ANALYZING SEARCH BEHAVIOR OF HEALTHCARE PROFESSIONALS FOR DRUG SAFETY SURVEILLANCE

DAVID J. ODGERS
Center for Biomedical Informatics, Stanford University, 1265 Welch Road
Stanford CA 94305-5479 USA
Email: djodgers@stanford.edu

RAVE HARPAZ
Center for Biomedical Informatics, Stanford University, 1265 Welch Road
Stanford CA 94305-5479 USA
Email: rharpaz@stanford.edu

ALISON CALLAHAN
Center for Biomedical Informatics, Stanford University, 1265 Welch Road
Stanford CA 94305-5479 USA
Email: alison.callahan@stanford.edu

GREGOR STIGLIC
Faculty of Health Sciences, University of Maribor, Zitna ulica 15
2000 Maribor, Slovenia
Email: gregor.stiglic@um.si

NIGAM H. SHAH
Center for Biomedical Informatics, Stanford University, 1265 Welch Road
Stanford CA 94305-5479 USA
Email: nigam@stanford.edu

Post-market drug safety surveillance is hugely important and is a significant challenge despite the existence of adverse event (AE) reporting systems. Here we describe a preliminary analysis of search logs from healthcare professionals as a source for detecting adverse drug events. We annotate search log query terms with biomedical terminologies for drugs and events, and then perform a statistical analysis to identify associations among drugs and events within search sessions. We evaluate our approach using two different types of reference standards consisting of known adverse drug events (ADEs) and negative controls. Our approach achieves a discrimination accuracy of 0.85 in terms of the area under the receiver operator curve (AUC) for the reference set of well-established ADEs and an AUC of 0.68 for the reference set of recently labeled ADEs. We also find that the majority of associations in the reference sets have support in the search log data. Despite these promising results additional research is required to better understand users' search behavior, biasing factors, and the overall utility of analyzing healthcare professional search logs for drug safety surveillance.

1. Background

Drug safety surveillance is a significant problem given the frequency of post-market adverse drug events (ADEs) that compromise patient health and result in increased costs and burden on the healthcare system [1-4]. The FDA adverse event reporting system (FAERS) is currently the main source for detecting post-market ADEs, but has recognized shortcomings [2, 5-8] such as under-reporting of adverse events (AEs), an issue that is not unique to FAERS [9]. The time required for a sufficient signal when using such sources can delay the release of ADE alerts [10, 11]. As alternatives to spontaneous ADE reports, other sources of data that have been used for detecting drug-AE associations include claims data [12], electronic health records (EHRs) [13-15] and consumer search logs [16-18]. There is a growing trend to use increasingly diverse sources for detecting ADEs, including online social networks. For example, Twitter has been used as a source for mining drug-AE associations [19] with the recognition that more work is required to establish threshold signal levels from such sources towards validating discovered associations. Another data source that is being investigated for ADE detection is the biomedical literature [20-22]. In addition to examining a larger variety of sources, approaches have also been recently developed to combine signals from data sources such as EHRs, claims, biomedical literature, and Internet search logs with signals from FAERS [18, 23, 24].

In this work we explore the potential for using search logs from healthcare professionals as an observational data source for drug safety surveillance. We use two years of search logs from UpToDate, an online source of health information provided by Wolters-Kluwer that includes detailed descriptions of symptoms, diseases, drugs and indications to support evidence based medicine. UpToDate is used on a subscription basis by institutions and any individual who purchases a license. Typically, medical and research institutions purchase licenses for UpToDate, which are then used by members of the institution – this can include physicians, researchers and students. UpToDate use in hospitals is associated with fewer patient complications and adverse events, shorter hospital stays, reduced mortality rates and higher quality performance measures [25, 26]. How medical professionals use UpToDate thus has potentially direct implications for patient health, some of which may be discovered by analyzing the records of this use. Logs of UpToDate use capture the source institution and machine used for a search, the search string entered, the time and date of the search, the type of search, and UpToDate pages visited as a result of the search. UpToDate search logs are thus a rich resource with the potential to enable time-sensitive and context-aware analyses of search behavior. UpToDate search logs are also an important source of observational data for drug safety surveillance because the majority of UpToDate users are health professionals who access UpToDate during their day-to-day practice of providing patient care. In contrast, web search logs and social media capture a broad range of online behavior with unknown context(s).

We investigate an approach to detect drug-AE pairs by first annotating individual UpToDate search logs to identify drugs and events, and then performing association analysis on drug-event pairs identified within a predefined surveillance period. The novelty of our approach is that it capitalizes on a previously untapped source of observational data for detecting drug-AE associations. We assess the performance of our approach by using a reference standard of well-

established associations from the EU-ADR project as well as a reference set of recently labeled ADEs obtained from 2013 FDA product labeling revisions. Our findings are two-fold: (1) the majority of associations from the two reference sets have support in the UpToDate search logs and (2) the approach we investigated can detect ADEs with an accuracy (measured using the area under the receiver operating curve, AUC) of 0.85 for the reference set of well-established ADEs, and an AUC of 0.68 for the reference set of recently labeled ADEs—a result that merits further investigation. It is possible that better methods are required for analyzing healthcare professional search logs to reliably detect more recently reported (and potentially unreported) ADEs.

2. Methods

2.1. *Search logs used for analysis*

We used two years of UpToDate search logs spanning January 2011 to December 2012 as the data source for exploring associations between drugs and AEs. The log for a single UpToDate search consists of a query string, unique session ID, Internet protocol (IP) address of the computer on which the search was performed and the timestamp of the search. The logs were pre-processed to include only those records for string search queries. Example search logs are shown in Table 1.

Table 1. Three example UpToDate search logs. Institution, session ID and IP addresses have been replaced with fake values but search terms are taken directly from existing records.

Search term	*Institution*	*Session ID*	*IP address*	*Time*
adrenal insufficiency	A	000018A4C27BDCC	123.456.789	2012-08-01 20:13:23
elevated LDH	B	002679154AE8055C	456.789.199	2012-08-01 15:14:31
effexor xr	C	01B6280B71230987	789.123.456	2012-08-01 11:16:15

The session ID allows us to identify queries performed within a single session, the IP address allows us to identify searches performed on a single machine, and the timestamp enables the temporal reconstruction of the click sequence within a given session.

2.2. *Query annotation*

We processed the query strings of the search logs using our previously described text processing workflow [27, 28] to annotate the queries with drug and disease terms from biomedical ontologies. Briefly, query strings were processed using MGREP with a lexicon of more than 3 million strings built from biomedical ontologies and terminologies, in which terms and concepts are mapped by synonymy and parent-child relationships. Query strings that match drug names (*e.g.* brand drugs, multi-ingredient drugs, different preparations, dose, form, and salt ingredients) are mapped to the drug's active ingredient by using RxNorm relationships to obtain all RxNorm codes associated with the given active ingredient. These were then mapped to UMLS concept codes (concept unique identifiers, or CUIs). Query strings that match condition mentions (*i.e.* potential adverse events) are also annotated with UMLS CUIs based on mappings between terms and concepts in UMLS.

2.3. *Modeling assumptions*

In performing our analysis we make several assumptions about the nature of the UpToDate log data and the manner in which healthcare professionals search for information about drugs and adverse events they are concerned about. We acknowledge that healthcare professionals may search for drug related information beyond the case of adverse events (*e.g.* drug indications), but in this study we restrict our analysis and target only adverse events. We assume that when a healthcare professional is concerned about a potential association between a drug and an AE, and desires to retrieve information about the association, they will perform searches for the drug and event within a short period of time (possibly even in the same query). It is assumed that the longer the time interval between the search for a drug and event, the less likely it is that the searches for drug and event are related to each other or part of the same thought or decision process. We do not assume an ordering on the search terms for the drug and event of interest (either the drug or the AE can be searched for first). We allow that a user may search for the drug and AE multiple times and cycle through searches of the same drug and AE within the same session. We further assume that each session (identified by the session ID) was performed by a single unique healthcare professional (but not the converse), and allow that a unique IP address may be associated with multiple different healthcare professionals. Given these assumptions, we use unique sessions, rather than unique queries or unique IPs as our basic unit for counting and subsequent association analyses.

2.4. *Association analysis*

We consider three scenarios related to the search order of given drug-event pair of interest: (1) the search for the drug precedes the search for the event; (2) the search for the event precedes the drug; and (3) order is irrelevant. We restrict the amount of time that may elapse between the search of a drug and an event (depending on the ordering) by defining a surveillance period that is indexed (starts) with the first mention of a drug or event (depending on ordering) within a session, and ends at a pre-specified amount of time later within the same session (see Figure 1). For a given drug-event pair of interest, a given pre-specified surveillance period, and a given pre-specified drug-event search ordering, we compute the following 2x2 contingency table

	event	no event
drug	**a**	**b**
no drug	**c**	**d**

where **a** is the number of unique sessions including search terms associated with the drug and event of interest that falls within the surveillance period indexed by the first mention of a drug (or event), **b** is the number of unique sessions including search terms associated with the drug of interest but not the event of interest that conform with the surveillance period restriction, **c** is the number of unique sessions including search terms associated with the event of interest but not the

drug of interest that conform with the surveillance period restriction, and **d** is equal to the total number of unique sessions minus the sum of the cells **a**, **b**, and **c**. Having computed the 2x2 contingency table for a given drug-event pair of interest, we use the Odds Ratio measure, with a zero-cell correction (adding 0.5 to each count in a 2x2 table that contains zero in any of its cells), to quantify the strength of association between the drug and event. As our final association score we use the 5th percentile of the odds ratio distribution, which provides an adjustment for small sample sizes and protects against false associations due to chance. This type of adjustment has been shown to provide greater accuracy than point estimates in a related signal detection study [29] and was also found to provide greater accuracy in this study. We note that by performing a two-dimensional (2x2) analysis we omit to account for potential biasing factors such as temporal trends, media influence, or the search habits of healthcare professionals. We plan to examine these and other types of biases in a follow-up study.

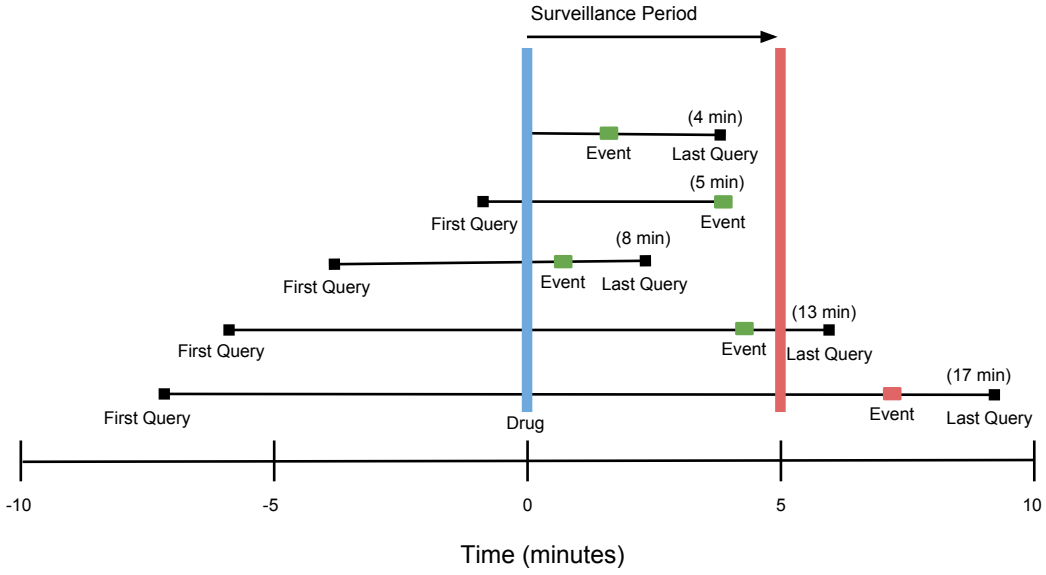

Figure 1. Illustration of a surveillance period relative to UpToDate search sessions that contain occurrences of 'drug then event' queries. A single search session is shown as a horizontal line that persists for the time length of the session. Sessions that contain the same drug and event as queries are aligned by assigning the time of drug mention as time 0. Within each session, the instance of drug and event queries with the shortest time difference is compared to the surveillance period. Sessions that have their corresponding shortest drug → event mention time within the surveillance period are counted towards the association analysis (shown as sessions with the event in green). Sessions that have a drug → event time difference that is greater than the surveillance period are not counted (shown as the session with the event in red).

2.5. *Evaluation*

We evaluate our approach using two reference standards, each consisting of a set of positive test cases of drug-event pairs recognized as truly associated (true ADEs), and a set of negative controls of drug-event pairs that are highly unlikely to be associated. The first reference standard we use was created by the EU-ADR project [30] for the retrospective evaluation of ADE signal detection

approaches based on EHRs and claims data. It includes 44 positive test cases and 50 negative controls spanning ten serious adverse events (*e.g.* acute myocardial infarction, rhabdomyolysis) and 68 unique drugs. The drug-event associations comprising the EU-ADR positive test cases have been well known for a relatively long period of time. This reference set—when used to evaluate methods applied to data representing recent events such as our queries from 2011-2012, which may be influenced by existing information about well-known ADEs—will only provide an indication of the retrospective performance of the method.

Therefore, we used a second reference standard [31] for evaluation that was built specifically for the prospective evaluation of ADE signal detection methods. This time-indexed reference standard was manually curated from all product label updates (*e.g.* warnings) communicated by the FDA in 2013 through monthly summaries posted at FDA's MedWatch website [32] and classified as Boxed Warning, Warning (or Precautions), and Adverse Reactions. In addition, we considered only drugs that are orally administered. Each candidate association extracted from the MedWatch website was manually verified (by reviewing the drug's labeling history obtained from drug@FDA [33]) to ensure that the association is indeed new and is not qualified by a co-existing contraindicated situation or risk factor (this type of association requires non-standard analysis approaches). The drugs underlying each association were normalized to RxNorm active ingredients. Closely related events underlying each association that are synonymous or describe the same clinical syndrome were grouped and given a unique representative name. The final reference set includes 62 positive test cases and 75 negative controls, and covers 44 drugs and 38 events ranging from common and mild to rare and serious. The negative controls were created by randomly pairing the same drugs and events that appear in the set of positive test case, and verifying that these random pairs are not reported in the underlying drug's label as an ADE. The time stamp attached to each positive test case reflects the date on which a label corresponding to a drug was revised to include the adverse event associated with the drug. It is assumed that this date reflects the time by which an association became known to the general public. Consequently, using data that pre-dates the label update year of 2013 (as with the UpToDate logs used in our method) provides an indication of the prospective performance of an association detection method.

Both the EU-ADR and the time-index reference standards define each of the included AEs as sets of UMLS concepts that relate to and are consistent with a description of the overall clinical condition associated with the particular AE (the definitions were created manually by researchers with medical training). For example, the UMLS concepts for bullous eruption (C0235818), erythema multiforme (C0014742), Stevens-Johnson syndrome (C0038325), toxic epidermal necrolysis (C0014518), and bullous dermatosis (C0085932), compose the group of UMLS concepts that define the 'bullous eruptions' AE concept in the EU-ADR reference standard. In our analysis, a query was considered to mention a given AE if any of its search terms (or term sequences) was mapped to one of the UMLS concepts that define the underlying adverse event.

Drugs in the EU-ADR reference standard are specified using bottom-level ATC codes (ATC5), whereas drugs in the time-indexed reference standard are specified using RxNorm codes at the base active ingredient level (IN). The ATC codes were mapped to RxNorm INs. A query was considered to mention a given drug if any of its search terms (or term sequences) was mapped

to one of the UMLS CUIs that represent the drug's active ingredient, thus our analysis and evaluation was performed at the drug ingredient level. Queries that included the drugs and adverse events were tagged with the corresponding drug and adverse event concept identifiers. The tags were then used to compute 2x2 contingency tables and statistical associations as described before. Using the calculated association score for each drug-event pair (from each of the reference standards), the performance of our approach to identify ADEs was quantified using the area under the receiver operating characteristic curve (AUC).

3. Results

3.1. *Descriptive statistics*

We processed 320,975,093 unique UpToDate search queries associated with 134,217,800 unique sessions and 3,986,773 unique IP addresses. The annotation process resulted in 144,888 unique CUIs which includes 40,416 unique drug concepts and 47,220 unique clinical condition concepts. Sessions constitute our basic unit of count, and so we present several session-related statistics (Table 2). Session length is determined as the difference between the time of the first and last search in the session, and thus sessions that consist of a single search have a recorded time of 0 seconds. An average session lasted 20 minutes and consisted of 5 queries.

Table 2. UpToDate search session statistics.

	Queries Per session	*Session Duration (secs)*
Min	1	0
Max	304283	17373968
Median	3	0
Mean	5	1202

For the positive test cases of the EU-ADR reference standard, the median time between the query of a drug and event (regardless of ordering) was 621 seconds. For the time-indexed reference standard the median was 1,979 seconds. A possible explanation for this difference is that the EU-ADR positive test cases are better documented in UpToDate due to the fact that they are well known, and thus lead to shorter search sessions. Table 3 displays the top three associations that were searched for regardless of ordering within a surveillance period of five minutes. The table also lists the most frequently occurring search terms associated with each drug and event.

Table 3. Top drug-AE associations from the EU-ADR and time-indexed reference sets, and the most frequently occurring search terms that are evidence for the association.

	Top association (count)	*Top search terms (count)*
EU-ADR	acetylsalicylic acid; upper gastrointestinal bleed (1683)	**Drug**: aspirin (1639), Aggrenox (32), Excedrin (7), alka-seltzer (3)
		Event: gastrointestinal hemorrhage (1652) melena (15), hematemesis (13), upper gastrointestinal hemorrhage (3)
	furosemide; acute renal failure (1026)	**Drug**: Lasix (639), furosemide (387)
		Event: acute kidney failure (879), anuria (92), acute kidney tubular necrosis (55)
	simvastatin; rhabdomyolysis (523)	**Drug**: simvastatin (410), Zocor (92), vytorin (21)
		Event: rhabdomyolysis (519), myoglobinuria (4)
Time-indexed	ursodeoxycholate; liver damage (2291)	**Durg**: ursodiol (1844), actigall (248), urso (187), ursofalk (11), urdox (1)
		Event: liver cirrhosis (933), liver diseases (451), hyperbilirubinemia (335), hepatitis (332), icterus (167)
	levetiracetam; movement disorders (597)	**Drug**: keppra (496), levetiracetam (101)
		Event: myoclonus (167), tremor (99), spasm (56), dyskinetic syndrome (44), ataxia (40)
	ketoconazole; liver damage (190)	**Drug**: ketoconazole (176), nizoral (14)
		Event: liver diseases (85), hepatitis (56), hepatotoxicity (29), liver cirrhosis (7), icterus (6)

3.2. *Performance evaluation*

We explored different drug-event orderings and different surveillance periods as defined in the Methods section. We varied the surveillance period from 0 seconds to 100 minutes at intervals of 100 seconds. The surveillance period accounts for the time between the first and second search corresponding to the annotated drug or event in the search logs. A temporal measurement was used to take into consideration the time required to read and process the first search result of a drug/event pair before searching for the corresponding drug or event in the pair. For each combination of a surveillance period length and a drug-event ordering we computed the AUC using the corresponding association scores, and the proportion of test cases that the data supported evaluation of. We define test cases with data support as those test cases that had at least one search session containing the drug and AE ($a>0$ in the 2x2 contingency table) during the surveillance period. This was used to provide additional insight into the analysis, and is also used to identify the optimal surveillance period as a tradeoff between discrimination accuracy (AUC) and the number of test cases that are supported by the data. The more test cases with data support, the more reliable the performance evaluation. A very small surveillance period that results in only a small portion of test cases that are supported by the data would not be useful in a real setting. Such results cannot be used as an indicator of a method's success in identifying true associations.

Table 4 and Figures 2 and 3 provide a summary of our performance evaluation. Table 4 shows the maximum AUCs obtained by our approach and their corresponding surveillance periods.

Expectedly, retrospective evaluation based on the EU-ADR reference standard yields much better performance (maximum AUC=0.85) than the prospective evaluation based on the time-indexed reference standard (maximum AUC=0.68). The results also suggest that drug-event ordering does not dramatically affect performance, though the best performance is achieved by using the 'drug *then* event' ordering for both reference standards. The optimal surveillance period is small (approximately 4-10 minutes), which is consistent with our modeling assumptions.

Table 4. Maximum AUCs and corresponding surveillance period lengths for sessions where a drug was searched for before an event, event before drug, and either ordering.

	EU-ADR		***Time-indexed***	
Search direction	**Max. AUC**	**Search window**	**Max. AUC**	**Search window**
Drug → event	0.85	665	0.68	303
Event→ drug	0.82	287	0.67	276
Either direction	0.84	245	0.67	296

Figures 2 and 3 display graphs of discrimination performance as a function of surveillance periods (0 seconds to 100 minutes at intervals of 100 seconds) for the EU-ADR and the time-indexed reference standards, respectively. For illustrative purposes we used the 'drug *then* event' ordering, and note that similar performance patterns were obtained for the two other drug-event orderings. Each figure displays the AUC (top sub-figure) and the proportion (bottom sub-figure) of both positive and negative test cases supported by the data (test cases that had at least one query within the surveillance period).

Results for both reference standards suggest that performance fluctuates for small surveillance periods and then generally stabilizes, albeit with a decreasing trend as the surveillance period increases. As expected, the figures show that (1) the positive test cases have consistently more data support than the negative test cases; (2) the test cases of the EU-ADR reference standard have more data support than those of the time-indexed reference standard with smaller surveillance periods; and (3) the difference between the proportion of positive test cases and negative test cases with data support is much larger for the EU-ADR reference standard than for time-indexed reference standard. This is consistent with the larger AUCs obtained for the EU-ADR reference standard.

An encouraging result provided by the graphs is that a relatively large proportion (>80%) of positive test cases from the time-indexed reference standard are supported by data with short surveillance periods, around 5 minutes in duration. We hypothesize that this is the time an engaged user would take to read drug information on UpToDate, reason about a connection to an adverse event and then search for the specific event. Almost all the positive test cases are eventually supported by data (queries) with large surveillance periods, though most likely due to chance. The surveillance period at which these proportions start to converge is an alternative point to assess performance since it balances the tradeoff between discrimination accuracy and increasing the test cases that can be identified in the data. For the time-indexed reference standard this optimal point is to the right of the surveillance period yielding the maximum AUC, whereas

for the EU-ADR reference standard it appears to its left. At these points only a modest amount of accuracy is sacrificed, with the benefit that maximum test cases can be identified through the data.

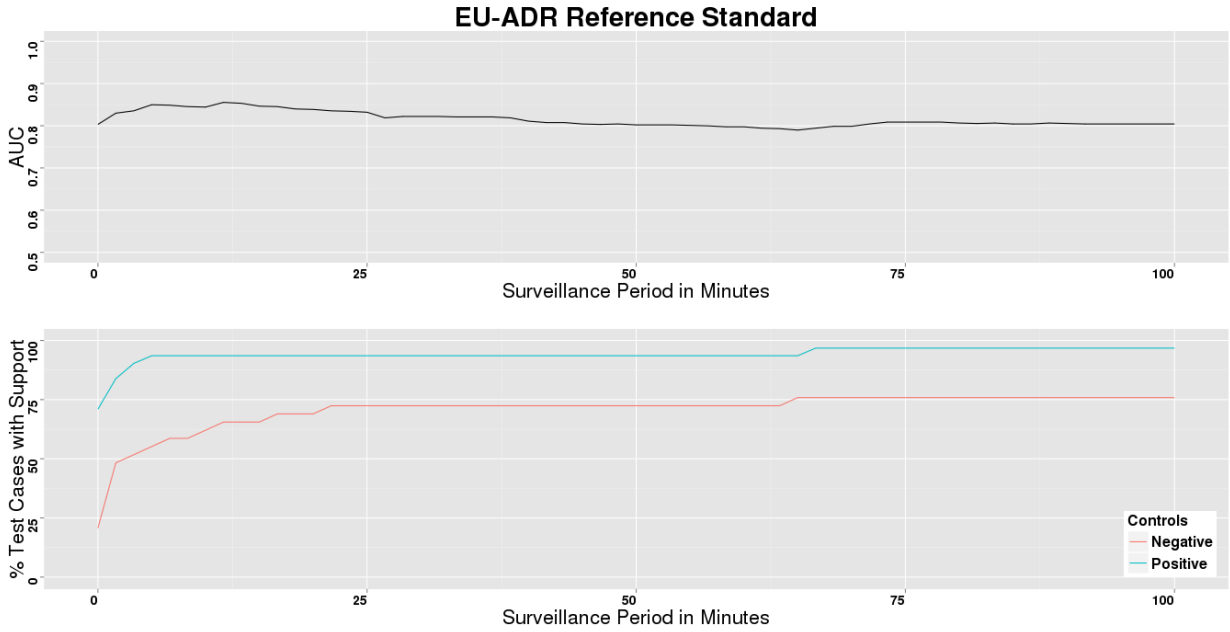

Figure 2. Performance as a function of surveillance period for the EU-ADR reference standard. The top graph shows the change in AUC as surveillance period increases. The bottom graph shows the percent of test cases with data support as surveillance period increases.

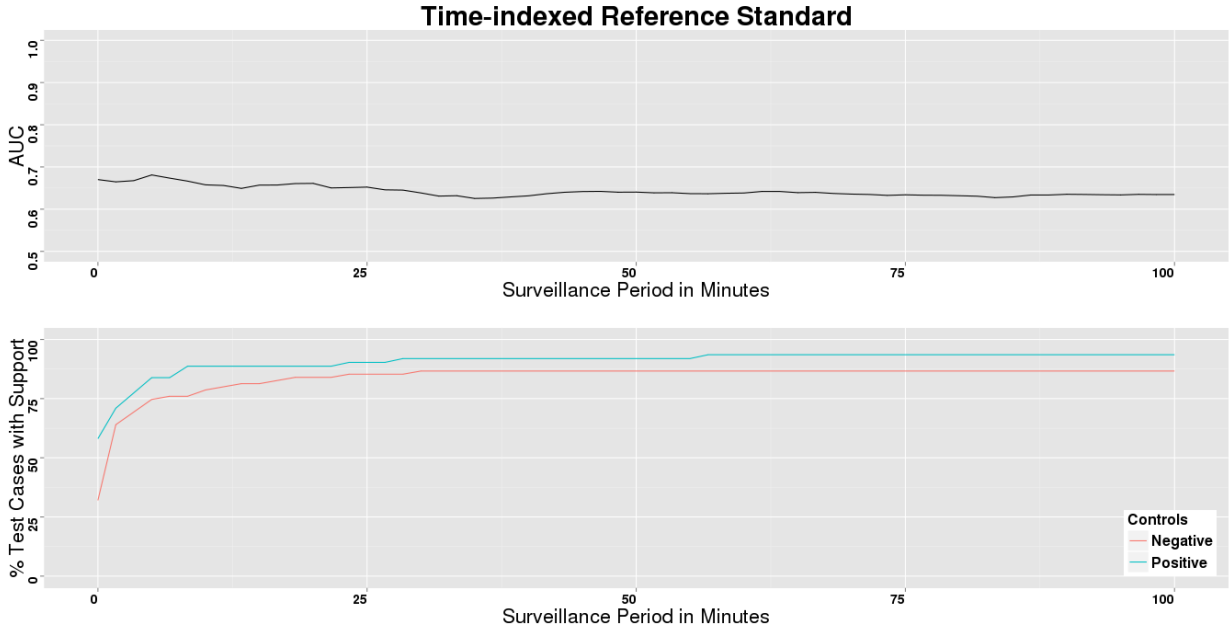

Figure 3. Performance as a function of surveillance period for the time-indexed reference standard. The top graph shows the change in AUC as surveillance period increases. The bottom graph shows the percent of test cases with data support as surveillance period increases.

4. Discussion

This work describes a preliminary investigation into mining drug-AE associations from healthcare professional search logs, a data source that to our knowledge has not been used for drug safety surveillance. In a recent commentary [34], Sarntivijai and Abernethy discuss the merits of using Internet search logs in general as a source for ADE signal detection and emphasize the need for more research in processing and assessing the validity of any discovered drug-AE associations. This work represents such an effort, and our findings suggest that physician search logs could be used as a data source for detecting ADEs.

The performance of our method on the drug-AE pairs (AUC of 0.85) from the EU-ADR project is not surprising, given that these ADEs are likely to be already described in UpToDate and also known to the medical professionals using UpToDate as a reference source; both factors that can lead users to query for the drugs and AEs. The lower AUC of 0.68 for the time-indexed reference standard of more recently labeled ADEs achieved by our method indicates that it is more challenging to detect relatively new or emerging ADEs. Despite the relatively high AUCs achieved by our method, we also observed lower association scores for drug-AE pairs from the time-indexed reference set. This finding warrants further investigation to determine which (if any) of the detected associations may be due to chance or if an odds ratio based association score is a discriminatory enough metric for search log analysis. Given the challenges in detecting relatively new or emerging ADEs in UpToDate logs, this datasource's utility might hinge on development of methods to combine data sources.

The possibility of chance ADE discoveries exists due to the fact that healthcare professionals may search for a given drug and disease within a single UpToDate search session for a variety of reasons, and the existence of multiple terms within a session is not a guarantee that the searches are associated. Caution must therefore be exercised when analyzing such records using data mining methods to avoid detection of spurious associations. Our experiments show that relatively short surveillance periods (and therefore search session length) are optimal for detecting ADEs, and this is consistent with our modeling assumptions regarding search behavior. With long enough surveillance periods, any drug and event may be found to be associated, and so the use of shorter periods can provide some protection against such effects.

The next frontier in drug safety surveillance involves determining time to signal, quantifying the plausibility of potential drug-AE pairs, estimating patients affected by ADEs, prioritizing ADEs by severity, and in combining data sources to improve drug safety surveillance. Our preliminary findings indicate that healthcare professional search logs can be a potential data source for advancing drug-safety research. We have recently demonstrated the performance of combining drug safety signals [23] and future work will build on these efforts by incorporating healthcare professional search logs as a complementary data source. We will also investigate the use of other methods to detect ADEs from physician search logs, develop additional benchmarks, develop more accurate search behavior models (perhaps by profiling or surveying users), and identify the type of events that are appropriate for surveillance via healthcare professional search logs.

Acknowledgements

We acknowledge funding support from NIH grant R01 GM101430. NHS and RH also received funding support from R01 LM011369 and U54 HG004028.

References

1. Classen, D.C., et al., JAMA, **277** (4), 301 (1997).
2. Lazarou, J., B.H. Pomeranz, and P.N. Corey, JAMA, **279** (15), 1200 (1998).
3. Classen, D.C., et al., Health Aff (Millwood), **30** (4), 581 (2011).
4. Hug, B.L., et al., Jt Comm J Qual Patient Saf, **38** (3), 120 (2012).
5. Honig, P.K., Clin Pharmacol Ther, **93** (6), 474 (2013).
6. Harpaz, R., et al., Clin Pharmacol Ther, **91** (6), 1010 (2012).
7. Platt, R., et al., N Engl J Med, **361** (7), 645 (2009).
8. Stang, P.E., et al., Ann Intern Med, **153** (9), 600 (2010).
9. Levinson, D.R., Department of Health and Human Services, Available from http://oig.hhs.gov/oei/reports/oei-06-09-00091.pdf (2012).
10. Avorn, J. and S. Schneeweiss, N Engl J Med, **361** (7), 647 (2009).
11. Hauben, M. and A. Bate, Drug Discov Today, **14** (7-8), 343 (2009).
12. Schneeweiss, S., et al., Epidemiology, **20** (4), 512 (2009).
13. Coloma, P.M., et al., Pharmacoepidemiol Drug Saf, **21** (6), 611 (2012).
14. LePendu, P., et al., Clin Pharmacol Ther, **93** (6), 547 (2013).
15. Iyer, S.V., et al., J Am Med Inform Assoc, **21** (2), 353 (2014).
16. Yom-Tov, E. and E. Gabrilovich, J Med Internet Res, **15** (6), e124 (2013).
17. White, R.W., et al., J Am Med Inform Assoc, **20** (3), 404 (2013).
18. White, R.W., et al., Clin Pharmacol Ther, **96** (2), 239 (2014).
19. Freifeld, C.C., et al., Drug Saf, **37** (5), 343 (2014).
20. Shetty, K.D. and S.R. Dalal, JAMIA, **18** (5), 668 (2011).
21. Avillach, P., et al., J Am Med Inform Assoc, **20** (3), 446 (2013).
22. Pontes, H., M. Clement, and V. Rollason, Drug Saf, **37** (7), 471 (2014).
23. Harpaz, R., et al. *KDD* (2013).
24. Xu, R. and Q. Wang, BMC Bioinformatics, **15** (2014).
25. Bonis, P.A., et al., Int J Med Inform, **77** (11), 745 (2008).
26. Isaac, T., J. Zheng, and A. Jha, J Hosp Med, **7** (2), 85 (2012).
27. Lependu, P., et al., J Biomed Semantics, **3 Suppl 1** S5 (2012).
28. Lependu, P., et al., AMIA Jt Summits Transl Sci Proc, **2012** 63 (2012).
29. Harpaz, R., et al., Clin Pharmacol Ther, **93** (6), 539 (2013).
30. Coloma, P.M., et al., Drug Safety, **36** (1), 13 (2013).
31. Harpaz, R., et al., Nature Scientific Data, **In press** (2014).
32. U. S. Food and Drug Administration. http://www.fda.gov/Safety/MedWatch/SafetyInformation/Safety-RelatedDrugLabelingChanges/default.htm (2014) [cited 2014 September 28].
33. U. S. Food and Drug Administration. http://www.accessdata.fda.gov/scripts/cder/drugsatfda/ (2014) [cited 2014 September 28].
34. Sarntivijai, S. and D.R. Abernethy, Clin Pharmacol Ther, **96** (2), 149 (2014).