Big data study yields insights into pregnancy symptoms and how they progress over time

December 22 2023, by Michael Nissen

Total number of symptoms recorded each week, for selected data set selection criteria. Some data set selection criteria have been omitted for clarity. See Table 1 for details on the respective data selection criteria. The highest number of symptoms is recorded around gestation weeks 4–5. Credit: *npj Digital Medicine* (2023). DOI: 10.1038/s41746-023-00935-3

Fatigue, backache or insomnia—during pregnancy almost all women suffer from the strain of symptoms such as these. An interdisciplinary team of researchers from FAU has now investigated when such
complaints are particularly common and how they progress. The team used an anonymized big data dataset from a pregnancy app.

Each pregnancy is unique, but nearly all pregnant women suffer from similar pregnancy symptoms: They are tired, have backache, suffer from constipation, trouble sleeping or shortness of breath.

"We have known about these symptoms for a long time. However, little research has been conducted to date into when they arise during pregnancy and how they affect each other," explains Prof. Dr. Björn Eiskofier. "We must learn to understand the background to these symptoms better if we wish to improve antenatal care and treatment options."

The head of the Machine Learning and Data Analytics Lab at FAU is coordinating the interdisciplinary research project SMART Start together with Prof. Dr. Matthias W. Beckmann (head of the department and Chair of Obstetrics and Gynecology) and Prof. Dr. Peter A. Fasching (Professor of Translational Gynecology and Obstetrics) from the Department of Obstetrics and Gynecology at Universitätsklinikum Erlangen.

Prof. Dr. Oliver Schöffski from the Chair of Health Management at FAU and Prof. Dr. Matthias Braun from the Chair of Systematic Theology and Ethics at the University of Bonn are also involved in the project. Together, the researchers hope to encourage digitalization in antenatal care in Germany on the basis of a broad dataset.

**Tired in the first trimester**

As part of the interdisciplinary research project, Michael Nissen, research associate and doctoral candidate at the Machine Learning and Data Analytics Lab, analyzed a big data dataset from the German
pregnancy app developer keleya. Expectant mothers can select their own individual symptoms in the keleya app. They then receive information and content tailored to their individual needs.

"The most common symptom among pregnant women is fatigue. That was selected by 92.9% of users. This was followed by backache (92.6%), shortness of breath (81.0%), and trouble sleeping (79.4%)", summarizes Nissen. "It is interesting that each individual symptom occurs in a specific time frame," explains the computer scientist.

Fatigue reaches its peak in the first trimester of pregnancy, headaches start particularly in the 15th week of pregnancy, and diarrhea tends to affect women mainly at the beginning and the end of pregnancy, with a clear minimum around week 20. Trouble sleeping increases steadily as the pregnancy advances.

There may be a link between trouble sleeping and pregnancy complaints.

Some of the symptoms affect more than just the expectant mother's quality of life. They may also be linked to unwanted consequences for the pregnancy. For example, literature has proven that trouble sleeping is linked to a higher risk of a Caesarian section, premature birth, and depression in pregnancy. It is, therefore, important to conduct research into symptoms.

Large dataset available for research purposes

Keleya provided FAU with a large, anonymized dataset from users of the app for research purposes, therefore making a direct contribution to gaining a better understanding of the topic, a good example of successful collaboration between industry and research. A total of 183,732 women tracked their pregnancy-related symptoms using the app's symptom tracker.
They recorded more than 1.5 million symptoms. The researchers analyzed this huge dataset and compiled symptom progression curves with weekly symptom curves for 15 different pregnancy-related symptoms. "The size of the dataset is considerably greater than previous work carried out in this area."

Furthermore, the dataset is an accurate reflection of the real situation, based as it is on real-world evidence. This may help to reduce any possible distortions and discrimination in medical research and deliver a wide picture beyond the scope of traditional medical studies.

User habits can make it problematic for health apps to be used in scientific studies. Some users only try the app one single time. "We were able to demonstrate that these data barely differed from very active users," explains Nissen. As a result, the data from one-time users were able to be used for research purposes.

Overall, the study includes information on several as yet unknown or controversial symptoms and how they progress over time, and is considerably larger in scope than any other studies conducted to date. "Our study underlines the potential of the secondary use of sector data," emphasizes the doctoral candidate. "Collaboration between science and industry can lead to new scientific findings."

The findings are published in the journal *npj Digital Medicine*.
