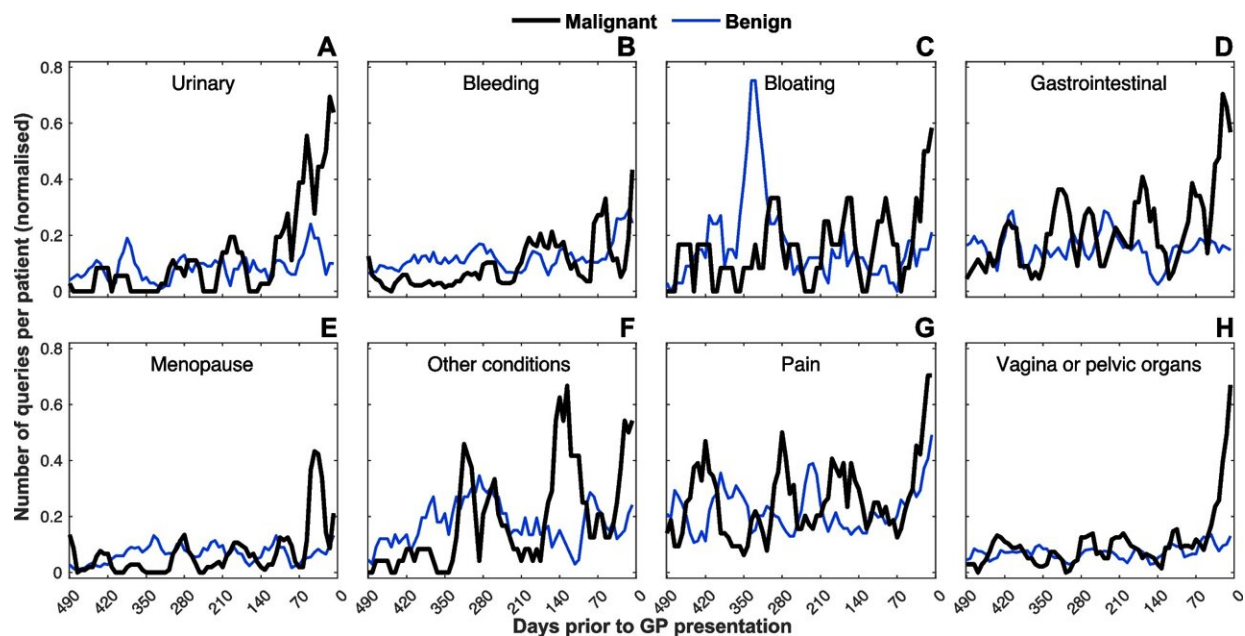# Online search data could be used to detect gynecological cancer cases earlier

March 11 2024, by Conrad Duncan

The time series chart outlining the number of online search queries per discrete category within the study cohort. The time series chart outlines the number of online search queries made per patient, within each distinct symptom category stratified by outcome (benign/malignant) up to 490 days in advance of GP referral. Online search activity can identify symptomatic individuals with gynecological cancer at an earlier stage. Credit: *BMC Public Health* (2024). DOI: 10.1186/s12889-024-17673-0

Search engine data could be used to detect gynecological cancer cases earlier, potentially months ahead of GP referrals, according to new research.

A study, led by Imperial researchers and colleagues from UCL's Department of Computer Science and Microsoft, has identified differences in online search data between people with a benign gynecological condition and those with a gynecological cancer, particularly those with ovarian cancer. The study has been published in *BMC Public Health*.

Ovarian cancer, the most lethal gynecological cancer, is the sixth most common cancer in the U.K., with about 7,400 people diagnosed each year and more than 4,000 deaths each year from the disease. There is currently no screening program in place for early disease detection.

Using Google data, the researchers noted a difference in search patterns as early as 360 days in advance of a GP referral and found that this provided a good prediction accuracy at about 60 days in advance of a GP referral.

Different symptom patterns were also noted between the benign and cancer groups. For example, a spike in urinary symptoms was noted 140 days in advance of GP referral, whereas bloating and pelvic pain symptoms appeared to present later (70 days in advance), compared to the benign group.

It is hoped that larger studies could confirm whether search engine data can be used as an early disease detection tool to help speed up diagnoses for gynecological cancer, particularly for ovarian cancer.

Dr. Srdjan Saso, senior author and chief investigator for the study, and gynecological cancer surgeon from Imperial's Department of Metabolism, Digestion and Reproduction, said, "Ovarian cancer is one of the most lethal cancers for women, with advanced stage carrying a poor prognosis, despite potential extensive surgery. The focus, therefore, remains on facilitating early disease detection. However, we do not have a screening program in place to enable this.

"Our research, in collaboration with computer scientists from UCL and Microsoft led by Professor Ingemar Cox and Dr. Elad Yom-Tov respectively, has shown that online search data may be used to differentiate between women with malignant and benign gynecological conditions. This suggests that it may be possible to build early detection tools using the internet which can identify women who may be at a higher risk of, in particular, ovarian cancers.

"We are hoping to raise funding in the near future to develop a multi-center study which can confirm these findings in a much larger cohort."

Enormous potential for disease screening

A previous study led by Imperial researchers showed that loyalty card data could be used to identify ovarian cancer cases earlier. However, this new study uses data from internet search engines which are used by a much larger proportion of the population (both in the U.K. and globally).

Dr. Jennifer F Barcroft, lead author for the study, from Imperial's Department of Metabolism, Digestion and Reproduction, said, "Our results show that it is possible to use search engine data to understand how conditions present, and that this may have use in early disease detection, given we have highlighted different symptom patterns between those with benign conditions and gynecological cancers.

"As nearly 98% of people in the U.K. have access to the internet, it is increasingly being used for health purposes. Online search data offers enormous potential within health and disease screening, given the widespread use of the internet worldwide. We hope that our research will drive interest in this novel area of research."

Professor Ingemar Cox, senior author for the study from UCL Computer Science, said, "Our current paper significantly expands on earlier work by incorporating a clinical study that confirms that an individual's risk of some diseases can be determined based on their Web search behavior. Web searches may therefore offer a simple, inexpensive method for disease screening. Of course, this does raise important ethical and privacy concerns, which need to be resolved."

Informed written consent was obtained from participants in the study to use online search data from Google Takeout files. This data was then anonymized.

More information: Jennifer F. Barcroft et al, Using online search activity for earlier detection of gynaecological malignancy, *BMC Public Health* (2024). DOI: 10.1186/s12889-024-17673-0

Provided by Imperial College London