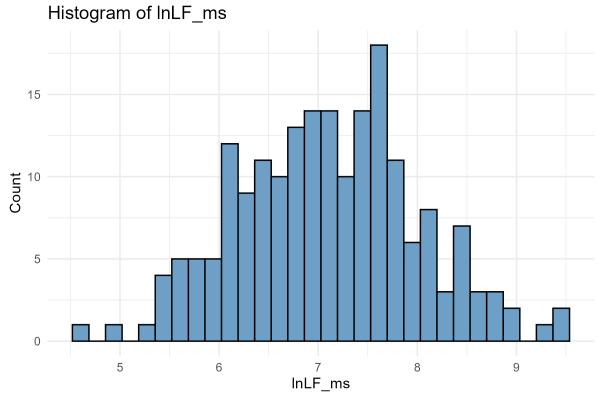

Figure 30: *
(c) Histogram of lnLF

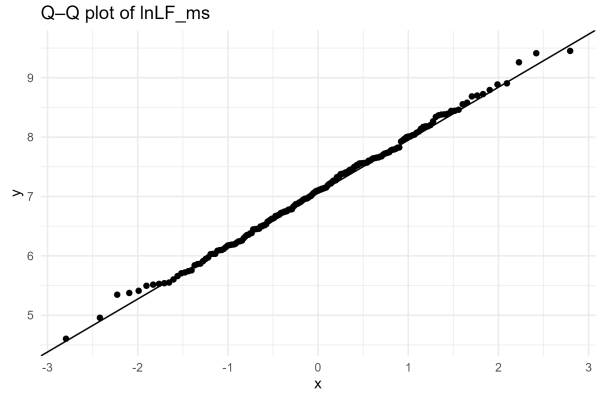

Figure 31: *
(d) Q–Q plot of lnLF

Figure 32: Distribution of LF before and after logarithmic transformation.

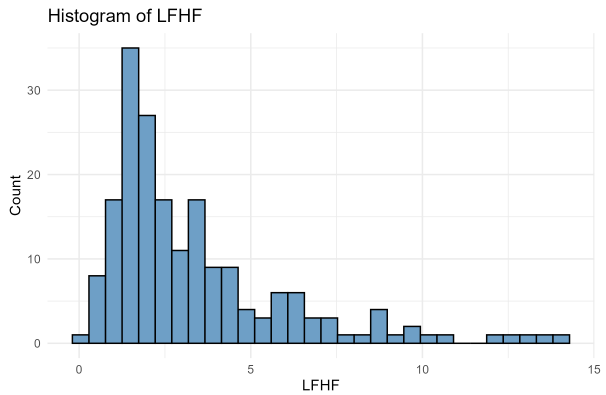

Figure 33: *
(a) Histogram of LFHF (raw scale)

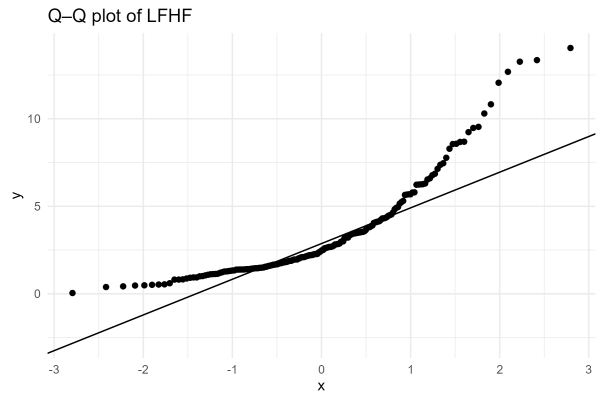

Figure 34: *
(b) Q–Q plot of LFHF (raw scale)

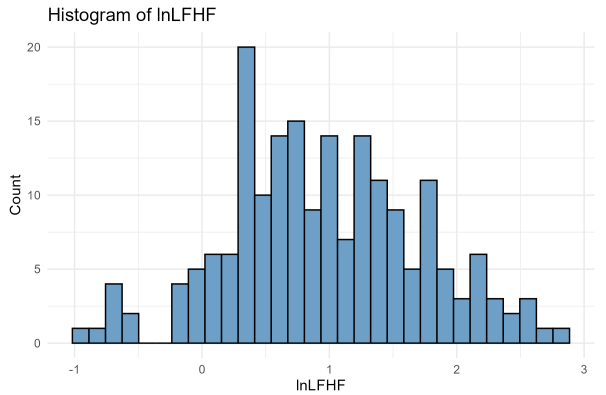

Figure 35: *
(c) Histogram of lnLFHF

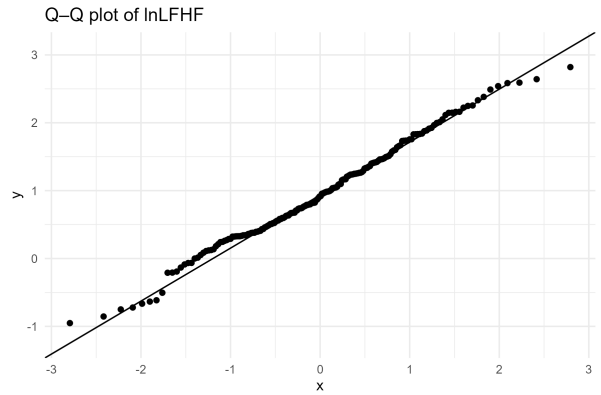

Figure 36: *
(d) Q–Q plot of lnLFHF

Figure 37: Distribution of LFHF before and after logarithmic transformation.

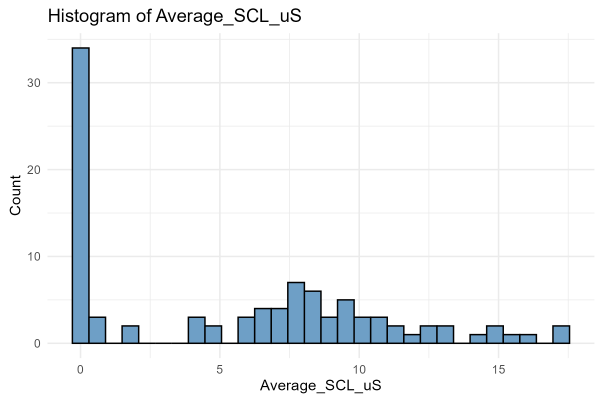

Figure 38: *
(a) Histogram of Average_SCL_uS (raw scale)

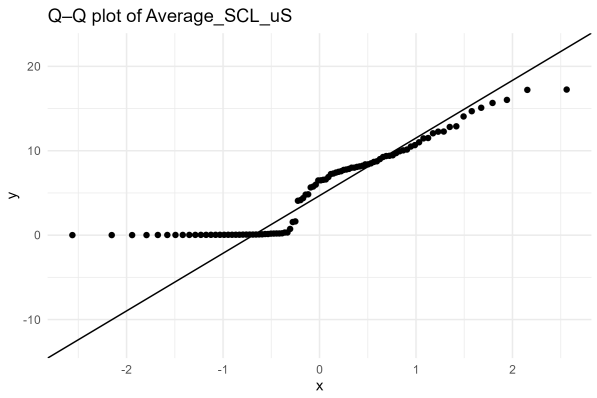

Figure 39: *
(b) Q–Q plot of Average_SCL_uS (raw scale)

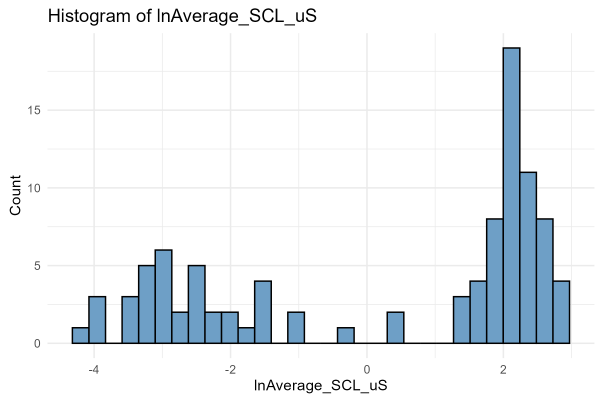

Figure 40: *
(c) Histogram of ln_Average_SCL_uS

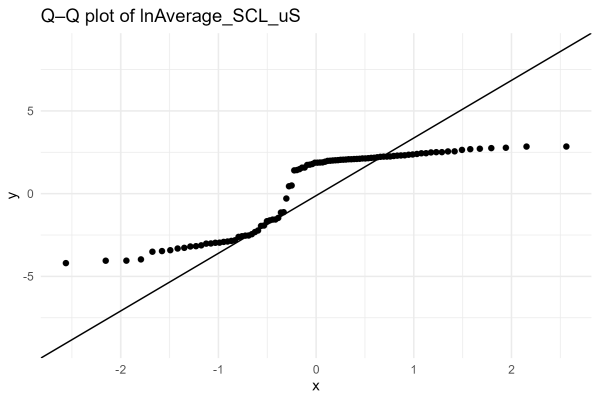

Figure 41: *
(d) Q–Q plot of ln_Average_SCL_uS

Figure 42: Distribution of Average_SCL_uS before and after logarithmic transformation.

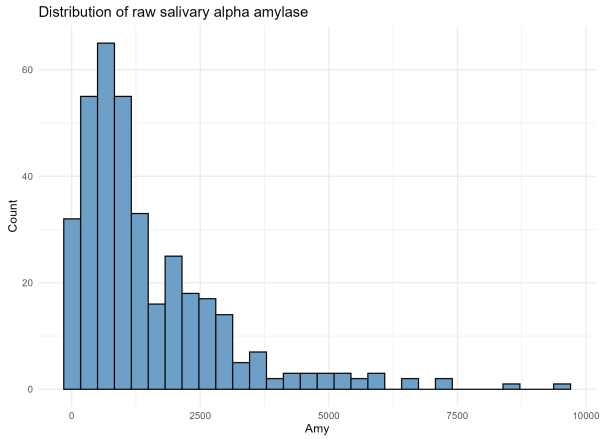

Figure 43: *
(a) Histogram of Salivary Alpha Amylase (raw scale)

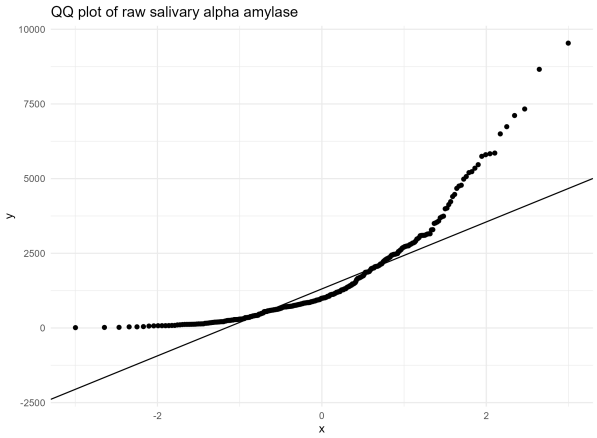

Figure 44: *
(b) Q–Q plot of Salivary Alpha Amylase (raw scale)

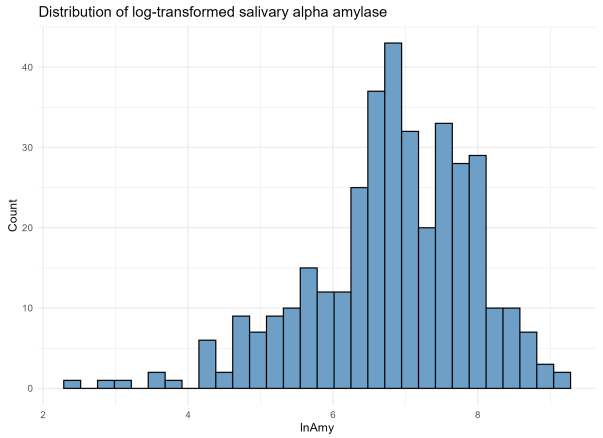

Figure 45: *
(c) Histogram of log-transformed Salivary Alpha Amylase

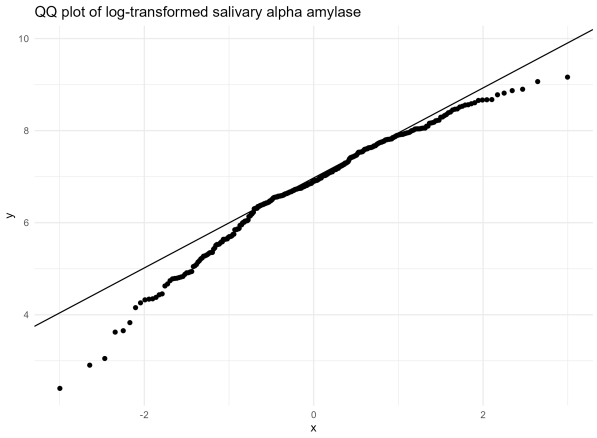

Figure 46: *
(d) Q–Q plot of log-transformed Salivary Alpha Amylase

Figure 47: Distribution of Salivary Alpha Amylase before and after logarithmic transformation.

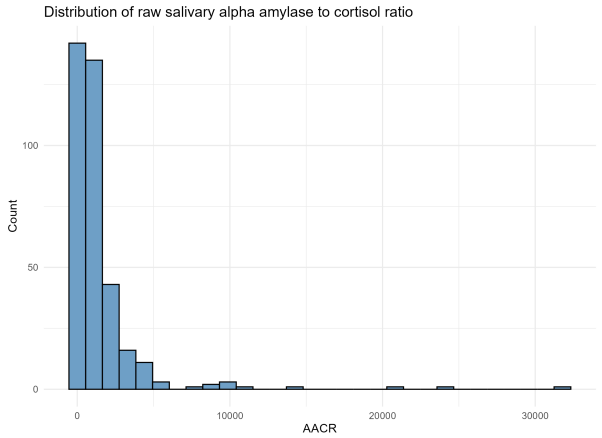

Figure 48: *
(a) Histogram of Salivary Alpha Amylase to Cortisol ratio(raw scale)

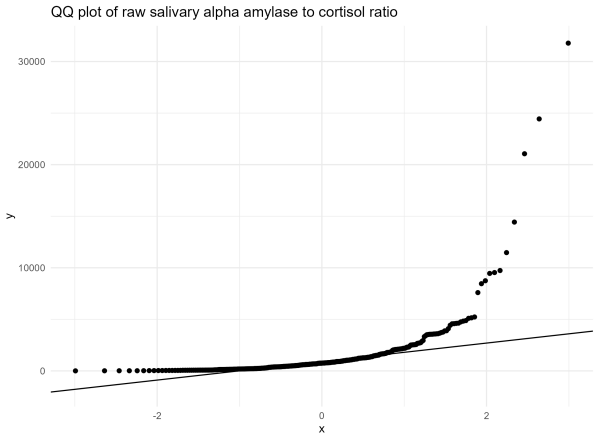

Figure 49: *
(b) Q–Q plot of Salivary Alpha Amylase to Cortisol ratio(raw scale)

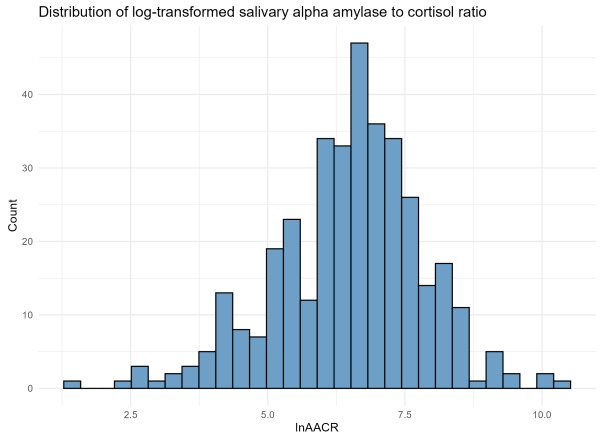

Figure 50: *
(c) Histogram of log-transformed Salivary Alpha Amylase to Cortisol ratio

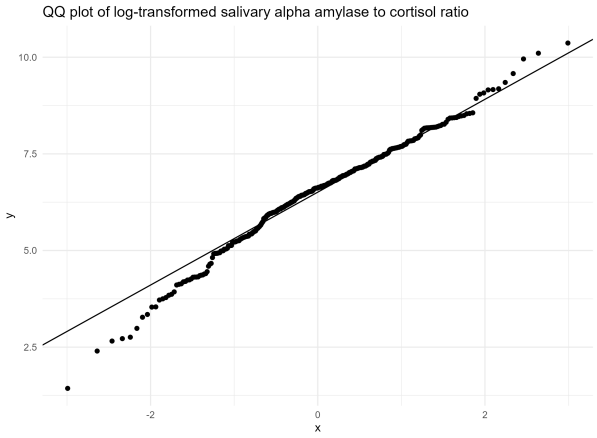

Figure 51: *
(d) Q–Q plot of log-transformed Salivary Alpha Amylase to Cortisol ratio

Figure 52: Distribution of Salivary Alpha Amylase to Cortisol ratio before and after logarithmic transformation.

Appendix B: Model diagnostics

For model diagnostics, linearity of the fixed effects was assessed using residuals versus fitted values, while residual normality was evaluated using Q–Q plots and residual histograms. These diagnostics are shown for two representative outcomes: lnRMSSD from the HRV analyses and lnCortisol from the salivary biomarker analyses. Other outcomes exhibited comparable diagnostic patterns and are therefore not shown for brevity.

For lnCortisol, which included time from CPT as a continuous predictor, additional influence diagnostics were examined. Participant-level Cook's distance and leverage measures were computed on the full dataset to assess potential influence associated with extreme predictor values. A sensitivity analysis excluding participants with extreme post-CPT sampling times was subsequently conducted, yielding comparable model estimates.

Figure 53: lnRMSSD and lnCortisol residual diagnostics plots

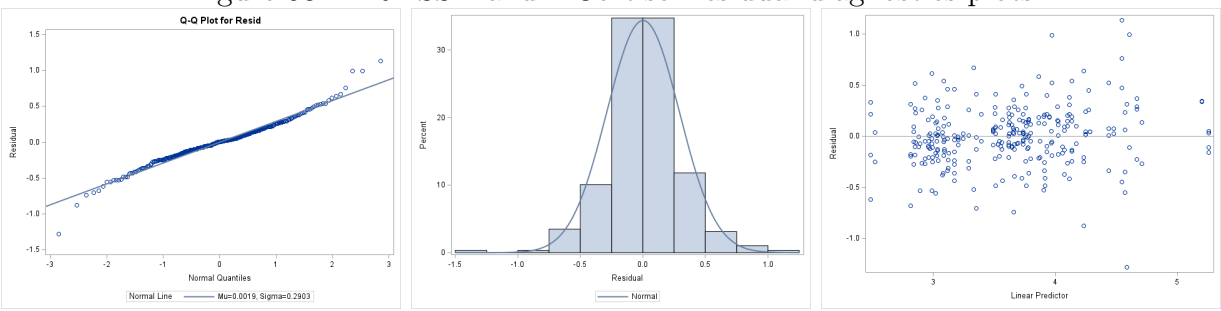

Figure 54: lnRMSSD model diagnostics

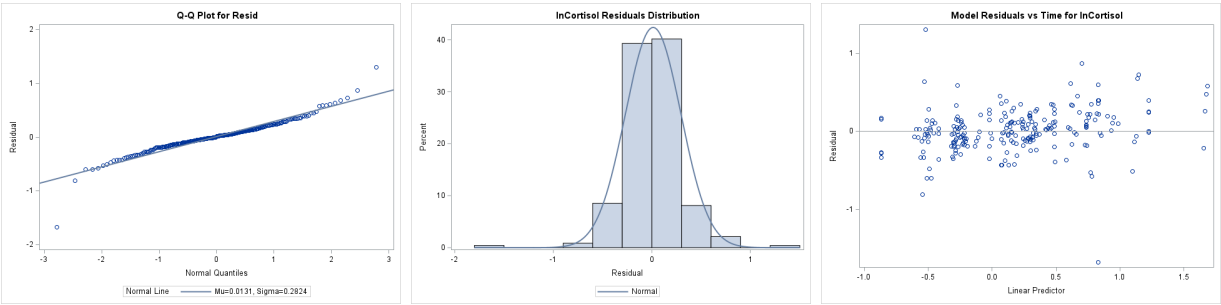

Figure 55: lnCortisol model diagnostics

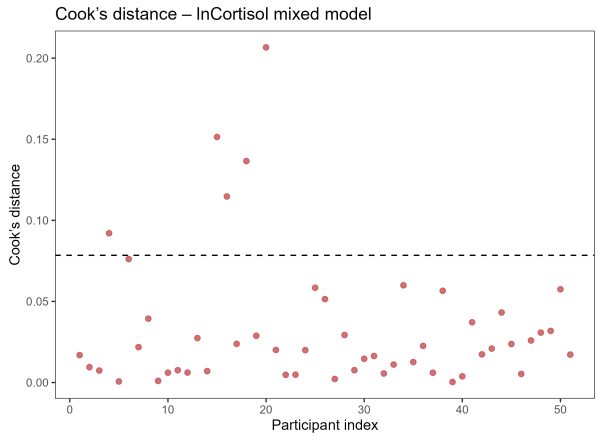

Figure 56: Participant-level Cook's distance for the lnCortisol model fitted on the full dataset. The dashed line indicates a heuristic threshold for potential influence.

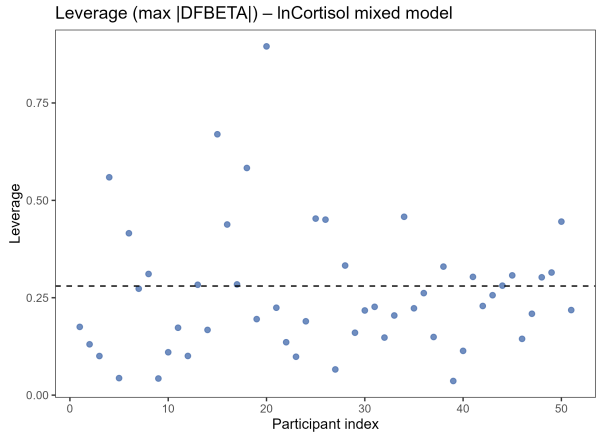

Figure 57: Participant-level leverage (maximum absolute DFBETA) for the lnCortisol model fitted on the full dataset. The dashed line denotes a heuristic reference threshold.

Appendix C: MNAR sensitivity analyses

Due to the size and complexity of the MNAR sensitivity analysis output tables for both the HRV/SCL and salivary biomarker datasets, detailed results are not reproduced here. Full MNAR model specifications, parameter estimates, and sensitivity tables are available in the accompanying Git repository.

References

Barker, D. J. (2004). The developmental origins of adult disease. *Journal of the American College of Nutrition*, 23(6 Suppl):588S–595S.

Bauer, A., Parsonage, M., Knapp, M., Iemmi, V., and Adelaja, B. (2016). Costs of perinatal mental health problems. *London School of Economics and Political Science*.

Braeken, M. A. K. A., Kemp, A. H., Outhred, T., Otte, R. A., and Van den Bergh, B. R. H. (2013). Pregnancy anxiety is associated with shorter gestational age and reduced heart rate variability in offspring. *Journal of Psychosomatic Research*, 75(6):556–562.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review*, 97(4):523–547.

Buss, C., Davis, E. P., Muftuler, L. T., Head, K., and Sandman, C. A. (2010). High pregnancy anxiety during mid-gestation is associated with decreased gray matter density in 6–9-year-old children. *Psychoneuroendocrinology*, 35:141–153.

de Geus, E. J. C., Willemsen, G. H. M., Klaver, C. H. A. M., and van Doornen, L. J. P. (1995). Ambulatory measurement of respiratory sinus arrhythmia and respiration rate. *Biological Psychology*, 41(3):205–227.

De Peuter, S., Victoir, A., Lemaigre, V., Van Diest, I., Verleden, G., Demedts, M., and Van den Bergh, O. (2006). Catastrophic thinking and symptom perception in asthma: validation of the catastrophizing about asthma scale. *The Internet Journal of Asthma, Allergy and Immunology*, 5(1). Catastrophizing about Asthma Scale (CAS) as a dyspnea-related catastrophizing measure.

Diggle, P. J., Heagerty, P., Liang, K.-Y., and Zeger, S. L. (2002). *Analysis of Longitudinal Data*. Oxford University Press, Oxford, 2 edition.

Kroenke, K., Spitzer, R. L., and Williams, J. B. W. (2002). The PHQ–15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosomatic Medicine*, 64(2):258–266.

Laborde, S., Mosley, E., and Thayer, J. F. (2017). Heart rate variability and cardiac vagal tone in psychophysiological research – recommendations for experiment planning, data analysis, and data reporting. *Frontiers in Psychology*, 8:213.

Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer Series in Statistics. Springer, New York.

Reitan, R. M. and Wolfson, D. (1985). *The Halstead–Reitan Neuropsychological Test Battery: Theory and Clinical Interpretation*. Neuropsychology Press, Tucson, AZ.

Schwabe, L., Haddad, L., and Schächinger, H. (2008). Hpa axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology*, 33(6):890–895.

Schwabe, L. and Schächinger, H. (2018). Ten years of research with the socially evaluated cold pressor test: Data from the past and guidelines for the future. *Psychoneuroendocrinology*, 92:155–161.

Shaffer, F. and Ginsberg, J. P. (2017). An overview of heart rate variability metrics and norms. *Frontiers in Public Health*, 5:258.

Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., Hergueta, T., Baker, R., and Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (m.i.n.i.): the development and validation of a structured diagnostic psychiatric interview for dsm-iv and icd-10. *The Journal of Clinical Psychiatry*, 59(Suppl 20):22–57.

Solomon, B. K., Wilson, K. G., Henderson, P. R., Poulin, P. A., Kowal, J., and McKim, D. A. (2015). A breathlessness catastrophizing scale for chronic obstructive pulmonary disease. *Journal of Psychosomatic Research*, 79(1):62–68.

Sullivan, M. J. L., Bishop, S. R., and Pivik, J. (1995). The pain catastrophizing scale: Development and validation. *Psychological Assessment*, 7(4):524–532.

Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology (1996). Heart rate variability: Standards of measurement, physiological interpretation, and clinical use. *Circulation*, 93(5):1043–1065.

Turk, E., van den Heuvel, M. I., Sleurs, C., Billiet, T., Uyttebroeck, A., Sunaert, S., Mennes, M., and Van den Bergh, B. R. (2023). Maternal anxiety during pregnancy is associated with weaker prefrontal functional connectivity in adult offspring. *Brain Imaging and Behavior*.

Van den Bergh, B. R. and Marcoen, A. (2004). High antenatal maternal anxiety is related to adhd symptoms, externalizing problems, and anxiety in 8- and 9-year-olds. *Child Development*, 75(4):1085–1097.

Van den Bergh, B. R., Sleurs, C., Geusens, B., Emsell, L., Sunaert, S., and Billiet, T. (2025). White matter microstructure and cognitive abilities in 28-year-old offspring prenatally exposed to maternal anxiety: A prospective exploratory multimodal brain imaging study. *Brain and Cognition*, 188:106319.

Van den Bergh, B. R., Van Calster, B., Smits, T., Van Huffel, S., and Lagae, L. (2008). Antenatal maternal anxiety is related to hpa-axis dysregulation and self-reported depressive symptoms in adolescence: A prospective study on the fetal origins of depressed mood. *Neuropsychopharmacology*, 33:536–545.

van den Heuvel, M. I., van Assen, M. A., Glover, V., Claes, S., and Van den Bergh, B. R. (2018). Associations between maternal psychological distress and salivary cortisol during pregnancy: A mixed-models approach. *Psychoneuroendocrinology*, 96:52–60.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer Series in Statistics. Springer, New York.

Vigo, D., Thornicroft, G., and Atun, R. (2016). Estimating the true global burden of mental illness. *The Lancet Psychiatry*, 3(2):171–178.

von Leupoldt, A., Mangelschots, E., Niederstrasser, N. G., Braeken, M., Billiet, T., and Van den Bergh, B. R. (2017). Prenatal stress exposure is associated with increased dyspnoea perception in adulthood. *European Respiratory Journal*, 50(1700642).

Wechsler, D. (2008). Wechsler adult intelligence scale–fourth edition (wais-iv). [Database record]. APA PsycTests.

Willemsen, G. H. M., de Geus, E. J. C., Klaver, C. H. A. M., van Doornen, L. J. P., and Carroll, D. (1996). Ambulatory monitoring of the impedance cardiogram. *Psychophysiology*, 33(2):184–193.