Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 December 2019 doi:10.20944/preprints201912.0413.v1

Peer-reviewed version available at *Behavioral Sciences* **2020**, *10*; doi:10.3390/bs10020055

Article

Assessing Mothers' Post-Partum Depression from Their Infants' Cry Vocalizations

Giulio Gabrieli [1], Marc H. Bornstein[2,3], Nanmathi Manian[4] and Gianluca Esposito [1,5]*

[1] Psychology Program, Nanyang Technological University, Singapore, 639818, Singapore; giulio001@e.ntu.edu.sg
[2] National Institute of Child Health and Human Development, Bethesda, MD, 20892, USA; marc.h.bornstein@gmail.com
[3] Institute for Fiscal Studies, WC1E 7AE, London, United Kingdom;
[4] Westat, Rockville, Maryland, United States of America; nanmathinanian@westat.com
[5] Department of Psychology and Cognitive Science, University of Trento, 38068, Trento, Italy; gianluca.esposito@unitn.it
* Correspondence: gianluca.esposito@ntu.edu.sg

Abstract: Postpartum depression (PPD), a condition that affects up to the 15% of mothers in high-income countries, reduces attention toward the needs of the child and it is among the first causes of infanticide. PPD is usually identified using self-report measures and therefore the diagnosis may not always be valid. Previous studies highlighted the presence of significant differences in the acoustical properties of the vocalizations of children of depressed and healthy mothers. In this study, cry episodes of infants of depressed and non-depressed mothers are analyzed to investigate the possibility that a machine learning model can identify PPD in mothers from the acoustical properties of infants' vocalizations. Acoustic features (F_0, $F_{1\text{-}4}$, Intensity) are first extracted from recordings of crying infants, then novel cloud-based artificial intelligence models are employed to identify maternal depression versus non depression from estimated features. Trained model shows that commonly adopted acoustical features can be successfully used to individuate Post-Partum Depressed mothers with very high accuracy (89.5%).

Dataset License: CC-BY-NC

Keywords: Infant Cry; Post-Partum Depression; Acoustic Analysis

---

1. Introduction

Cry is an innate behavior and constitutes the first form of communication newborns use to interact with their caregivers [1]. Similar to speech in adults, cry vocalizations are produced by the vibration of the vocal folds, which are controlled by the Central Nervous System (CNS). Therefore, acoustical analysis of cry can identify pathological conditions associated with the vocal tract, the brain, and the spinal cord, as demonstrated in previous research[2,3]. The functional utility of infant cry is to elicit a response in an infant's caregiver, but, as proved by previous works, some situations and conditions diminish adults' sensitivity and responsiveness to cry [4–8]. Mothers who suffer from Postpartum Depression (PPD), a condition that is reported by 10-15% of mothers in high-income countries [9,10], and up to 50% in low- and middle-income countries reduces the level of stimulation produced by infant cry and decreases mothers' level of responsiveness toward the needs of their children [11–14]. Infants of depressed mothers are therefore exposed to an increased developmental risk [15].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 December 2019 doi:10.20944/preprints201912.0413.v1

Peer-reviewed version available at *Behavioral Sciences* **2020**, *10*; doi:10.3390/bs10020055

1.1. Post-Partum Depression Identification

Post-Partum Depression, a very common childbearing complication, is defined as a major depression condition that involves decreased interest or pleasure in activities, or sadness over an extended period of time [16]. Development of Post-Partum Depression is not only connected to previous episodes of depression, but it seems also to be more common when paired with other stressful events, or in women with a family history of mood disorder[9,17]. Rapid hormonal changes after delivery seem to play a primary role in the development of this disorder[18].
Currently, the presence of Post-Partum Depression in new mothers is assessed through questionnaires, for example, the Edinburgh Postnatal Depression Scale, a 10-item questionnaire that uses 4-point Likert scale responses[19,20] and the Beck Depression Inventory (BDI-II), a 21-item self-report questionnaire of the presence and related degree of depressive symptoms, consistent with the DSM-IV. An alternative approach is the Structured Clinical Interview per DMS-IV Axis I disorders (SCID-I).
Because identification is often based on self-reported measures or reports to interview questions, an estimated 60% of mothers with depressive symptoms receive no treatment or a clinical diagnosis [21]. The development of a tool to identify PPD in mothers in an objective way may improve diagnosis and thereby enhance the quality of life of children of depressed mothers.

1.2. Infant Cry

Infants' actively regulate acoustic information in their vocalizations to express specific needs. For example, acoustical analysis of cries has been used to identify the reason that induced a baby to cry, whether hunger, pain, or discomfort [22]. Similarly, babies vocalize differently according to their health status. Analysis of infants' cries has shown that specific patterns of cry vocalizations reflect infants' health status [23]. For example, Sheinkopf et al. [24], found different patterns of acoustical properties of cry vocalizations in children at risk for ASD compared to vocalizations from a healthy control group. Likewise, Garcia & Garcia [25] distinguished cry samples collected from deaf and hearing infants.
In a typical study, cry vocalizations are elicited in babies using a trigger (e.g., heel prick) and recorded on digital or analog sources [26]. Cry signals are then filtered to remove higher frequencies components. Finally, acoustic features are estimated from the signals. Commonly used acoustic features are the Fundamental Frequency (F0), which is the lowest pitch of the periodic signals, and its formants (F_1-F_4), which are frequency peaks which wavelength is a multiple of the fundamental frequency.
Different techniques are used to estimate acoustic features from cry samples, automatically (by means of a peak detection algorithm) or manually (by visual inspection of the spectrogram). Estimated features are then compared using statistical methods (to investigate the existence of specific patterns associated with a pathology) or fed to a classifier (to investigate whether those differences are strong enough to be used to identify a clinical situation reliably).

1.3. Aim and Hypothesis

Because of depressed mothers' reduced sensitivity and reactions to infants' cries, children may regulate the frequencies of their vocalization to maximize the responses of their caregivers. Previous studies have reported significant differences between the vocalizations of infants of depressed and non-depressed mothers [12]. Therefore, an analysis of the acoustical properties of cry vocalizations could identify in an objective way mothers who suffer from PPD. In this study, we investigated the possibility of using cry samples to identify PPD in mothers. More specifically, we hypothesized that a Cloud Computing based model will be able to identify children of mothers suffering from PPD, by using recordings of their cry vocalizations.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 31 December 2019 doi:10.20944/preprints201912.0413.v1

Peer-reviewed version available at Behavioral Sciences 2020, 10; doi:10.3390/bs10020055

2. Methods

In this work, acoustical features (F_0, $F_{1\text{-}4}$, Intensity) have been estimated from cry vocalizations collected in a previous, Then, a cloud-based AI model has been trained and tested. A visual representation of the procedure is reported in Figure 1

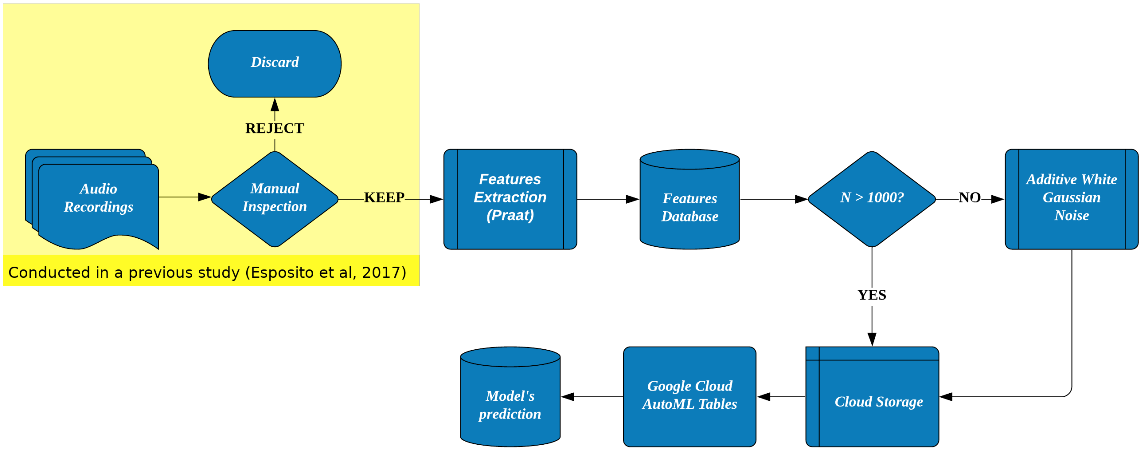

Figure 1. Summary of the steps employed in the development of the model for the diagnosis of PPD from infants' cry vocalizations.

2.1. Data

To test our hypothesis, we adopted a subset of a dataset used in a previous publication on the acoustical differences in cry vocalizations of children of depressed and healthy mothers [12]. Vocalizations from children of depressed (N = 29, 8 infant girls) and non-depressed mothers (N = 26, 7 infant girls) were collected at home when the infants were about 5 months of age (mean age = 157.4 days ± 8.5). 57 mothers (mean age = 31.1 years ± 3.9) were recruited from the Washington DC metropolitan area by mailing lists and newspaper advertisements; they included European Americans (n = 36), African-American (n = 10), Asian Americans (n = 7), American Indians (n = 1), and Latin Americans (n = 3). The study was approved by the IRB of the Eunice Kennedy Shriver National Institute of Child Health and Human Development (protocol code: 02-CH-0278) and was conducted according to the principles expressed in the Declaration of Helsinki. Written informed consent was obtained from all the participants prior to each recording session. Additional information on the demographic information of the participants are reported in the original work [12].
To increase the ecological validity of collected data, data were collected in the mothers' homes and mothers were asked to behave as they normally would, ignoring the presence of the experimenters. Infants and mothers were audio and video-recorded for at least 50 min, an amount of time that according to Holden and Millers [27] falls in the optimal time-frame for mother-infant observation. PPD was assessed using the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID-I) and the Beck Depression Inventory (BDI-II)[28], a 21-item survey which allows for a self-report of the presence and related degree of depressive symptoms, consistent with the DSM-IV. Mothers categorized as depressed had a high score on the BDI scale (>12) and had been diagnosed as having minor or major depression (SCID) by the time their infants were 5 months old. Collected cry samples (N = 715) were then digitalized in *WAVE* (*wav* file format, two channels) at 44.1kHz (16 bit). This format has been selected to preserve frequency information convoyed by the cry signals, as it is a lossless compression format [26]. Additional information on the data collection procedure, as well as the results of the statistical analysis, can be found in the original publication. [12].

2.2. Features extraction

Collected cry samples (N = 715) were then digitalized in *WAVE* (*wav* file format, two channels) at 44.1kHz (16 bit). This format has been selected to preserve frequency information convoyed by the cry signals, as it is a lossless compression format. Moreover, the sampling rate allows for analysis of frequencies up to 22kHz, which makes it suitable for a reliable analysis of up to the fourth formant. No further preprocessing was conducted on recorded signals to avoid alterations of frequency information conveyed within the signal.
Features ($F_{0\text{-}4}$) were extracted using Praat (v 6.0.50, Windows 64 bit), an open-source software design for voice analysis [29]. This software is based on the spectrographic analysis of a signal by means of a Long-Term Average Spectrum (LTAS). Specifically, the signal is first segmented into windows of a pre-specified length, then each segment is analyzed by means of an auto-correlation algorithm that works in the lag-domain (or $\tau - domain$).
Software's settings were adapted to correctly identify F_0 (Lower cutoff = 250Hz, upper cutoff = 800Hz) and the first four harmonics (Number of harmonics = 5, upper cutoff = 6000Hz) in a range that covers the spectrum in which cry vocalizations properties usually lie [30]. A copy of the script used for feature estimation is available online [31].
To investigate the possibility of using advanced Cloud Computing techniques to verify whether novel machine learning and neural networking techniques could be used to verify the presence of PPD in mothers, we relied on the Google Cloud Platform: *Google AutoML Tables*[1] [32]. A binary classification model was employed to discriminate between the cries of infants of mothers suffering from PPD from those of healthy infants. AutoML Tables were configured so that 80% of imported data was used for training, 10% for validation, and 10% for testing. The model was executed for up to two *node hours* (total running time of the training phase spread across the different machines that compose a node). Accuracy of the model was evaluated in terms of Precision (expressed in percentage), Area under the precision-recall curve (AUC PR, a value between 0 and 1, such that the higher the value, the higher the quality of the model), area under the curve of the receiver operative characteristics (AUC ROC, a value between 0 and 1, such that the higher the value, the higher the quality of the model), and logarithmic loss (a value between 0 and 1, such that the lower the value, the higher the quality of the model

Data Augmentation

AutoML Tables requires at least 1000 samples to executed (Beta version), therefore a data augmentation technique was applied to increase the number of samples of the dataset. Additive White Gaussian Noise ($\pm 1STD$) [33,34] was applied to a copy of the dataset and then merged with the original samples to obtain a dataset about twice the size of the original set of data (N = 1413). Augmented data, containing both acoustic (F_0, $F_{1\text{-}4}$, and Intensity) and demographic information (infants' gender, mothers' age) were employed for classification purposes. A copy of the final dataset is available online on the data repository of the Nanyang Technological University[31].

3. Results

Model's training stopped after 0.916 node hours, reporting an average accuracy on the test set of 89.5%, as well as robust values for AUC PR (0.954), AUC ROC (0.969), and Logarithmic Loss (0.250). Overall, the model achieved more than the 90% of precision (90.4%), with a true positive recall of 88.8% and an almost null false positive rate (0.09). Metrics of the score of the different evaluations are reported in Table 1.
For what concerns the model's error distribution, the confusion matrix of the model is reported in Table 2.

[1] https://cloud.google.com/automl-tables/

Metric	Score
AUC PR	0,954
AUC ROC	0,969
Logarithmic Loss	0,250
Accuracy	89,5%
Precision	90.4%
True positive rate (Recall)	88.8%
False positive rate	0.090

Table 1. Google's AutoML Model Evaluation Metrics.

	Predicted Label	
True Label	False	True
False	88%	12%
True	9%	91%

Table 2. Google's AutoML Model Confusion Matrix.

4. Discussion and Conclusions

In this work, we tested the possibility of using machine learning models to identify Post-Partum Depression in mothers from their infants' vocalizations.
Results of the model trained on Google's cloud computing service demonstrate the robustness of the method based on infants' cry analysis. The model, based on estimated acoustical properties of cry, identified at a high level of accuracy (89.5%) the children of depressed mothers. Our results suggest that machine learning models, trained in cloud environments, can support clinicians in the diagnosis of PPD.
Despite these promising results, some limitations need to be addressed. First, our model has been tested on a single dataset. Future studies should address the performance of models on data collected from different participants to verify the broader utility of the methods. Moreover, we trained the models using only acoustical features and demographic information about mothers (age) and infants (gender). Future studies might also address how including additional data, such as the questionnaires score (BDI) or the gestational age of the baby at birth, might improve predictive models by reducing the ratio of false positives and false negatives.
In conclusion, we show the possibility of using an objective measure of infants' cries subjected to machine learning models to advance beyond commonly used subjective reports to identify infants of Post-Partum depressed mothers.

Author Contributions: Conceptualization, Giulio Gabrieli and Gianluca Esposito; Data curation, Giulio Gabrieli, Marc H. Bornstein, Nanmathi Manian and Gianluca Esposito; Formal analysis, Giulio Gabrieli; Funding acquisition, Marc H. Bornstein and Gianluca Esposito; Methodology, Giulio Gabrieli and Gianluca Esposito; Software, Giulio Gabrieli; Supervision, Gianluca Esposito; Visualization, Giulio Gabrieli and Gianluca Esposito; Writing – original draft, Giulio Gabrieli; Writing – review & editing, Giulio Gabrieli, Marc H. Bornstein, Nanmathi Manian and Gianluca Esposito.

Funding: This research was supported by the Intramural Research Program of the NIH/NICHD, USA, and an International Research Fellowship at the Institute for Fiscal Studies (IFS), London, UK, funded by the European Research Council (ERC) under the Horizon 2020 research and innovation programme (grant agreement No 695300-HKADeC-ERC-2015-AdG).

Acknowledgments: In this section you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: The authors declare no conflict of interest. The founders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

 doi:10.20944/preprints201912.0413.v1

Peer-reviewed version available at *Behavioral Sciences* **2020**, *10*; doi:10.3390/bs10020055

Abbreviations

The following abbreviations are used in this manuscript:

PPD	Post-Partum Depression
CNS	Central Nervous System
SVC	Support Vector Machine
LTAS	Long-Term Average Spectrum
AUC PR	Area Under the Curve: Precision-Recall
AUC ROC	Area Under the Curve: Receiver Operative Characteristics

1. Esposito, G.; Venuti, P. Understanding early communication signals in autism: a study of the perception of infants' cry. *Journal of Intellectual Disability Research* **2010**, *54*, 216–223.
2. Mende, W.; Wermke, K.; Schindler, S.; Wilzopolski, K.; Hock, S. Variability of the cry melody and the melody spectrum as indicators for certain CNS disorders. *Early Child Development and Care* **1990**, *65*, 95–107.
3. Lester, B.M. Spectrum analysis of the cry sounds of well-nourished and malnourished infants. *Child Development* **1976**, pp. 237–241.
4. Bornstein, M.H. Children's parents. *Handbook of child psychology and developmental science* **2015**, pp. 1–78.
5. Sroufe, L.A.; Egeland, B.; Carlson, E.A.; Collins, W.A. *The development of the person: The Minnesota study of risk and adaptation from birth to adulthood*; Guilford Press, 2009.
6. Ainsworth, M.D.S.; Blehar, M.C.; Waters, E.; Wall, S.N. *Patterns of attachment: A psychological study of the strange situation*; Psychology Press, 2015.
7. Higley, E.; Dozier, M. Nighttime maternal responsiveness and infant attachment at one year. *Attachment & Human Development* **2009**, *11*, 347–363.
8. Esposito, G.; Nakazawa, J.; Venuti, P.; Bornstein, M. Perceptions of distress in young children with autism compared to typically developing children: A cultural comparison between Japan and Italy. *Research in Developmental Disabilities* **2012**, *33*, 1059–1067. doi:10.1016/j.ridd.2012.01.014.
9. O'hara, M.W.; Swain, A.M. Rates and risk of postpartum depression—a meta-analysis. *International review of psychiatry* **1996**, *8*, 37–54.
10. Paulson, J.F.; Bazemore, S.D. Prenatal and postpartum depression in fathers and its association with maternal depression: a meta-analysis. *Jama* **2010**, *303*, 1961–1969.
11. Donovan, W.L.; Leavitt, L.A.; Walsh, R.O. Conflict and depression predict maternal sensitivity to infant cries. *Infant Behavior and Development* **1998**, *21*, 505–517.
12. Esposito, G.; Manian, N.; Truzzi, A.; Bornstein, M.H. Response to infant cry in clinically depressed and non-depressed mothers. *PloS one* **2017**, *12*, e0169066.
13. Bornstein, M.H.; Arterberry, M.E.; Mash, C.; Manian, N. Discrimination of facial expression by 5-month-old infants of nondepressed and clinically depressed mothers. *Infant Behavior and Development* **2011**, *34*, 100–106.
14. Esposito, G.; Del Carmen Rostagno, M.; Venuti, P.; Haltigan, J.; Messinger, D. Brief report: Atypical expression of distress during the separation phase of the strange situation procedure in infant siblings at high risk for ASD. *Journal of Autism and Developmental Disorders* **2014**, *44*, 975–980. doi:10.1007/s10803-013-1940-6.
15. Murray, L.; Hipwell, A.; Hooper, R.; Stein, A.; Cooper, P. The cognitive development of 5-year-old children of postnatally depressed mothers. *Journal of Child Psychology and Psychiatry* **1996**, *37*, 927–935.
16. Wisner, K.L.; Parry, B.L.; Piontek, C.M. Postpartum depression. *New England Journal of Medicine* **2002**, *347*, 194–199.
17. Beck, C.T. Predictors of postpartum depression: an update. *Nursing research* **2001**, *50*, 275–285.
18. Bloch, M.; Schmidt, P.J.; Danaceau, M.; Murphy, J.; Nieman, L.; Rubinow, D.R. Effects of gonadal steroids in women with a history of postpartum depression. *American Journal of Psychiatry* **2000**, *157*, 924–930.
19. Cox, J.L.; Holden, J.M.; Sagovsky, R. Detection of postnatal depression: development of the 10-item Edinburgh Postnatal Depression Scale. *The British journal of psychiatry* **1987**, *150*, 782–786.
20. Cox, J. Origins and development of the 10item Edinburgh Postnatal Depression Scale. *Perinatal psychiatry* **1994**, pp. 115–124.

21. Ko, J.Y.; Rockhill, K.M.; Tong, V.T.; Morrow, B.; Farr, S.L. Trends in postpartum depressive symptoms—27 states, 2004, 2008, and 2012. *MMWR. Morbidity and mortality weekly report* **2017**, *66*, 153.
22. Tejaswini, S.; Sriraam, N.; Pradeep, G. Recognition of infant cries using wavelet derived mel frequency feature with SVM classification. 2016 International Conference on Circuits, Controls, Communications and Computing (I4C). IEEE, 2016, pp. 1–4.
23. Esposito, G.; Hiroi, N.; Scattoni, M.L. Cry, baby, cry: Expression of distress as a biomarker and modulator in autism spectrum disorder. *International Journal of Neuropsychopharmacology* **2017**, *20*, 498–503.
24. Sheinkopf, S.J.; Iverson, J.M.; Rinaldi, M.L.; Lester, B.M. Atypical cry acoustics in 6-month-old infants at risk for autism spectrum disorder. *Autism Research* **2012**, *5*, 331–339.
25. Garcia, J.O.; Garcia, C.R. Mel-frequency cepstrum coefficients extraction from infant cry for classification of normal and pathological cry with feed-forward neural networks. Proceedings of the International Joint Conference on Neural Networks, 2003. IEEE, 2003, Vol. 4, pp. 3140–3145.
26. Gabrieli, G.; Scapin, G.; Bornstein, M.H.; Esposito, G. Are Cry Studies Replicable? An Analysis of Participants, Procedures, and Methods Adopted and Reported in Studies of Infant Cries. *Acoustics* **2019**, *1*, 866–883.
27. Manian, N.; Bornstein, M.H. Dynamics of emotion regulation in infants of clinically depressed and nondepressed mothers. *Journal of Child Psychology and Psychiatry* **2009**, *50*, 1410–1418.
28. Beck, A.T.; Steer, R.A.; Brown, G.K. Beck depression inventory-II. *San Antonio* **1996**, *78*, 490–498.
29. Boersma, P.; Weenink, D. Praat: doing phonetics by computer. 2009. *Computer program available at< http://www. praat. org* **2005**.
30. Gabrieli, G.; Leck, W.Q.; Bizzego, A.; Esposito, G. Are Praat's default settings optimal for infant cry analysis? Linux Audio Conference 2019. LAC, 2019, pp. 83–88.
31. Gabrieli, G.; Esposito, G. Related Data for: Assessing mothers' post-partum depression from their infants' cry vocalizations, 2019. doi:10.21979/N9/IU0UOB.
32. Bisong, E. An Overview of Google Cloud Platform Services. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*; Springer, 2019; pp. 7–10.
33. Grover, P.; Sahai, A. Shannon meets Tesla: Wireless information and power transfer. 2010 IEEE international symposium on information theory. IEEE, 2010, pp. 2363–2367.
34. Hughes, B. On the error probability of signals in additive white Gaussian noise. *IEEE Transactions on Information Theory* **1991**, *37*, 151–155.

Sample Availability: The dataset generated for this publication is available on the Data Repository of the Nanyang Technological University https://doi.org/10.21979/N9/IU0UOB [31].