Background: Allergen immunotherapy (AIT) treatment for allergic rhinitis and asthma is used by 2.6 million Americans annually. Clinical and sterility testing studies identify no risk of contamination or infection from extracts prepared using recommended aseptic techniques, but regulatory concerns persist. Social media can be used to investigate rare adverse effects not captured by traditional studies.

Objective: We sought to investigate large social media databases for suggestion of AIT skin and soft tissue infection (SSTI) risk and compare this risk to a comparator procedure with a sterile pharmaceutical.

Methods: We analyzed US-restricted data from more than 10 common text-based social media platforms including Facebook, Twitter, and Reddit between 2012 and 2016. We used natural language processing (NLP) to identify posts related to AIT and, separately, influenza vaccination ( comparator procedure). NLP was followed by manual review to identify posts suggesting a possible SSTI associated with either AIT or influenza vaccination. SSTI frequencies with 95% CIs were compared.

Results: We identified 25,126 AIT posts, which were matched by social media platform to 25,126 influenza vaccination–related SSTI. NLP identified 2689 posts that required manual review, with 7 posts (0.02%; 95% CI, 0.005%-0.043%) indicative of possible AIT-related SSTI. NLP identified 4088 posts that required manual review, with 6 posts (0.02%; 95% CI, 0.005%-0.043%) indicative of possible influenza-related SSTI. NLP identified 2689 posts that required manual review, with 7 posts (0.03%; 95% CI, 0.007%-0.048%) indicative of possible influenza vaccination–related SSTI.

Conclusions: Social media data suggest that SSTI from AIT and influenza vaccination are equally rare events. Given that AIT’s SSTI risk appears comparable to the risk using a sterile pharmaceutical based on social media data, current aseptic technique procedures seem safe. (J Allergy Clin Immunol 2019;144:129-34.)

Key words: Informatics, NLP, infection, posts, allergen immunotherapy, preparation, aseptic, influenza, vaccination, safety, compounding, shot
extract large volumes of web content from health-related websites (eg, weightwatchers.com), social media (eg, Facebook, Twitter, and Reddit), forums (eg, forums.thebump.com), and blogs. Their database, freely available on their website, includes more than 2 billion users’ discussions, reflective of more than 40,000 medications and medical conditions.

Within Treato’s database, data are analyzed using a patented algorithm (patent #US 8,612,455 B2) that applies natural language processing (NLP) to online free-text discussions. NLP is a subfield of computer science that uses computers to process human (natural) language.17 NLP has been used in medicine and clinical research to process and analyze large amounts of free-text documents, such as clinical notes/reports and social media posts.18,19 Treato’s analysis engine combines a diverse range of medical ontologies, organized in a concept-based structure and coded similarly to the Unified Medical Language System (UMLS).20 and patient language vocabulary for concept extraction and analysis. Among the ontologies used is the Medical Dictionary for Regulatory Activities,21 which is used by the US Food and Drug Administration for adverse drug reaction term coding. These ontologies provide a set of terms to extract relevant concepts from free text and their internal hierarchical structure, which is used for grouping and aggregation. Synonymous terms (eg, “allergy shot” and “allergy desensitization”) are assigned the same ontology code, whereas ambiguous terms (eg, abbreviations such as “AIT” that could mean “allergen immunotherapy” or “advanced individual training”) have multiple candidate codes from which only 1 code is chosen using an ambiguity resolver component within Treato’s algorithm.

In this study, we considered social media posts generating only from the United States (as designated on the basis of internet protocol address) from January 1, 2012, through December 31, 2016.

**AIT social media data set.** After extracting 18 unique AIT synonym terms (Concept UMLS ID: C0162352) from the UMLS,20 such as “desensitization to allergens” or “desensitizing immunotherapy,” we expanded the AIT vocabulary by adding synonyms based on clinical expertise, common misspellings (eg, “allergy”), and lexical variants (eg, plurals “immunotherapies”). We manually reviewed a random subset of 1000 social media posts to verify that synonyms (eg, words such as “allergy,” “allergen,” and “shots”/“injections”) appeared together to determine whether any key synonyms were missed. This process notably did not identify any new synonyms. The final list of 33 query terms (Table I) was used to extract social media posts from Treato’s database, resulting in an AIT sample consisting of 25,126 social media posts.

**Influenza vaccination social media data set.** To generate a comparison sample of social media reports on influenza vaccination, we used Treato’s existing terminology list of influenza vaccination synonyms (eg, “flu shot/s,” “flu jab”; n = 428,832 posts). Given the high volume of influenza posts, we restricted the sample by matching posts to the AIT sample. To create this matched sample, we first randomly extracted the same number of posts for each one of the top 10 most common websites in the AIT sample; the remaining posts were randomly chosen from other websites with social media reports of influenza vaccination.

**Skin and soft tissue infections.** To identify potential SSTIs, we compiled a list of reactions and indicators suggestive of SSTI and other effects, based on clinical experience and literature review (Table I).22-24 SSTIs were defined as the presence of signs of bacterial or other epidermal or dermal infection related to the administration of AIT or influenza vaccination. We distinguished between reports that contained (1) mentions of potential SSTI (eg, impetigo, folliculitis, furuncles/carbuncles, and cellulitis); (2) mentions of other infection signs or symptoms (eg, words like “infections,” “fever,” “leukocytosis,” and “delirium”); and (3) mentions of potential SSTI treatment or service use (eg, antibiotics and expressions such as “hospitalization” or “emergency room [ER]”). We also investigated mentions of common and expected local reactions to AIT (eg, redness, swelling, and itch at the injection site). Searches were implemented using regular expressions in R statistical software (R Foundation for Statistical Programming, Vienna, Austria).25

Each post containing potential AIT or influenza vaccination–related SSTI, treatments, or service use was manually reviewed. All posts describing symptoms that were possibly an SSTI were reviewed by 2 internists and allergist-immunologists with a consensus achieved.

We compared possible SSTIs, and other symptoms, between AIT and influenza vaccine by determining frequencies with exact (ie, Clopper Pearson) 95% CIs in SAS version 9.4 (Cary, NC). We determined the precision of our methodology, the fraction of relevant instances among the retrieved instances.26

The research protocol was reviewed and deemed exempt by the Partners Human Research Committee.

**RESULTS**
We identified 25,126 AIT posts, which were matched by social media platform to 25,126 influenza vaccine posts (Table III). NLP identified 4088 AIT posts (16.2%) that required manual review. The most common term used to describe AIT in social media was “allergy shot” and its lexical variations (eg, plurals and misspellings). There were 6 posts (0.02%; 95% CI, 0.005%-0.043%) indicative of possible AIT-related SSTI (Table IV). Of the 6 possible SSTI posts, 4 included the terms “infection” or “cellulitis” (Fig 1, A). Two posts were included for possibly describing an SSTI: the first patient described symptoms most consistent with a local reaction from AIT, although because antibiotics were prescribed, investigators included it as a possible SSTI. The other post described a high fever, a hot itchy swollen arm, and cough, included as a possible SSTI, but may have been an allergic reaction to AIT or a respiratory infection.

NLP identified 2689 influenza posts (10.7% of total matched posts) that required manual review, with 7 posts (0.03%; 95% CI, 0.007%-0.048%) indicative of possible influenza vaccination–related SSTI. These 7 posts described SSTIs using terms including “infection” (n = 4) or “cellulitis” (n = 3; Fig 1, B).
Slightly more than 1% of posts (n = 279) described other local or systemic reactions to AIT, commonly local swelling (0.32%; 95% CI, 0.252%-0.392%), hospital service use (0.29%; 95% CI, 0.229%-0.362%), and skin redness (0.24%; 95% CI, 0.182%-0.304%; Table IV). Other infection indicators were uncommonly described, largely fever (0.02% [95% CI, 0.002%-0.037%] for AIT and 0.35% [95% CI, 0.274%-0.419%] for influenza vaccination). All related side effects were identified in 1.13% (95% CI, 1.003%-1.265%) of AIT posts, compared with 1.35% (95% CI, 1.203%-1.488%) of influenza posts.

Frequencies of AIT signs and symptoms of interest were all comparable to those observed with influenza vaccination, with the exceptions of itch, which was more common with AIT (0.22% [95% CI, 0.161%-0.277%] vs 0.04% [95% CI, 0.018%-0.070%] for influenza vaccination), and fever, which was more common with influenza vaccination (0.35% [95% CI, 0.274%-0.419%] vs 0.02% [95% CI, 0.002%-0.037%] for AIT).

Our NLP approach achieved 7% precision (285 of 4088 potential AIT-related SSTI and other events identified by NLP were confirmed) in identifying AIT-related side effects and 12.6% precision (338 of 2689 potential influenza vaccination–related SSTI and other events identified by NLP were confirmed) in identifying influenza vaccination–related side effects reported in the social media. Irrelevant posts included posts about other allergic conditions (eg, food allergies and atopic dermatitis) and posts mentioning AIT/influenza vaccination in other contexts (eg, medical personnel described unwillingness to undergo mandatory influenza vaccination).

**DISCUSSION**

In this study, we analyzed large US-restricted public text–based social media data from sites including Facebook, Twitter, and Reddit, with the aim of discovering rare adverse events less likely to be identified by traditional research methods. We identified mentions of symptoms consistent with SSTI that the poster associated with AIT, a procedure prepared using aseptic technique, or influenza administration, a procedure using a sterile pharmaceutical. In total, only 6 of 25,126 posts related to AIT described a possible AIT-related SSTI, whereas only 7 of 25,126 posts related to influenza vaccination described a possible vaccination-related SSTI. Slightly more than 1% of all posts described other local or systemic reactions from each procedure.

We identified a small number of posts describing a possible AIT-related SSTI using informatics methods (0.02%; 95% CI, 0.005%-0.043%). This study confirms findings to date from traditional studies that have found a negligible risk of infection for patients receiving AIT.5,10-12,14 Specifically, 1 previous retrospective study of 272 patients on AIT did not identify a single patient with SSTI from AIT, nor any patient with a fever, discharge from the injection site, or cellulitis.14 Our previous study used clinical use data to identify potential AIT-related infections, and of 136,322 separate AIT administration episodes over a 10-year period at 2 academic medical centers, no infectious complications of AIT were identified. Previous data relied on patients reporting the symptoms to their clinician in follow-up or seeking medical care for adverse symptoms within the same health system. Thus, it was plausible that infections existed but...
were not being identified through these traditional research methods. As such, we used advanced informatics methods on large amounts of social media data over 5 years to detect rare signals not easily identifiable in clinical studies. Our methods identified a few reports of possible SSTIs from AIT and influenza vaccine, data that support that both procedures pose no meaningful infection risk.

The frequency with which we observed AIT-related posts that suggest a potential associated SSTI and the frequency with which we identified influenza vaccination–related posts that suggest a potential associated SSTI were almost identical, with overlapping CIs. This consistency suggests that the observed frequencies are comparable despite their preparation differences and that aseptic technique should not be a concern for treating allergists, patients on AIT, or regulatory bodies. Given other potential consequences predicted should substantial changes to AIT preparation be recommended by a regulatory body, including a contracted AIT supply, as well as consequences from changing the beyond-use date (eg, increased allergic reactions), allergen extract preparation need not be changed under a guise of improving patient safety without thorough consideration of these data.

Surprisingly, slightly fewer local or systemic reactions were reported in AIT-related posts than were reported in influenza vaccination–related posts. As with any treatment, both AIT and influenza vaccinations may result in unintended side effects; in sum, side effects investigated comprised just 1.13% of AIT posts and 1.35% of influenza vaccination posts. This finding supports the large body of existing literature on the tolerability of AIT for the treatment of patients with allergic rhinitis and allergic asthma. Furthermore, social media posts described swelling, redness, and exudate with similar frequency for both injections. Notably, itch was the only symptom significantly more common for AIT than for influenza vaccination, and fever was the only symptom significantly more common for influenza vaccination than for AIT.

Our study has notable limitations. First, it is possible that adverse events from an injection might be underreported on social media, even on health-related sites. However, data suggest that adverse drug reactions reported on social media generally align with those reported in the electronic health record, and rare events are more prevalent in social media than in the electronic health record. We could not study all social media data and could not study private posts. Nonetheless, we used large amounts of social media data from numerous social media sites that included the most common text-based platforms Facebook, Twitter, and Reddit. Because frequencies of US-restricted social media posts do not correlate with epidemiology estimates of prevalence and incidence, we used a matching process by which we were able to compare AIT SSTI–observed frequencies to influenza vaccine SSTI–observed frequencies. Although a new approach, our methodology was shown to have validity given that we were able to identify the expected findings of itch being more common for AIT and fever being more common in influenza vaccination. Still, there is no perfect AIT comparator; influenza vaccination posts contained messages about involuntary vaccination and general antivaccination sentiment. Despite using random sampling for the comparator group posts, the sample represented less than 10% of influenza posts, and may not have been representative. Although NLP was used to mine social media data and find potential SSTI–related posts, it is possible that the search terms and algorithms missed posts that were actual matches. However, our NLP method used a broad set of key words designed to be inclusive and capture all possible relevant mentions, and reassuringly, the algorithms detected many possibly relevant cases and cases where the text was suggestive of an infection in both AIT and influenza vaccination data sets. It is possible that more cases may have been identified with a manual review; however, manual review of 52,000 posts would not be feasible and, practically, could be subject to more error (eg, natural human error) as much as NLP algorithms. Although we aimed to study conventional AIT, we may have captured another injection for allergies that the patient interpreted as an “allergy shot” (eg, neutralization-provocation therapy and Rinkel technique). Although the most prevalent method to prepare AIT in the United States is aseptic technique, the preparation method of AIT could not be discerned from social media posts and other preparation methods (eg, sterile technique) might have been possible. Although we reviewed several random subsamples of the social media posts that did not include SSTI–related keywords (and did find additional SSTI–related posts), we cannot calculate NLP recall in a reliable manner. In addition, social media posters are likely younger patients more familiar with technology. Finally, although it is possible that SSTIs were reported by the same person on different websites under different user names, reviewers considered the positive events distinct.
Mining of social media data suggests that suspected infections from AIT, prepared with aseptic technique, are rare, as rare as suspected infections from influenza vaccination, a sterile pharmaceutical. Current AIT preparation guidelines do not appear to pose an infectious risk to patients, as previously determined in translational and clinical studies, and now
confirmed in an informatics study. AIT safety efforts should focus on operational improvements to prevent, diagnose, and treat the allergic reactions that predictably complicate the administration of this effective treatment for allergic rhinitis and asthma.

Clinical implications: Current preparation guidelines for allergen immunotherapy do not pose an infectious risk to patients.

REFERENCES