Healthcare professionals’ queries on oseltamivir and influenza in Finland 2011-2016—Can we detect influenza epidemics with specific online searches?

Samuli Pesälä1,2 | Mikko J. Virtanen2 | Milla Mukka1 | Kimi Ylilammi3 | Pekka Mustonen4 | Minna Kaila5 | Otto Helve2,6

Background: Healthcare professionals (HCPs) search medical information during their clinical work using Internet sources. In Finland, Physician’s Databases (PD) serve as an Internet medical portal aimed at HCPs. Influenza epidemics appear seasonal outbreaks causing public health concern. Oseltamivir can be used to treat influenza. Little is known about HCPs’ queries on oseltamivir and influenza from dedicated online medical portals and whether queries could be used as an additional source of information for disease surveillance when detecting influenza epidemics.

Methods: We compared HCPs’ queries on oseltamivir and influenza from PD to influenza diagnoses from the primary healthcare register in Finland 2011-2016. The Moving Epidemic Method (MEM) calculated the starts of influenza epidemics. Laboratory reports of influenza A and influenza B were assessed. Paired differences compared queries, diagnoses, and laboratory reports by using starting weeks. Kendall’s correlation test assessed the season-to-season similarity.

Results: We found that PD and the primary healthcare register showed visually similar patterns annually. Paired differences in the mean showed that influenza epidemics based on queries on oseltamivir started earlier than epidemics based on diagnoses by −0.80 weeks (95% CI: −1.0, 0.0) with high correlation ($\tau = 0.943$). Queries on influenza preceded queries on oseltamivir by −0.80 weeks (95% CI: −1.2, 0.0) and diagnoses by −1.60 weeks (95% CI: −1.8, −1.0).

Conclusions: HCPs’ queries on oseltamivir and influenza from Internet medical databases correlated with register diagnoses of influenza. Therefore, they should be considered as a supplementary source of information for disease surveillance when detecting influenza epidemics.

KEYWORDS
health personnel, influenza, information systems, oseltamivir, primary health care, public health surveillance
Accessing the Internet enables users, including healthcare professionals (HCPs), to search medical information on the web. However, general search engines, such as Google, can lead to websites’ distributing medical information of varying quality to the general public, thus making it unreliable for use in clinical work. Online professional medical databases, such as MEDLINE or PubMed, provide information not only for scientific purposes but also for clinical decision making and continuing professional development. These databases are used to find information on current diseases and medications, and the information is used by HCPs in patient care.

Influenza appears worldwide as a possibly severe infectious disease with a major public health concern. Influenza outbreaks follow a temporal pattern and typically occur during the cold seasons of the year. People are infected mainly by two types of influenza viruses, A and B, spread via air or contaminated surfaces. Influenza epidemics may cause work absence, severe complications, hospitalizations, and even deaths, thus having a large economic burden on the community. For the prevention of influenza, immunization and good personal health and hygiene habits are suggested to control infection. To treat influenza, antiviral medications can be used. Oseltamivir, an antiviral agent, is used to treat seasonal or pandemic influenza in adults and children. Oseltamivir is a neuraminidase inhibitor, which prevents the reproduction of the influenza virus. It is available both as a tablet and in a liquid form and is recommended for people with a high risk of complications and should be taken within 48 hours of the first symptoms of influenza. The World Health Organization (WHO) has classified oseltamivir as complementary on their list of essential medicines in a health system.

There are several methods for the detection of influenza epidemics. The number of influenza diagnoses has traditionally been collected during medical visits to primary healthcare professionals and also from positive findings at microbiological laboratories. The accumulation of cases in a defined time period in a population determines the start of an influenza epidemic and indicates the intensity of an outbreak. Along with epidemiological and virological data from clinical encounters and microbiological laboratories, influenza search trends from Internet search engines have also been studied to detect epidemics. Certain Google queries coincided highly with medical visits related to influenza-like symptoms, thus making influenza activity estimation geographically possible. However, this surveillance method contained several flaws in timing and location of an influenza outbreak by overestimating the intensity of an epidemic and missing the first wave of an influenza pandemic. In addition, searches of the general public may be affected by issues not related to an actual epidemic, such as publicity on a given disease while searches by HCPs may be less affected. As general search engines are used by both the general public and HCPs, these search data can yield heterogeneous results. Our study on HCPs’ information seeking behavior from Internet medical databases showed that searches and diagnoses on Lyme borreliosis associated with each other. Therefore, we concluded that Internet searches could be used as an additional source of information for disease surveillance.

The Moving Epidemic Method (MEM) has been developed to assess the timing of influenza epidemics and to estimate their baseline and threshold. It is implemented by WHO and the European Centres for Disease Prevention and Control (ECDC) to monitor influenza circulation in European countries. Historical data on influenza weekly rates are analyzed with MEM, and the method includes three stages. In the first stage, the length of each influenza season with start and end points is determined forming a pre-epidemic, an epidemic, and a post-epidemic period. In the second stage, the epidemic baseline and thresholds are calculated using pre-epidemic and post-epidemic values from historical seasons. In the third stage, low, medium, and high intensity thresholds are computed. Although early detection of influenza to detect outbreaks using epidemiological and virological data has been studied before, including general search engine queries, little is known about queries on influenza and oseltamivir on Internet databases by HCPs.

When searching for medical information on influenza and oseltamivir, HCPs access dedicated medical databases on the Internet. Duodecim Medical Publications Ltd (owned by the Finnish Medical Society Duodecim) produces and maintains an Internet-based portal called Physician’s Databases (PD). It is available throughout the Finnish healthcare system, and its users can be tracked in primary health care in Finland by an Internet Protocol address included in a log file. PD includes point-of-care evidence-based medicine articles aimed at HCPs, mostly comprising physicians, pharmacists, and nurses. In 2016, there were nearly 21 000 working-age physicians in Finland. The medical articles in the databases are in Finnish and are opened over 15 million times a year. The National Institute for Health and Welfare (NIHW) is the research and development institute in Finland maintaining the register of public primary healthcare diagnoses and the National Infectious Diseases Register (NIDR). The data from both NIHW registers can be used in research, decision making, and planning of healthcare services.

When searching for medical information on influenza in clinical work, primary care physicians in Finland may access the medical Internet portal, PD, where a medical article on influenza aimed at HCPs can be found. During or immediately following a patient visit, a primary care physician reports an influenza diagnosis in the electronic patient record, where it will then be automatically transferred to the national register of public primary healthcare diagnoses. NIDR includes laboratory reports of influenza A and influenza B notified electronically by microbiological laboratories. If the prescription of an antiviral medication for influenza is needed, PD’s pharmaceutical database is available for a physician to search for information on oseltamivir. Queries on oseltamivir (openings of the page with information on oseltamivir) and influenza (openings of the page with information on influenza) during a physician’s encounter can be tracked in the log files of PD. The aim of our study was to assess whether HCPs’ queries on oseltamivir and
FIGURE 1  Influenza epidemic thresholds (pre-epidemic, post-epidemic) across Finland during 2011-2016 by season for (A) weekly queries on oseltamivir and (B) weekly influenza diagnoses. The start and end of an epidemic period placed in the patterns of queries and diagnoses corresponds to the underlined epidemic week.
influenza from Internet medical databases could be used as an additional source of information for disease surveillance when detecting influenza epidemics. We hypothesized that queries on oseltamivir would share similar seasonal trends seen in epidemiological data on influenza and that they could be measured with MEM.

2 | MATERIAL AND METHODS

We carried out a register-based study by collecting weekly log data on the number of queries on oseltamivir and influenza from PD, in order to compare logs to influenza diagnoses (J09-11 according to the International Statistical Classification of Diseases and Related Health Problems, 10th Revision [ICD-10] disease classification code system9 and R80 in the International Classification of Primary Care, Second Edition [ICPC2] coding system23) and laboratory reports of influenza A and influenza B found from NIDR. Queries on oseltamivir included log data on oral capsules (30 mg, 45 mg, and 75 mg) and a powder for oral suspension (6 mg/mL) of oseltamivir. The data were collected across Finland during 2011-2016 comprising five seasons of influenza (2011/12, 2012/13, 2013/14, 2014/15, and 2015/16) with five indicators (queries on oseltamivir, influenza diagnoses, laboratory reports of influenza A and influenza B, and queries on influenza).

We used the MEM model to calculate the starts and ends of an epidemic period and influenza thresholds (pre-epidemic, post-epidemic) [R language, 2.12 version24]. We analyzed the starting weeks of the epidemic periods consisting of queries on oseltamivir, influenza diagnoses, queries on influenza, and laboratory reports of influenza A and B pairwise comprising a total of ten pairs. The starting weeks correspond to week numbers for a calendar year starting from the beginning of January (week 1). To assess if each indicator reaches the epidemic threshold at similar times, paired differences in the starting weeks were calculated. Due to a small number of observations (starting weeks), the bootstrapping method25,26 was used to estimate the distribution of observations. We bootstrapped paired differences comprising five observations with 1,000 replications resulting in bootstrapped mean, bias-corrected and accelerated (BCa) (adjusted for ties) 95% confidence interval (CI) of the mean, and p-value of the mean. Kendall's rank correlation coefficient (τ) was used to assess the statistical season-to-season similarity between a pair (Kendall's tau formula in Appendix 1 files). Paired differences and Kendall's correlation tested only the starting weeks comprising five seasons with five indicators.

3 | RESULTS

Visually similar patterns were found between annual queries on oseltamivir and influenza diagnoses during 2011-2016 by season (Figure 1, panel A and B). In addition, laboratory reports of influenza A and influenza B and queries on influenza shared similar seasonal patterns (Figure 2). The MEM-calculated weekly queries on oseltamivir start during weeks 1-5 (alert weeks) and end during weeks 10-15.

**FIGURE 2** Queries on oseltamivir, influenza diagnoses, laboratory reports of influenza A and influenza B, and queries on influenza across Finland during 2011-2016 by season
(Figure 1, panel A and Table 1). The seasons peak during weeks 4-9 (Figure 3). Pre- and post-epidemic thresholds throughout the seasons were at 271 and 291 queries on oseltamivir, respectively. Influenza diagnoses calculated by MEM started during weeks 2-5 (alert weeks) and end during weeks 11-14 (Figure 1, panel B and Table 1). The seasons peak during weeks 4-9 (Figure 3). Pre- and post-epidemic thresholds throughout the seasons were at 146 and 162 influenza diagnoses, respectively. Table 1 and Figure 3 show the MEM-calculated starts and ends of the epidemic periods. Paired differences were computed by using the starting weeks. Paired differences in the mean showed statistical significance between queries on oseltamivir and influenza diagnoses (-0.80 weeks, 95% BCa CI: -1.0, 0.0, \( P = 0.000 \)) and between diagnoses and laboratory reports of influenza A (2.00 weeks, 95% BCa CI: 1.0, 3.2, \( P = 0.000 \)). Very high positive correlations were found between queries on oseltamivir and influenza diagnoses (\( r = 0.943 \)) and between diagnoses and laboratory reports of influenza A (\( r = 0.943 \)). Queries on influenza preceded queries on oseltamivir by -0.80 weeks (95% BCa CI: -1.2, 0.0, \( P = 0.015 \)), diagnoses by -1.60 weeks (95% BCa CI: -1.8, -1.0, \( P = 0.000 \)), and laboratory reports of influenza A by -0.80 weeks (95% BCa CI: -1.8, 0.4, \( P = 0.166 \)) and B by -2.40 weeks (95% BCa CI: -3.8, -1.0, \( P = 0.002 \)) (Figure 2, Table 2). In addition, paired differences in the mean showed statistical significance between queries on oseltamivir and laboratory reports of influenza A (1.20 weeks, 95% BCa CI: 0.0, 2.0, \( P = 0.000 \)), and statistically significant difference was found when laboratory reports of influenza A were paired with laboratory reports of influenza B (-2.80 weeks, 95% BCa CI: -5.6, -0.4, \( P = 0.041 \), \( r = -0.252 \)). The results of the paired differences in the mean, BCa CIs, \( p \)-values, and correlations are shown in Table 2.

### Table 1

<table>
<thead>
<tr>
<th>Season</th>
<th>Queries on oseltamivir</th>
<th>Influenza diagnoses</th>
<th>Laboratory reports of influenza A</th>
<th>Laboratory reports of influenza B</th>
<th>Queries on influenza</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Epidemic starts</strong></td>
<td><strong>Epidemic starts</strong></td>
<td><strong>Epidemic starts</strong></td>
<td><strong>Epidemic starts</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Date</strong></td>
<td><strong>Week</strong></td>
<td><strong>Date</strong></td>
<td><strong>Week</strong></td>
<td><strong>Date</strong></td>
</tr>
<tr>
<td>Start of epidemics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011/12</td>
<td>Jan 30 5</td>
<td>Jan 30 5</td>
<td>Jan 23 4</td>
<td>Jan 23 4</td>
<td>Jan 16 3</td>
</tr>
<tr>
<td>2012/13</td>
<td>Jan 14 3</td>
<td>Jan 21 4</td>
<td>Jan 14 3</td>
<td>Jan 14 3</td>
<td>Jan 7 2</td>
</tr>
<tr>
<td>2013/14</td>
<td>Jan 20 4</td>
<td>Jan 27 5</td>
<td>Jan 20 4</td>
<td>Jan 27 5</td>
<td>Jan 20 4</td>
</tr>
<tr>
<td>2014/15</td>
<td>Dec 29 1</td>
<td>Jan 5 2</td>
<td>Dec 15 51</td>
<td>Jan 26 5</td>
<td>Dec 29 1</td>
</tr>
<tr>
<td>2015/16</td>
<td>Jan 4 1</td>
<td>Jan 11 2</td>
<td>Dec 21 52</td>
<td>Feb 1 5</td>
<td>Dec 28 53</td>
</tr>
<tr>
<td>End of epidemics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011/12</td>
<td>Mar 26 13</td>
<td>Mar 26 13</td>
<td>Mar 19 12</td>
<td>May 7 19</td>
<td>Mar 12 11</td>
</tr>
<tr>
<td>2012/13</td>
<td>Apr 1 15</td>
<td>Apr 1 14</td>
<td>Apr 1 14</td>
<td>Apr 15 16</td>
<td>Mar 25 13</td>
</tr>
<tr>
<td>2013/14</td>
<td>Mar 17 12</td>
<td>Mar 24 13</td>
<td>Mar 24 13</td>
<td>May 19 21</td>
<td>Mar 10 11</td>
</tr>
<tr>
<td>2014/15</td>
<td>Apr 6 15</td>
<td>Mar 30 14</td>
<td>Mar 30 14</td>
<td>Apr 27 18</td>
<td>Mar 16 12</td>
</tr>
<tr>
<td>2015/16</td>
<td>Mar 7 10</td>
<td>Mar 14 11</td>
<td>Feb 22 8</td>
<td>May 2 18</td>
<td>Feb 22 8</td>
</tr>
</tbody>
</table>

4 | DISCUSSION

Our study showed similar patterns and statistically significant paired differences and high correlations between HCPs’ queries on oseltamivir and primary healthcare influenza diagnoses during the starting weeks of an epidemic in Finland during 2011-2016 (Figure 1, panel A and B, Table 2). In addition, high correlations and statistically significant paired differences were found between queries on oseltamivir and laboratory reports of influenza A, queries on influenza and laboratory reports of influenza A, queries on influenza and influenza diagnoses (Table 2). Paired differences estimated how much earlier the epidemic started. The smaller the \( p \)-value, the more statistically significant the paired difference was. The higher the correlation, the more similarly a paired indicator appeared between seasons. Pairs with negative correlations related to laboratory reports of influenza B (Table 2) indicate that the epidemics of influenza B appear seasonally later than influenza A (Figures 2 and 3).

Using the Internet enables users, including HCPs, to search for health-related information on diseases and medications. Internet searches for influenza-related health information on Google were found to be predictive of primary healthcare visits for influenza-like illnesses. Our study showed similar patterns and statistically significant paired differences and high correlations between HCPs’ queries on oseltamivir and primary healthcare influenza diagnoses during the starting weeks of an epidemic in Finland during 2011-2016 (Figure 1, panel A and B, Table 2). In addition, high correlations and statistically significant paired differences were found between queries on oseltamivir and laboratory reports of influenza A, queries on influenza and laboratory reports of influenza A, queries on influenza and influenza diagnoses (Table 2). Paired differences estimated how much earlier the epidemic started. The smaller the \( p \)-value, the more statistically significant the paired difference was. The higher the correlation, the more similarly a paired indicator appeared between seasons. Pairs with negative correlations related to laboratory reports of influenza B (Table 2) indicate that the epidemics of influenza B appear seasonally later than influenza A (Figures 2 and 3).
platforms may provide HCPs with medical information of question-
able quality. Although influenza searches from Google have been assessed in terms of disease surveillance, this method included several weaknesses in timing and regional scales to predict influenza epidemics. Notably, general search engines cannot characterize the users performing the searches. We have shown here that HCPs’ queries on oseltamivir and influenza from the dedicated Internet portal highly coincided with diagnoses and laboratory reports of influenza A and influenza B.

This study includes certain limitations to be taken into consideration. In this work, we studied queries on oseltamivir and influenza in the whole country including no data on geographical variations. The

| TABLE 2 | Pairs, paired differences with the mean, bias-corrected and accelerated confidence intervals and p-values, and Kendall’s correlations. Queries on oseltamivir, influenza diagnoses, laboratory reports of influenza A and influenza B, and queries on influenza were paired and calculated and bootstrapped according to the start of the epidemic weeks shown in Table 1 |

<table>
<thead>
<tr>
<th>Pair</th>
<th>Paired differences</th>
<th>Bias-corrected and accelerated 95% confidence interval of the mean (adjusted for ties)</th>
<th>p-value of the mean</th>
<th>Kendall’s correlation coefficient (τ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Queries on oseltamivir – Influenza diagnoses</td>
<td>−0.80</td>
<td>−1.0</td>
<td>0.0</td>
<td>0.000</td>
</tr>
<tr>
<td>Influenza diagnoses – Laboratory reports of influenza A</td>
<td>2.00</td>
<td>1.0</td>
<td>3.2</td>
<td>0.000</td>
</tr>
<tr>
<td>Queries on influenza – Influenza diagnoses</td>
<td>−1.60</td>
<td>−1.8</td>
<td>−1.0</td>
<td>0.000</td>
</tr>
<tr>
<td>Queries on oseltamivir – Laboratory reports of influenza A</td>
<td>1.20</td>
<td>0.0</td>
<td>2.0</td>
<td>0.021</td>
</tr>
<tr>
<td>Queries on influenza – Queries on oseltamivir</td>
<td>−0.80</td>
<td>−1.2</td>
<td>0.4</td>
<td>0.015</td>
</tr>
<tr>
<td>Queries on influenza – Laboratory reports of influenza A</td>
<td>−0.80</td>
<td>−1.8</td>
<td>0.4</td>
<td>0.166</td>
</tr>
<tr>
<td>Queries on influenza – Laboratory reports of influenza B</td>
<td>−2.40</td>
<td>−3.8</td>
<td>−1.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Laboratory reports of influenza A – of influenza B</td>
<td>−2.80</td>
<td>−5.6</td>
<td>−0.4</td>
<td>0.041</td>
</tr>
<tr>
<td>Influenza diagnoses – Laboratory reports of influenza B</td>
<td>−0.80</td>
<td>−2.4</td>
<td>1.0</td>
<td>0.354</td>
</tr>
<tr>
<td>Queries on oseltamivir – Laboratory reports of influenza B</td>
<td>−1.60</td>
<td>−3.2</td>
<td>0.4</td>
<td>0.088</td>
</tr>
</tbody>
</table>
influenza epidemic may start in different regions of a country at different times, and the detection of this would require more accurate location data. PD and the national register of public primary healthcare diagnoses are separate databases; thus, queries and diagnoses cannot be connected to one another nor the same patient visited primary care physician’s encounter. It is possible that the great number of queries on oseltamivir (Figure, panel A) compared to influenza diagnoses (Figure, panel B) also includes the presence of some secondary healthcare queries on oseltamivir opened from PD. In addition, during influenza epidemics, the national public primary healthcare register may exclude some influenza diagnoses that were misdiagnosed and reported as within a broader category of infectious diseases, such as acute respiratory infections. Also, not every patient presenting influenza symptoms is tested for the virus. In these cases, the prescription is based on symptoms and knowledge on the epidemic situation of the influenza virus in the given population. In the Finnish primary health care, oseltamivir is rarely prescribed outside influenza epidemic periods and is, therefore, less well-known than medications that are constantly prescribed. This results in queries on oseltamivir specifically during epidemics. However, when treating influenza patients, physicians may still only check up on indications of oseltamivir from PD without then prescribing the medication, thus increasing the number of queries in a log file. Also, media coverage may affect HCPs’ queries during and outside epidemics. Not only influenza publications in the media may impact HCPs themselves, but also patients may face influenza-related media publications. During clinical encounters, patients may present certain symptoms or demand influenza medication to stockpile them worrying pandemics, resulting in possible search trends from PD. Since queries on influenza from PD preceded other indicators (Table 1, Figures 2 and 3), HCPs may have searched influenza-related information during encounters prior to an epidemic. In addition, the magnitudes between the patterns of queries on oseltamivir and diagnoses may also vary due to different qualities of the databases. However, since MEM was used to calculate only starts and ends of the epidemic periods, the heights of the spikes did not have an effect on the outcomes of the analyses. The strength of our study was that HCPs (representativeness) search for information from real-time Internet databases (timeliness).

This is the first study, to our knowledge, to demonstrate high correlation and statistical significance between queries on oseltamivir and primary healthcare influenza diagnoses by using the MEM model. We found that HCPs’ information searching behavior strongly associates with epidemiological data on influenza. Therefore, we state that HCPs' queries could be used as a supplementary source of information for disease surveillance when detecting influenza epidemics. Our study depicts a possible development for infectious disease surveillance systems. While our study utilizes a database unique for Finland, similar medical databases can be used to assess data in European countries and internationally. The combination of Internet-based query data and other surveillance data could enhance current surveillance systems. In the future, it may be possible to create algorithms that analyze HCPs’ queries in real time in order to help the detection of the beginning epidemic. Information from these different sources could be combined and delivered to primary healthcare units facing the first patients at the start of an epidemic to estimate the need for primary healthcare services and workforce during epidemics. Further studies should focus on the applicability of these results in different pathologies and other medical databases in other countries.

CONFLICT OF INTERESTS

MK has held various trustee positions in the Finnish Medical Society Duodecim since the late 1990s. OH has held various trustee positions in the Finnish Medical Society Duodecim and Duodecim Medical Publications Ltd since 2009 and is a partner at iHealth Finland Ltd. The other authors have no competing interests.

AUTHORS’ CONTRIBUTIONS

SP, MJV, MM, KY, PM, and OH designed the study concept. SP, MJV, MM, and OH performed the literature research. SP, MJV, KY, and OH gathered and supplied the study data. SP, MJV, and KY carried out the data analysis. SP, MJV, and OH carried out the data interpretation. SP, MJV, MM, KY, PM, MK, and OH involved in the critical revision and final approval. SP and MM drafted the manuscript.

ORCID

Samuli Pesälä https://orcid.org/0000-0002-6090-1367
Mikko J. Virtanen https://orcid.org/0000-0003-4120-4432
Milla Mukka https://orcid.org/0000-0001-6148-4158
Kimi Ylilammi https://orcid.org/0000-0002-0601-1363
Pekka Mustonen https://orcid.org/0000-0002-5077-7788
Minna Kaila https://orcid.org/0000-0002-9645-4925
Otto Helve https://orcid.org/0000-0001-6254-3180

REFERENCES


APPENDIX
Kendall’s tau formula.