Accurate early identification of postpartum depression using demographic, clinical and digital phenotyping

Lisa Hahn [1,2], Simon B. Eickhoff [1,2], Ute Habel [3,7], Elmar Stickeler [4], Tamme W. Goecke [5], Susanne Stickel [3], Patricia Schnakenberg [3], Matthias Franz [6], Jürgen Dukart [1,2] *, & Natalia Chechko [1,3] *

[1]Institute of Neuroscience and Medicine, Brain & Behaviour (INM-7), Research Centre Jülich, Jülich, Germany; [2] Institute of Systems Neuroscience, Medical Faculty, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; [3] Department of Psychiatry, Psychotherapy and Psychosomatics, Medical Faculty, Uniklinik RWTH Aachen University, Aachen, Germany; [4] Department of Gynecology and Obstetrics, Medical Faculty, Uniklinik RWTH Aachen University, Aachen, Germany; [5] RoMed Hospital Rosenheim, Department of Obstetrics, Rosenheim, Germany; [6] Clinical Institute of Psychosomatic Medicine and Psychotherapy, Medical Faculty, Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany; [7] Institute of Neuroscience and Medicine, Brain Structure Function Relationship (INM-10), Research Centre Jülich, Jülich, Germany

* contributed equally

l.hahn@fz-juelich.de

Introduction

- **Postpartum depression** (PPD) is diagnosed in up to **13 % of women** after childbirth [1-2]
- Development of PPD depends on **many factors**, but its **definite cause** is **unknown.** Several known risk factors are associated with PPD, such as history of depression, postpartum blues or premenstrual syndrome [1, 4-9]
- In contrast to other psychiatric disorders, PPD is **more easily treatable with most effective prevention/ intervention** shortly after delivery in at-risk mothers [3, 5, 10-11]
- Most attempts for the **prediction** have either been **late** in the postpartum period (e.g. after 8-32 weeks) [15] **or** only reached a **low sensitivity** [16]
- There are **no accurate predictors** for PPD to such an extent that at-risk mothers can be identified and can benefit from early interventions

→ Here, we **evaluate** the **potential predictive power** of baseline **demographic, clinical and digital phenotyping** for early identification of PPD

Methods

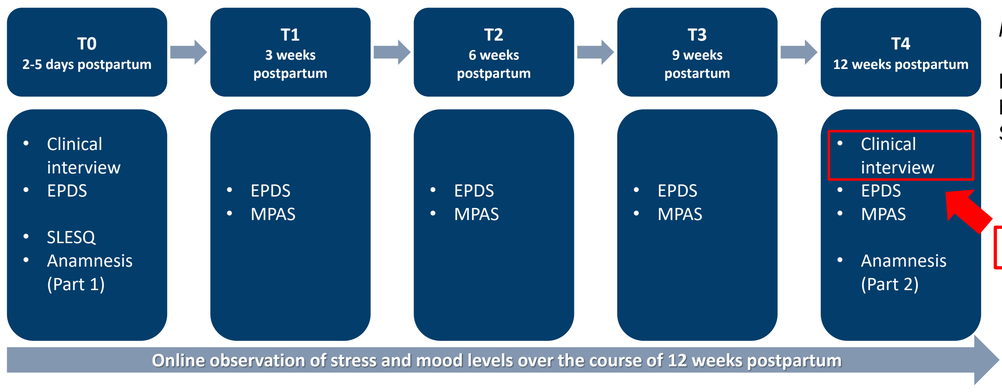


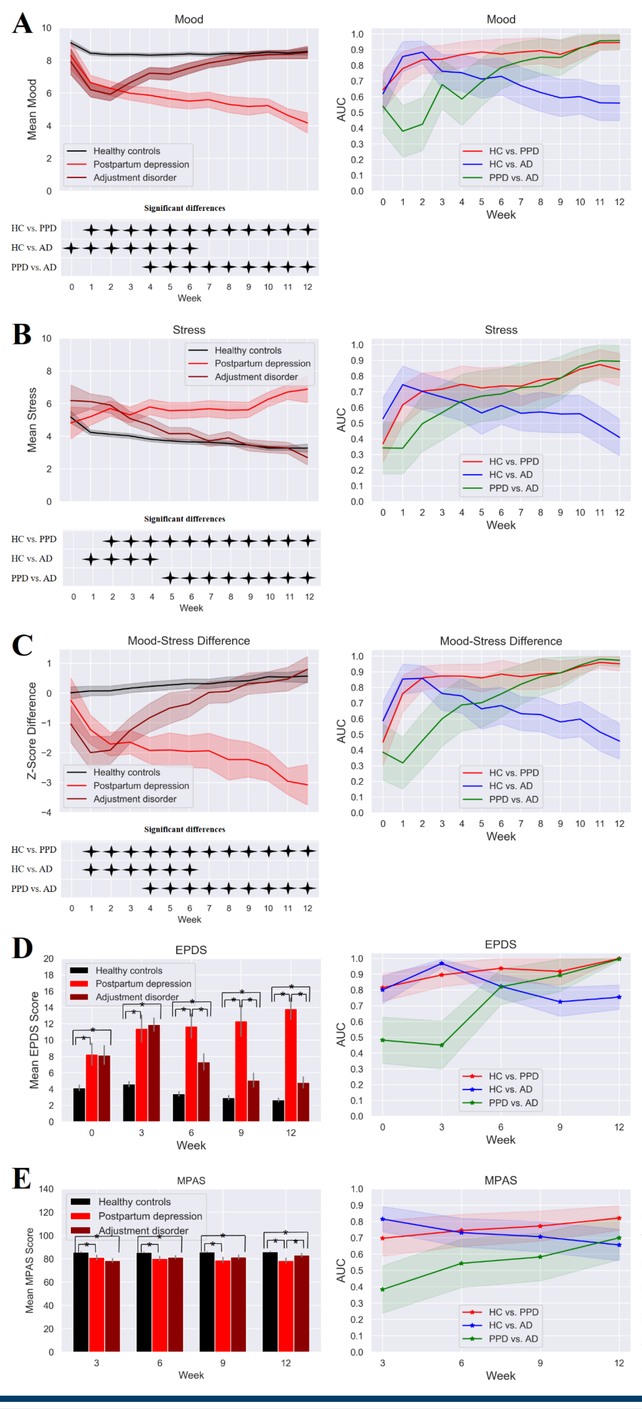
Figure 1. Study design.

EPDS: Edinburgh Postnatal Depression Scale
MPAS: Maternal Postnatal Attachment Scale
SLESQ: Stressful Life Events Screening Questionnaire

- **308 mothers** for the **training cohort** (mean age = 31.7 ± 4.76) and **193 mothers** (mean age = 32.7 ± 4.78) for the **validation cohort** recruited after giving birth at the University Hospital Aachen
- **Defined into three groups at week 12** (according to DSM-5 [12]): Women with PPD, Women with adjustment disorder (AD), and Healthy controls (HC)
- Measurements at **five different time points (T0 - T4)** separated by three-week intervals
- **Digital phenotyping: mood and stress levels** (i.e. scale from one to ten) were filled in online on a daily basis

- **Statistical analysis:**
  - Anamnestic data incl. SLESQ: Pearson X^2 test and logistic regression
  - Mood and stress levels, MPAS and EPDS scores: mixed ANOVA
- **Machine learning analysis:**
  - Logistic regression classifier with 1000 permutations of three-fold cross-validation for each group comparison for training
  - Evaluation based on balanced accuracy, sensitivity, specificity and area under the curve (AUC)
    → Application to validation cohort

Results



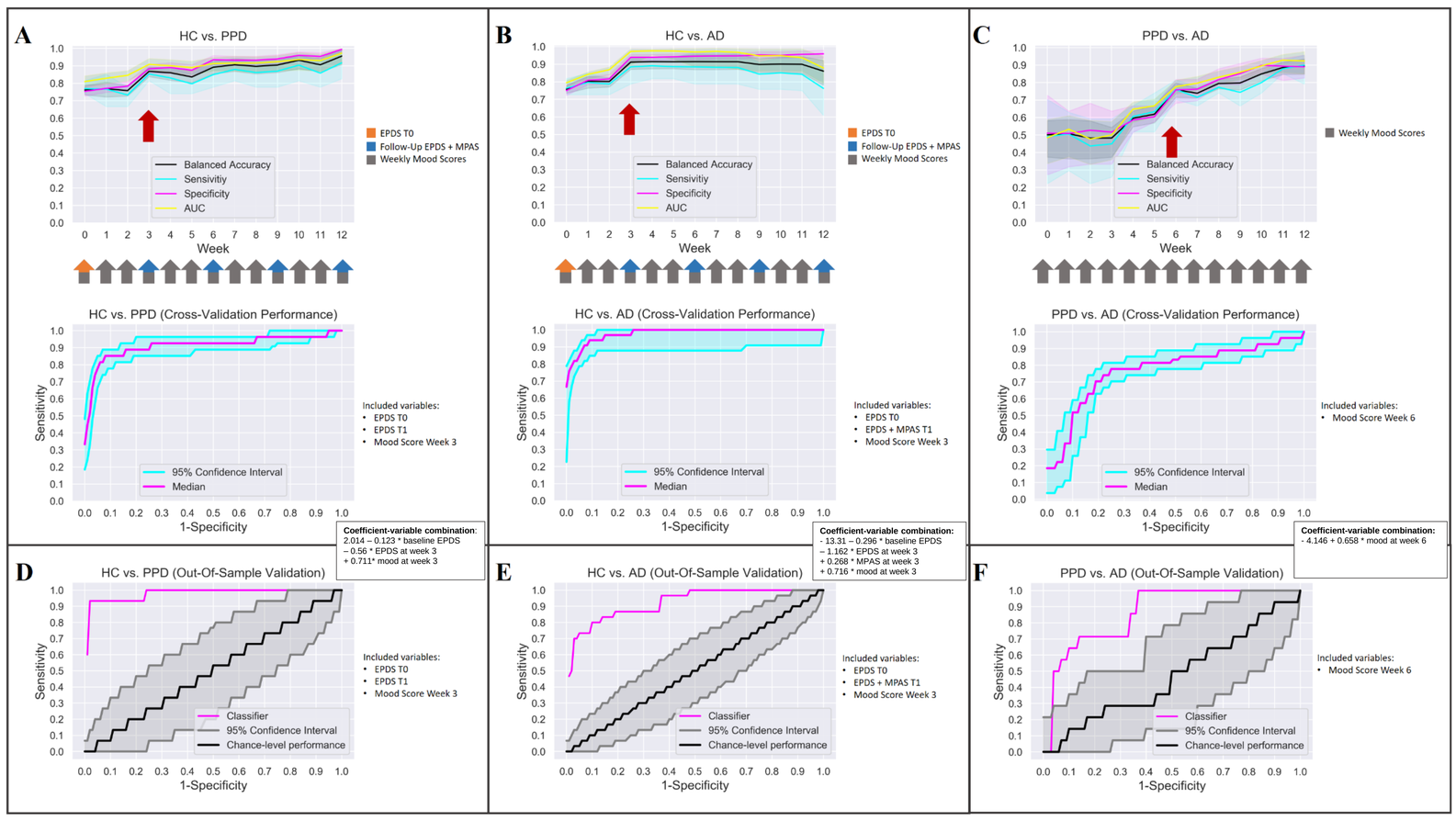


Figure 2. Results of machine learning analysis. Balanced accuracy, sensitivity, specificity and out-of-sample AUC for each group comparison are displayed for the exploration cohort (A-C). For HC vs. PPD (A), the values are displayed for EPDS at baseline and follow-up incl. mood scores. For HC vs. AD (B), the values are displayed for EPDS at baseline, EPDS and MPAS at follow-up incl. mood scores. For PPD vs. AD (C), the values are displayed for mood scores. (D-F) AUCs obtained for the replication cohort are displayed for the classifier selected based on the exploration results aside with chance-level performance.

Figure 1. Mood, stress, mood-stress difference, EPDS and MPAS scores. Weekly mood (A), stress (B) and mood-stress difference scores (C) incl. 95 % confidence intervals, results of the simple effects analyses and AUCs incl. 95 % confidence interval for each group comparison. *EPDS (D) and MPAS (E) mean scores and associated AUCs for each time point and group separately incl. their standard error and 95 % confidence interval. Statistically significant t-tests for group comparisons are marked with *.*

Table 1. Anamnestic data.

Anamnestic variable	HC	PPD	AD	Statistical test
Personal psychiatric history (no/yes)	221/27	16/12	19/14	$X^2(2, N= 309) = 34.7, p < .001$ *1,2
Familial psychiatric history (no/yes)	195/53	16/12	18/15	$X^2(2, N = 309) = 13.4, p = .001$ *1,2
Subjective birth-related psychological traumas	215/30	19/9	20/13	$X^2(2, N= 306) = 20.2, p < .001$ *1,2
Premenstrual syndrome (no PMS/mild PMS/PMS)	111/85/29	4/12/12	7/16/10	$X^2(4, N= 286) = 27.9, p < .001$ *1,2
Baby blues (no/yes)	152/93	8/20	7/26	$X^2(2, N= 306) = 28.0, p < .001$ *1,2
Stressful life events (no/yes)	145/103	11/17	12/20	$X^2(2, N= 308) = 7.92, p = .019$
Breastfeeding T4 (no/yes)	63/183	14/14	8/25	$X^2(2, N= 307) = 7.69, p = .021$ *1

* Bonferroni-corrected significant difference ($p < .05$) between [1] HC and PPD, [2] between HC and AD and/or [3] between PPD and AD

Discussion

- **Demographic and clinical risk factors** alone **did not differentiate** between women with **PPD and** women with **AD**
- Significant **risk factors for PPD** were largely **in accordance with** the **literature** [1, 4-9]
  - Breastfeeding (T4) as consequence and not as protective factor [13-14]
- **EPDS** and **MPAS** scores, **mood** and **stress** levels **displayed** a **distinctive pattern** for PPD and AD as compared to HC
  - EPDS was more sensitive than MPAS
  - Mood levels allowed for an accurate early differentiation of PPD and AD from HC
- **Accurate early differentiation for PPD vs. HC** was achieved by using baseline EPDS, and EPDS and mood scores **at week 3** with a **balanced accuracy of 0.87** in the **training** and **0.93** in the **validation cohort**
- **Accurate early differentiation for AD vs. HC** was achieved by using baseline EPDS, and EPDS, MPAS and mood scores **at week 3** with a **balanced accuracy of 0.91** in the **training** and **0.79** in the **validation cohort**
- **Accurate differentiation of PPD vs. AD** was possible at **week 6** with mood scores alone resulting in a balanced accuracy of **0.76** in the **training** and **0.73** in the **validation cohort**
- Combinations of mood, EPDS, and MPAS scores allowed for an **accurate identification of women at risk for PPD**

References: [**1**] O'Hara & Swain 1996, International Review of Psychiatry, 8:1, 37-54; [**2**] Goodman et al. 2008, Handbook of Clinical Psychology, 2, 937-965; [**3**] O'Hara 2009, Journal of Clinical Psychology, 65(12), 1258-1269; [**4**] Beck 2001, Nursing Research, 50(5), 275-285; [**5**] O'Hara & McCabe 2013, Annual Review of Clinical Psychology, 9, 379-407; [**6**] Robertson et al. 2004, General Hospital Psychiatry, 26(4), 289-295; [**7**] Buttner et al. 2013, Archives of Women's Mental Health, 16(3), 219-225; [**8**] Lee et al. 2015, Obstetrics & Gynecology Science, 58(5), 353-358; [**9**] Sylvén et al. 2013, Acta Obstetricia et Gynecologica Scandinavica, 92(2), 178-184; [**10**] Miller 2002, JAMA, 287(6), 762-765; [**11**] Stewart & Vigod 2016, New England Journal of Medicine, 375(22), 2177-2186; [**12**] American Psychological Association 2013, Diagnostic and statistical manual of psychiatric disorders (5th ed.); [**13**] Henderson et al. 2003, Birth, 30(3), 175-180; [**14**] Taveras et al. 2003, Pediatrics, 112(1), 108-115; [**15**] Camdeviren et al. 2007, Expert Systems with Applications, 32(4), 987-994; [**16**] Tortajada et al. 2009, Methods of Information in Medicine, 48(03), 291-298; [**17**] Jiménez-Serrano et al. 2015, Telemedicine and e-Health, 21(7), 567-574