Data-Tracking on Government, Non-profit, and Commercial Health-Related Websites

INTRODUCTION

Health information privacy is a core value of American health. The privacy of communication between healthcare providers and patients is protected by law, but no equivalent protection exists for individuals performing health-related searches on patient-facing websites. Consumer aggregated data obtained from searches on patient-facing websites can be used to create a health profile and even merge it with non-health-related information. These data profiles can be sold to other companies or used to curate targeted advertisement that follows the individual user. While this information collection is acknowledged, the degree of such data releases and specificity to the individual is often unknown to individual users. Therefore, we used a privacy inspection tool to determine prevalence and type of data tracking from commonly searched government and non-government health-related websites.

METHODS

Our sample was restricted to US-based websites from the most trafficked health websites measured by the website traffic monitoring service SimilarWeb (www.similarweb.com) and all sites from the Medical Library Association’s recommended websites for health information as of October 17, 2020. To identify website monitoring and data collection, each website URL was examined with Blacklight, an internet-based tool which tests how website use surveillance on their users. We present data for each website including use of ad tracking, use of third-party tracking and identification “cookies,” and availability of use data for Facebook and Google Analytics tracking.

Data was described using frequencies and measures of central tendencies. The ANOVA test and chi-square test were used to assess differences in means and frequencies between government, non-profit, and commercial health websites. All analyses were conducted in R version 4.0.2 (R Foundation).

RESULTS

The average number of ad trackers across included websites was 2.11 (SD 0.60), 7.15 (SD 6.26), and 15.84 (SD 10.29) for government, non-profit, and commercial health websites, respectively (p < 0.001). The average number of third-party cookies across included websites was 1.11 (SD 1.05), 10.85 (SD 11.9), and 25.08 (SD 25.45) for government, non-profit, and commercial health websites, respectively (p=0.003). Regarding websites informing Facebook of user activity, 0 (0.0%) government website, 10 (50.0%) non-profit websites, and 15 (60.0%) of commercial websites provided user data to Facebook. Regarding websites informing Google analytics of user activity, 6 (67.7%) government websites, 14 (70.0%) non-profit websites, and 16 (64.0%) of commercial websites provided user data to Google analytics. Average search results are recorded in Table 1, while individual results can be found in Table 2.

DISCUSSION

All health websites studied provide data to ad trackers and third-party cookies. Popular commercial websites used substantially more third-party cookies and ad-trackers than non-profit websites, which had more than the average government website.

Health-related websites often serve as a supplement, in which patients can find answers and further explanations for disease and treatment options. Data provided this way can be used to construct a personal profile of personal health information, and subsequently be provided or sold to other...
companies for the purposes of improving advertising targeting, as is the case in the current report in which this relationship with Google and Facebook is observed. The degree to which this information is used this way is unknown. However, this appears to be a common practice, and a recent publication of COVID-19-related websites found similar results related to the prevalence of third-party tracking. Greater clarity of how websites use collected health data may allow better identification of online resources that maximize the privacy of users, while also provided helpful medical information.

This study has several limitations. First, we used one software program to examine website tracking, and the algorithm for monitoring data-privacy may vary. Furthermore, we

| Reference       | Ad trackers | Third-party cookies | Information provided to Facebook | Information provided to Google Analytics |
|-----------------|-------------|---------------------|-----------------------------------|-----------------------------------------|
| **Government websites** |             |                     |                                   |                                         |
| Cdc.gov         | 7           | 3                   | *                                 | *                                       |
| Medlineplus.gov | 3           | 2                   |                                  |                                         |
| NIH.gov         | 2           | 2                   | Yes                              |                                        |
| NIA.NIH.gov     | 2           | 0                   | Yes                              |                                        |
| Health.gov      | 2           | 2                   | Yes                              |                                        |
| Cancer.gov      | 1           | 2                   |                                  |                                        |
| NIDDK.nih.gov   | 2           | 0                   | Yes                              |                                        |
| Healthfinder.gov| 2           | 2                   | Yes                              |                                        |
| Nei.nih.gov     | 2           | 0                   | Yes                              |                                        |
| **Commercial websites** |         |                     |                                   |                                         |
| Healthline.com  | 8           | 16                  |                                  | Yes                                     |
| Webmd.com       | 28          | 33                  | Yes                              |                                         |
| Medicalnewstoday.com | 13    | 18                  |                                  | Yes                                     |
| Cvs.com         | 3           | 3                   |                                  |                                         |
| Walgreens.com   | 30          | 54                  | Yes                              |                                        |
| Athenaehealth.com| 20          | 27                  | Yes                              |                                        |
| Myfitnesspal.com| 2           | 0                   |                                  | Yes                                     |
| Fitbit.com      | 16          | 16                  | Yes                              |                                         |
| Drugs.com       | 9           | 16                  | Yes                              |                                         |
| Psychologytoday.com | 6        | 4                   | Yes                              |                                         |
| Menshealth.com  | 17          | 15                  | Yes                              |                                        |
| Womenshealthmag.com | 17     | 17                  | Yes                              | Yes                                     |
| Msdmanuals.com  | 6           | 3                   | Yes                              |                                         |
| Medicine.net.com| 15          | 20                  | Yes                              |                                        |
| Medscape.com    | 20          | 16                  | Yes                              |                                        |
| Healthgrades.com| 10          | 4                   | Yes                              | Yes                                     |
| Babycenter.com  | 38          | 92                  | Yes                              |                                         |
| Practicefusion.com | 8        | 8                   | Yes                              | Yes                                     |
| Weightwatchers.com | 25      | 51                  | Yes                              |                                         |
| Yahoo.com/lifestyle/tagged/health |         |                     |                                   |                                         |
| Everydayhealth.com | 28       | 47                  | Yes                              | Yes                                     |
| Healthgrades.com | 10          | 4                   | Yes                              | Yes                                     |
| Mercola.com     | 5           | 4                   | Yes                              |                                         |
| Health.com      | 12          | 53                  |                                  | Yes                                     |
| Rxlist.com      | 12          | 20                  | Yes                              |                                         |
| **Non-profit**  |             |                     |                                   |                                         |
| Mayoclinic.org  | 15          | 31                  |                                  | Yes                                     |
| Clevelandclinic.org | 4          | 3                   | Yes                              |                                         |
| Kidshealth.org  | 6           | 11                  | Yes                              |                                         |
| Hivinsite.uesf.edu | 0      | 1                   | Yes                              |                                         |
| Familydoctor.org| 9           | 6                   | Yes                              | Yes                                     |
| Cancer.org      | 21          | 40                  | Yes                              | Yes                                     |
| Breastcancer.org| 10          | 4                   | Yes                              | Yes                                     |
| Cancercare.org  | 3           | 4                   | Yes                              |                                         |
| Foundationforwomenscancer.org | 3    | 17                  | Yes                              |                                         |
| Oncolink.org    | 7           | 17                  | Yes                              |                                         |
| Diabetes.org    | 13          | 19                  | Yes                              | Yes                                     |
| Joslin.org      | 2           | 0                   | Yes                              |                                        |
| Childrenwithdiabetes.com | 0 | 0                   |                                 |                                         |
| Afb.org         | 5           | 6                   | Yes                              |                                         |
| w-e-h.org       | 0           | 0                   | Yes                              |                                         |
| Heart.org       | 16          | 34                  | Yes                              | Yes                                     |
| Thebody.com     | 16          | 10                  | Yes                              | Yes                                     |
| Americanstroke.org | 1        | 4                   |                                 |                                         |
| Health.harvard.edu | 3       | 3                   |                                 | Yes                                     |
| Stroke.org      | 9           | 18                  | Yes                              | Yes                                     |

*Data abstracted on October 17, 2020, using Blacklight a publicly available service, with code made available on GitHub, which traverses individual websites and collects information such as third-party cookies, ad trackers, key logging, session recording, canvas fingerprinting, Facebook tracking, and Google Analytics use.*
limited our search to selected commonly trafficked websites. A broader sample may provide more granularity among health-related websites. Finally, it is not entirely known what the websites do with this information, or its overall benefits or harms.

We found that every category of health-related website examined provided information to ad trackers and created third-party cookies. Furthermore, providing data to Facebook and Google for targeted advertising was found to be relatively common, particularly among our sample of commercial and non-profit health-related websites. Searching for personal health information is not a private action and patients and providers must account for this when they search for health information online.

**Corresponding Author:** Alexander R. Zheutlin, MD, MS; Department of Internal Medicine, University of Utah School of Medicine, 30 N 1900 E, Room 4C104, Salt Lake City, UT 84132, USA (e-mail: alexander.zheutlin@hsc.utah.edu).

**Funding** Alexander Zheutlin is supported by the Utah Stimulating Access to Research in Residency Transition Scholar (StARRTS) Award Number 1R38HL143605-01.

**Declarations:**

**Conflict of Interest:** The authors declare no competing interests.

**REFERENCES**

1. Libert T. Privacy implications of health information seeking on the web. Commun ACM 2015;58(3):68-77.
2. Blacklight – The Markup. Accessed January 4, 2021. https://themarkup.org/blacklight
3. Mattu S, Sankin A. How We Built a Real-time Privacy Inspector. The Markup, September 22, 2020. Accessed September 29, 2020. https://themarkup.org/blacklight/2020/09/22/how-we-built-a-real-time-privacy-inspector#survey.
4. Savage M, Savage LC. Doctors Routinely Share Health Data Electronically Under HIPAA, and Sharing With Patients and Patients’ Third-Party Health Apps is Consistent: Interoperability and Privacy Analysis. J Med Internet Res 2020 Sep 2;22(9):e19818.
5. McCoy MS, Libert T, Buckler D, Grande DT, Friedman AB. Prevalence of Third-Party Tracking on COVID-19-Related Web Pages. JAMA. 2020 8:e2016178.
6. Murgia M, Hallow M. How top health websites are sharing sensitive data with advertisers. Financial Times. November 12, 2019. https://www.ft.com/content/0bf44d8e-022b-11ea-bc59-e49d2a136b8d. Accessed on October 17, 2020.

**Publisher’s Note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.