Using Naturally Occurring Data to Understand Post-Natal Depression

Miriam Koschate & Elahe Naserian

Post-natal depression (PND) affects between 10-15% of mothers and about 10% of fathers, and can have detrimental consequences for the family as a whole as well as the new-born child well into their adulthood. Prospective longitudinal data are difficult and expensive to collect leading to many studies being either cross-sectional or retrospective. PND is still highly stigmatised making self-report answers less reliable. The rising use of online forums, such as Netmums, Mumsnet and Reddit, by millions of parents and parents-to-be creates anonymous naturally occurring, longitudinal data in the form of online posts. Computational analysis of such posts may provide a new way of examining changes during pregnancy and after birth.

To assess how early parents-to-be develop a sense of being a parent, and how well they can access this new identity after birth, we developed a novel classification model that can detect from short written text which of two identities is cognitively active. With regular expressions and human labelling, we identified a set of 92 first-time pregnant women with PND symptoms and 557 first-time pregnant women without apparent PND symptoms where the date of birth could be identified. Posts during the nine months of pregnancy and up to three months after birth were analysed using probabilities from our classification model. Multilevel modelling shows that our two cohorts differ substantially in their development of a parent identity during pregnancy. Remarkably, mothers with PND symptoms eventually develop a parent identity around the time of birth but find it increasingly hard to access this identity cognitively over time. Significant differences emerge again at three months after birth, coinciding with the population median onset of clinically diagnosed PND. Our current research seeks to understand the reasons for differences in development and accessibility and evaluates whether our method can be used to detect early warning signs of PND during pregnancy to allow for better mental health care.