Open Forum Infectious Diseases

MAJOR ARTICLE

Text-Based Identification of Herpes Zoster Ophthalmicus With Ocular Involvement in the Electronic Health Record: A Population-Based Study

Chengyi Zheng,[1] Lina S. Sy,[1] Hilary Tanenbaum,[1] Yun Tian,[1] Yi Luo,[1] Bradley Ackerson,[2] and Hung Fu Tseng[1]

[1]Department of Research & Evaluation, Kaiser Permanente Southern California, Pasadena, California, USA, [2]South Bay Medical Center, Kaiser Permanente Southern California, Harbor City, California, USA

Background. Diagnosis codes are inadequate for accurately identifying herpes zoster ophthalmicus (HZO). Manual review of medical records is expensive and time-consuming, resulting in a lack of population-based data on HZO.

Methods. We conducted a retrospective cohort study, including 87 673 patients aged ≥50 years who had a new HZ diagnosis and associated antiviral prescription between 2010 and 2018. We developed and validated an automated natural language processing (NLP) algorithm to identify HZO with ocular involvement (ocular HZO). We compared the characteristics of NLP-identified ocular HZO, nonocular HZO, and non-HZO cases among HZ patients and identified the factors associated with ocular HZO among HZ patients.

Results. The NLP algorithm achieved 94.9% sensitivity and 94.2% specificity in identifying ocular HZO cases. Among 87 673 incident HZ cases, the proportion identified as ocular HZO was 9.0% (n = 7853) by NLP and 2.3% (n = 1988) by *International Classification of Diseases* codes. In adjusted analyses, older age and male sex were associated with an increased risk of ocular HZO; Hispanic and black race/ethnicity each were associated with a lower risk of ocular HZO compared with non-Hispanic white.

Conclusions. The NLP algorithm achieved high accuracy and can be used in large population-based studies to identify ocular HZO, avoiding labor-intensive chart review. Age, sex, and race were strongly associated with ocular HZO among HZ patients. We should consider these risk factors when planning for zoster vaccination.

Keywords. herpes zoster ophthalmicus; natural language processing; retrospective cohort study; shingles.

After primary infection, varicella zoster virus (VZV) resides in dorsal root ganglia and can reactivate later in life when immune function weakens [1, 2]. Once reactivated, VZV travels along sensory nerves and reaches the skin, manifesting as herpes zoster ([HZ] shingles) [3]. Herpes zoster is characterized by its distinctive unilateral dermatomal distribution of vesicular rash accompanied by pain [1]. When HZ affects the ophthalmic division (V1) of the fifth cranial nerve (trigeminal nerve), it is defined as HZ ophthalmicus (HZO) [4–6]. Although the term "ophthalmicus" is used, HZO does not necessarily affect the eye. The 3 branches of the ophthalmic division of the trigeminal nerve cover a large area of the upper face and head, including the scalp, forehead, eyelid, eye, and nose [2]. When the eye or its surrounding area is affected (called "ocular HZO"), complications with severe sequelae may occur, including chronic ocular inflammation, loss of vision, lingering pain [6, 7], and cerebrovascular events [8].

A previous study estimated that the lifetime risk of HZO is approximately 1% [2]. However, the incidence of HZO is poorly quantified due to methodological challenges. Both HZO and ocular HZO are substantially underdetected by diagnosis codes [9–11]. Manual chart review is a potentially more sensitive method used to identify HZO and ocular HZO patients. However, it would be labor-intensive to perform chart reviews to obtain population-wide estimates [7, 12–15]. Developing a robust method to identify ocular HZO is crucial for understanding factors associated with its occurrence, studying its overall burden, and assessing the impact of possible interventions in preventing ocular HZO.

Building on our prior work using natural language processing (NLP) to identify HZO, we refined the algorithm and expanded it to differentiate between ocular HZO and nonocular HZO cases [11]. We then applied the algorithm to a large cohort of HZ patients to examine the characteristics of patients with ocular HZO.

Received 6 October 2020; editorial decision 23 December 2020; accepted 29 December 2020.

Correspondence: Chengyi Zheng, PhD, MS, Department of Research and Evaluation, Kaiser Permanente Southern California, 100 S. Los Robles Ave., 2nd Floor, Pasadena, CA 91101 (chengyi.x.zheng@kp.org).

Open Forum Infectious Diseases®2021

DOI: 10.1093/ofid/ofaa652

METHODS

Study Setting

This study was conducted at Kaiser Permanente Southern California (KPSC), an integrated healthcare organization with 15 hospitals, 234 medical offices, and more than 7600 physicians. Kaiser Permanente Southern California provides prepaid comprehensive healthcare to 4.7 million racially and socioeconomically diverse members [16]. Members can receive medical care either at KPSC-owned or contracted facilities. Kaiser Permanente Southern California maintains extensive electronic medical records (EMRs) that store both structured data (including sociodemographics, diagnoses, laboratory tests, pharmacy utilization, immunization records, membership history, death, and utilization of outpatient, emergency department, and inpatient services) and unstructured, free text data such as clinical notes, pathology reports, and radiology reports.

Patient Consent Statement

The KPSC Institutional Review Board approved this study with a waiver of written patient informed consent.

Training and Validation Datasets

Our goal for NLP algorithm development was to distinguish ocular HZO from nonocular HZO, as well as from HZ at nonophthalmicus locations. We first refined our previously validated HZO NLP algorithm [11] to increase the sensitivity for identifying HZO cases. We then applied the improved HZO NLP algorithm to an existing study cohort [14] of incident HZ cases in 2011–2015 identified by *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) codes 053.xx and ICD-10-CM codes B02.xx among immunocompetent KPSC members aged ≥50 years. The HZO NLP algorithm was applied to all clinical notes from 7 days before the incident HZ diagnosis date to 30 days after the diagnosis date. We chose this interval based on the typical timeframe for the progression of HZO [17]. Clinical notes included outpatient (office, urgent care), emergency department, inpatient, and telehealth (telephone, email, virtual) visit notes.

We identified a random sample of 500 potential NLP-identified HZO patients for training and validating the ocular HZO NLP algorithm. We used a sample of HZO patients instead of HZ patients to enrich the sample with potential cases of ocular HZO. The 500 patients were randomly split into a training dataset (n = 300) for developing the new NLP algorithm and a validation dataset (n = 200) for evaluating the performance of the NLP.

The reference standard for the development and evaluation of the NLP algorithm was manual chart review. Four trained research associates were assigned into 2 pairs, with each pair assigned to review the same 250 cases independently. An infectious disease specialist (B.A.) helped to resolve discrepancies between the results of the double reviews.

We conducted chart reviews on these patients to categorize them into 1 of 4 groups: ocular HZO, nonocular HZO, HZ at nonophthalmicus locations (non-HZO HZ), and non-HZ patients. Herpes zoster was defined as HZ confirmed in the chart note or presence of HZ-related signs or symptoms in the absence of non-HZ causes. Ocular HZO was defined as HZ-related signs or symptoms in the eye, eye adnexa, upper eyelid, or periorbital area [6, 17–19]. Nonocular HZO was defined as HZ-related signs or symptoms in the frontal scalp, forehead, eyebrow, glabella, and nose. Non-HZO HZ was defined as HZ at locations other than the HZO locations listed above.

Algorithm Development

We developed a rule-based NLP algorithm using the training dataset. We iteratively refined the algorithm to match the chart review reference standard. We described the detailed NLP processes previously [11, 20]. First, we converted the clinical notes extracted from the EMR system into formats suitable for the NLP search. A preprocessing step removed ill-formatted text and detected sections and sentence boundaries. We created terminologies to capture HZ-related information [3, 5, 21]. The NLP search was performed at different levels: section, sentence, and neighboring sentences [11]. We used the relationship detection algorithm to identify associated entities such as signs or symptoms and their associated dermatomal locations. We used the negation and temporal relationship detection algorithms to identify negated, uncertain, historical, and future statements. We developed a postprocessing step to integrate and finalize the results based on the case definitions described above.

Algorithm Validation Using the Validation Dataset

The performance of the NLP algorithm was evaluated against the reference standard validation dataset. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and positive and negative likelihood ratios were calculated. To account for the artificially higher prevalence of ocular HZO in the validation dataset, we also reported adjusted performance metrics based on the estimated ocular HZO prevalence of the HZ source population.

Algorithm Application

We applied the validated ocular HZO NLP algorithm to a cohort of KPSC members aged ≥50 years with a new HZ ICD diagnosis code between 2010 and 2018 and antiviral medication (acyclovir, famciclovir, or valacyclovir) within ±7 days of diagnosis [14]. We excluded patients with an HZ diagnosis in the prior year. The first diagnosis date during the period was defined as the index date. Patients were required to have at least 6 months of membership before the index date so that comorbidities and previous healthcare utilization could be ascertained.

Identification of Ocular Herpes Zoster Ophthalmicus by Natural Language Processing Versus International Classification of Diseases Codes

Among the cohort above with incident HZ, we calculated the frequencies and percentages of ocular HZO cases identified by both NLP and ICD codes, by NLP only, by ICD codes only, or by neither. The ocular HZO ICD-9 codes are 053.2x, and the ICD-10 codes are B02.3x.

Statistical Analysis

To compare the characteristics of patients with ocular HZO, nonocular HZO, and non-HZO HZ, we calculated distributions of demographic variables, prior healthcare utilization, and comorbidities previously found to be associated with HZ or HZO [13, 22]. We calculated *P* values using the χ^2 test for all categorical variables. We used robust Poisson models to estimate the factors associated with the proportion of ocular HZO among all HZ patients. We calculated the prevalence ratio (PR) and 95% confidence intervals (CIs) associated with age, sex, race/ethnicity, body mass index (BMI), smoking history, marital status, healthcare utilization, and comorbidities. The multivariable model was constructed using stepwise regression. All statistical analyses were conducted using SAS version 9.4 (SAS Institute, Cary, NC).

RESULTS

Training and Validation Datasets

For the training and validation datasets, the assigned pairs of chart reviewers were largely in agreement in classifying HZO and ocular HZO cases. The kappa scores between the 2 pairs of reviewers were 0.89 (0.83–0.96) and 0.85 (0.78–0.91). Herpes zoster ophthalmicus was found among 357 (71.4%) of the 500 chart-reviewed records, 243 (68.1%) of which were classified as having ocular HZO.

Algorithm Validation Using the Validation Dataset

The validation results of the ocular HZO NLP algorithm are summarized in Table 1. The NLP algorithm achieved 94.9% sensitivity, 94.2% specificity, 93.9% PPV, and 95.1% NPV for ocular HZO identification (Table 2). The positive and negative likelihood ratios were 16.3 and 0.05, respectively.

Table 1. Cross Table for Comparing NLP Results and Chart Review Reference Standard for Identification of Ocular HZO Cases

	Reference Standard		
NLP Result	Positive	Negative	Total
NLP Positive	92	6	98
NLP Negative	5	97	102
Total	97	103	200

Abbreviations: HZO, herpes zoster ophthalmicus, NLP, natural language processing.

Table 2. Accuracy Measurements of NLP in identifying ocular HZO cases

Evaluation Statistics % (95% CI)	Validation Dataset	Adjusted for Population Prevalence
Sensitivity	94.9 (88.4–98.3)	-
Specificity	94.2 (87.8–97.8)	-
Positive Likelihood Ratio	16.3 (7.5–35.4)	-
Negative Likelihood Ratio	0.05 (0.02–0.13)	-
Positive Predictive Value[a]	93.9 (87.6–97.1)	61.7 (42.5–77.8)
Negative Predictive Value[a]	95.1 (89.2–97.9)	99.5 (98.7–99.8)

Abbreviations: CI, confidence interval; HZO, herpes zoster ophthalmicus, NLP, natural language processing.

[a]Positive and negative predictive values are dependent on disease prevalence. Based on this study, the prevalence of ocular HZO disease among HZ cases was 9%.

Table 3. Comparison of NLP to ICD Codes for Identification of Ocular HZO Cases

		ICD Codes		
Ocular HZO		Yes	No	Total
NLP	Yes	1910 (2.2%)	5943 (6.8%)	7853 (9.0%)
	No	78 (0.1%)	79 742 (91.0%)	79 820 (91.0%)
	Total	1988 (2.3%)	85 685 (97.7%)	87 673 (100.0%)

Abbreviations: HZO, herpes zoster ophthalmicus; ICD, *International Classification of Diseases*; NLP, natural language processing.

NOTE: This table lists the numbers and percentages of ocular HZO cases identified by NLP and/or ICD codes. The total number of cases (n = 87 673) included all patients aged ≥50 years with incident HZ identified using ICD codes and medications. One patient could have multiple HZ-related ICD codes for the same encounter. The case was marked as Yes for the ICD codes column if at least one of the HZ codes was an ocular HZO ICD code (053.2x, B02.3x).

Identification of Ocular Herpes Zoster Ophthalmicus by Natural Language Processing Versus International Classification of Diseases Codes

There were 87 673 KPSC members aged ≥50 years with incident HZ identified using ICD codes and medications. For the application of the ocular HZO NLP algorithm, NLP processed 1 053 241 clinical notes, consisting of 224 204 887 words. Among 87 673 incident HZ cases, the proportion identified as ocular HZO was 9.0% (n = 7853) by NLP and 2.3% (n = 1988) by ICD codes (Table 3). Among the ocular HZO cases coded by ICD, NLP confirmed 96.1% of them as ocular HZO cases. The NLP confirmation rates were similar for the ICD-9-CM and ICD-10-CM codes (data not shown in Table 3). There were 5943 ocular HZO cases (6.8% of incident HZ cases) identified by NLP that were not identified by ICD codes.

Characteristics of Ocular Herpes Zoster Ophthalmicus (HZO), Nonocular HZO, and Non-HZO HZ Patients

The characteristics of the study population are presented in Table 4. Among the 87 673 HZ cases, 7853 (9.0%) were classified as ocular HZO and 3095 (3.5%) as nonocular HZO. The proportion of ocular HZO among HZO cases was 72%.

Compared with nonocular HZO patients, ocular HZO patients were older (mean age 66.8 vs 64.8 years), had more comorbidities, and had higher healthcare utilization in the

Table 4. Baseline Characteristics of Ocular HZO, Nonocular HZO, and Non-HZO Herpes Zoster Patients

Characteristics	HZO[a]		Non-HZO HZ[b]	*P* Value[c]
	Ocular	Nonocular		
Total	7853 (9.0)	3095 (3.5)	76 725 (87.5)	
Age in Years				<.001
50–59	2386 (7.9)	1117 (3.7)	26 653 (88.4)	
60–69	2519 (8.7)	1012 (3.5)	25 305 (87.8)	
70–79	1819 (9.7)	660 (3.5)	16 253 (86.8)	
≥80	1129 (11.3)	306 (3.1)	8514 (85.6)	
Mean (SD)	66.8 (10.7)	64.8 (10.1)	65.3 (10.3)	
Sex				<.001
Female	4611 (8.6)	1771 (3.3)	47 010 (88)	
Male	3242 (9.5)	1324 (3.9)	29 715 (86.7)	
Race/Ethnicity				<.001
Non-Hispanic white	4160 (10.3)	1753 (4.3)	34 474 (85.4)	
Hispanic	1946 (6.8)	744 (2.6)	25 751 (90.5)	
Asian	612 (9)	176 (2.6)	5985 (88.4)	
Black	978 (9.6)	341 (3.3)	8902 (87.1)	
Other[d]	157 (8.5)	81 (4.4)	1613 (87.1)	
BMI, kg/m^2				<.001
<18.5	68 (10.9)	14 (2.2)	541 (86.8)	
18.5–24.9	1877 (9.6)	698 (3.6)	17 015 (86.9)	
25.0–29.9	2363 (8.9)	934 (3.5)	23 228 (87.6)	
30.0–34.9	1369 (8.8)	502 (3.2)	13 755 (88)	
≥35.0	911 (9.2)	353 (3.6)	8629 (87.2)	
Missing	1265 (8.2)	594 (3.9)	13 557 (87.9)	
Smoking History				<.001
Current/passive smoker	1720 (9.3)	668 (3.6)	16 090 (87.1)	
Former smoker	885 (9.3)	319 (3.3)	8320 (87.4)	
Never smoker	3870 (9.1)	1467 (3.4)	37 419 (87.5)	
Unknown	1378 (8.1)	641 (3.8)	14 896 (88.1)	
Marital Status				.34
Married	4151 (8.9)	1643 (3.5)	40 796 (87.6)	
Never married	987 (8.9)	384 (3.5)	9674 (87.6)	
Previously married[e]	1220 (9.4)	434 (3.3)	11 331 (87.3)	
Other/Missing	1495 (8.8)	634 (3.7)	14 924 (87.5)	
Outpatient Visits in 6 Months Before Herpes Zoster Diagnosis Date				<.001
0	836 (8.2)	375 (3.7)	8978 (88.1)	
1–5	4005 (8.8)	1632 (3.6)	40 018 (87.7)	
≥6	3012 (9.5)	1088 (3.4)	27 729 (87.1)	
Mean (SD)	6.3 (7.6)	5.8 (7.4)	5.9 (7.0)	
Emergency department visits in 6 Months Before Herpes Zoster Diagnosis Date				<.001
0	6495 (8.8)	2675 (3.6)	64 532 (87.6)	
≥1	1358 (9.7)	420 (3)	12 193 (87.3)	
Mean (SD)	0.2 (0.7)	0.2 (0.6)	0.2 (0.7)	
Hospitalizations in 6 Months Before Herpes Zoster Diagnosis Date				.05
0	7383 (8.9)	2946 (3.6)	72 279 (87.5)	
≥1	470 (9.3)	149 (2.9)	4446 (87.8)	
Mean (SD)	0.1 (0.5)	0.1 (0.4)	0.1 (0.4)	
Comorbidity				
Allergic rhinitis	333 (9.3)	111 (3.1)	3141 (87.6)	.29
Asthma	605 (9.6)	191 (3)	5510 (87.4)	.02
Atopic dermatitis	34 (10.1)	10 (3)	292 (86.9)	.67
Cancer (solid)	434 (9.3)	152 (3.3)	4088 (87.5)	.43
Cancer (nonsolid)	154 (10.1)	60 (3.9)	1309 (85.9)	.18
Cerebrovascular disease	275 (11.2)	77 (3.1)	2101 (85.7)	<.001
Chronic kidney disease	1111 (10.1)	337 (3.1)	9578 (86.9)	<.001
Chronic obstructive pulmonary disease	348 (10.5)	92 (2.8)	2887 (86.8)	<.001
Coronary artery disease	674 (10.9)	195 (3.2)	5287 (85.9)	<.001

Table 4. Continued

Characteristics	HZO[a]		Non-HZO HZ[b]	*P* Value[c]
	Ocular	Nonocular		
Dementia	151 (10.9)	47 (3.4)	1186 (85.7)	.04
Depression	892 (8.9)	376 (3.8)	8756 (87.4)	.44
Diabetes	1664 (9.1)	573 (3.1)	15 962 (87.7)	.01
Hypertension	3370 (9.6)	1165 (3.3)	30 506 (87.1)	<.001
Inflammatory bowel disease	153 (9.2)	61 (3.7)	1442 (87.1)	.86
Obstructive sleep apnea	297 (9.7)	111 (3.6)	2645 (86.6)	.29
Peripheral vascular disorder	140 (10.6)	35 (2.7)	1144 (86.7)	.03
Rheumatoid arthritis	141 (8.2)	52 (3)	1520 (88.7)	.28
Systemic lupus erythematosus	28 (8.4)	9 (2.7)	298 (89)	.64

Abbreviations: BMI,body mass index; HZO,herpes zoster ophthalmicus; SD, standard deviation.
NOTE: Values are mean ± standard deviation (SD) or n (%), unless otherwise indicated.
[a]HZO cases were identified by natural language processing (NLP).
[b]Non-HZO HZ cases were coded HZ cases not identified as HZO by NLP.
[c]*P* value was calculated using χ^2 test.
[d]"Other" includes Native American, mixed, or unknown race/ethnicity.
[e]"Previously married" includes divorced, separated, or widowed.

6 months before the HZ diagnosis date. The proportion of ocular HZO patients among all HZ patients increased with age. In contrast, the proportion of nonocular HZO patients among all HZ patients did not increase with age (Figure 1).

Risk Factors for Ocular Herpes Zoster Ophthalmicus

Table 5 presents the adjusted PR for the factors associated with the prevalence of ocular HZO compared with all other HZ patients. Overall, older age was associated with ocular HZO. Patients with age ≥80 years had an adjusted PR of 1.31

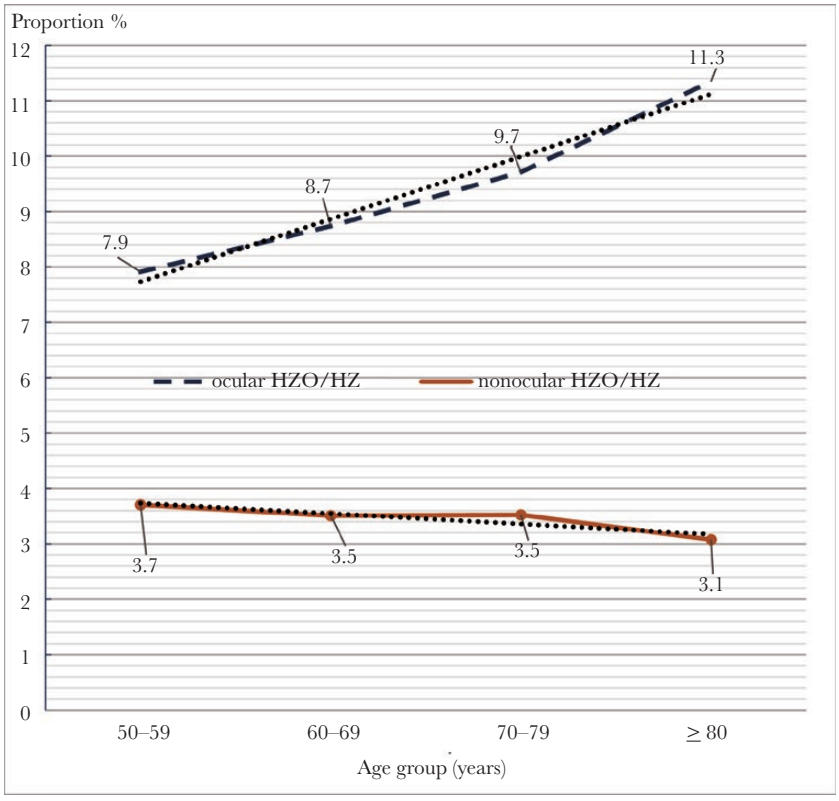

Figure 1. Proportion of ocular herpes zoster ophthalmicus (HZO) and nonocular HZO among HZ patients by age group. The lines show the proportion of ocular HZO and nonocular HZO cases among all HZ cases by age group. The accompanying linear trendlines demonstrate an approximately linear increase of ocular HZO by age and a slight downward trend for nonocular HZO.

Table 5. Prevalence Ratio of Ocular HZO vs All Other Herpes Zoster Cases

Characteristic	Adjusted Prevalence Ratio (95% CI) Ocular HZO vs All Other Herpes Zoster
Age in Years	
50–59	1 [Reference]
60–69	1.07 (1.01–1.13)
70–79	1.16 (1.09–1.23)
≥80	1.31 (1.22–1.41)
Sex	
Male vs female	1.10 (1.05–1.15)
Race/Ethnicity	
White	1 [Reference]
Hispanic	0.68 (0.64–0.72)
Black	0.89 (0.82–0.97)
Asian/Pacific Islander	0.94 (0.88–1.01)
Other	0.87 (0.74–1.02)

Abbreviations: CI, confidence interval; HZO, herpes zoster ophthalmicus.

NOTE: The Poisson regression model was selected based on stepwise variable selection, with entry-level $P = .05$ and stay-level $P = .1$. All variables listed in Table 4 were considered in the multivariate model. Only variables with statistically significant adjusted prevalence ratios in the final model are displayed.

compared with those 50–59 years old. The prevalence of ocular HZO was higher among males than females (adjusted PR, 1.10; 95% CI, 1.05–1.15) and lower among both Hispanic and black race/ethnicity compared with non-Hispanic whites (adjusted PR = 0.68, 95% CI = 0.64–0.72 and adjusted PR = 0.89, 95% CI = 0.82–0.97, respectively).

DISCUSSION

Our study demonstrates the ability of a computer-based NLP algorithm applied to vast quantities of free-text EMR to identify ocular HZO cases with high accuracy. In contrast to claims-based data in which limited clinical information is coded into structured data, the EMR contains rich, detailed free-text clinical information that can be searched using NLP. Using the NLP algorithm, we significantly improved our ability to identify ocular HZO patients compared with using ICD codes only. We successfully applied the NLP algorithm to more than 80 000 HZ patients to identify the characteristics associated with the occurrence of ocular HZO. The proportion of ocular HZO among HZ cases was 9% (7853 of 87 673), similar to the proportion (9%, 184 of 2035) reported in a previous chart review study conducted in Olmsted County, Minnesota, despite different methodologies [7]. The proportion of ocular HZO among HZO cases was 72%, similar to the proportion (69%, 62 of 90) reported in a previous chart review study performed in the Veterans Administration Healthcare System [23].

Epidemiological studies of ocular HZO rely heavily on ICD codes to identify patients. However, code-based approaches suffer from low sensitivity due to undercoding [9, 15]. The problem of undercoding ocular HZO is aggravated by the fact that many ocular HZO patients without ocular symptoms often do not visit an ophthalmologist. Therefore, studies requiring ocular HZO confirmation by an ophthalmologist would not capture the true number of ocular HZO cases [11], underestimating the incidence of ocular HZO and the proportion of ocular HZO among all HZ patients. Our NLP approach not only identified HZO with high accuracy [11], but it also distinguished ocular HZO from HZO without ocular involvement, a significant improvement compared with previous ocular HZO studies based on ICD codes [7, 9, 10, 15, 24, 25]. The number of ocular HZO cases identified by the NLP algorithm was significantly higher than the number identified by the ICD codes alone.

Historically, chart review of a large number of HZ cases, which would be resource-intensive, was the only accurate way to identify and classify HZO and ocular HZO [4, 7, 13]. As a result, even the largest previous chart-based studies of ocular HZO included a total of a few hundred patients evaluated for ocular HZO [7, 23, 25]. In addition, past chart-based studies used time frames of 2 weeks [7] and 30 days [17] to identify ocular HZO. With our NLP algorithm, the additional computer processing time for those additional 2 weeks was minimal, whereas with manual chart review, the increased effort needed would be substantial. The NLP approach solves the problem of chart review resource limitations for population-based studies of HZO.

To our knowledge, this is the largest study that investigates the differences between ocular HZO and nonocular HZO patients. One of the major novel findings of the current study is that the proportion of ocular HZO increased with age; however, the proportion of nonocular HZO did not. The ratio of ocular HZO versus nonocular HZO cases increased by more than 70% from age 50 to 69 to age ≥80. This suggests that ocular HZO is related to the age-dependent waning of immunity. Because immunity against VZV wanes with age, ocular HZO becomes more common. This age-related trend was not observed for nonocular HZO because its proportion remained constant with age.

Besides age, ocular HZO was also associated with sex and race/ethnicity. We examined an extensive list of comorbidities that were reported to be associated with HZ or HZO [26–28]. After adjustment, none were found to be associated with ocular HZO. It is worth noting that we examined the risk factors of ocular HZO among HZ patients, not risk factors for HZ. In our study, we found that older age and white versus Hispanic or black race/ethnicity was associated with higher risk of ocular HZO among HZ patients, similar to the risk factors for HZ [29]. However, we found that male sex was associated with higher risk of ocular HZO among HZ patients, in contrast to previous reports of women being at increased risk of HZ compared with men [22].

This study had some potential limitations. First, we limited our study population to patients who had a new HZ diagnosis and associated antiviral prescription because HZ coding is less

reliable for patients without an antiviral prescription [14]. The HZ population with or without an antiviral prescription may be different. In addition, not all coded HZ cases included in the study are likely to be true HZ. The actual proportion of ocular HZO among all HZ patients could therefore be higher. Furthermore, due to the lack of availability of certain variables in the EMR system, we did not include other risk factors related to HZ, such as family history of HZ. In the future, NLP could be used to help overcome these limitations by increasing the accuracy of coded HZ, detecting miscoded HZ cases in the general population, and identifying additional risk factors. Finally, zoster vaccination was outside the scope of this study, because our focus was on (1) using an NLP algorithm to identify ocular HZO among HZ patients and (2) determining risk factors for ocular HZO. Future studies could examine the impact of zoster vaccination on HZ, HZO, and ocular HZO.

CONCLUSIONS

In summary, we developed an automated algorithm to identify ocular HZO patients with high accuracy. Among HZ cases, the proportion identified as ocular HZO was 9.0%, much higher than the 2.3% from the ICD codes. Our findings suggest that ocular HZO is associated with age, sex, and race/ethnicity. This approach allows us to perform large-scale population-based studies of ocular HZO that would otherwise be infeasible, enabling us to study the risk factors that lead to this severe form of HZ and to study the prognostic factors that affect the risk of developing unwanted sequelae after the episode, such as blindness or cerebrovascular events.

Acknowledgments

We thank Anna Lawless, Cindy Agus, Nehaa Khadka, Jonathan Arguello, and Raul Calderon for their efforts in the chart review.

Financial support. This work was supported by Kaiser Permanente Southern California internal research funds.

Potential conflicts of interest. All authors received research support from GlaxoSmithKline for work unrelated to this study. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest.

References

1. Cohen JI. Clinical practice: herpes zoster. N Engl J Med **2013**; 369:255–63.
2. Liesegang TJ. Herpes zoster ophthalmicus natural history, risk factors, clinical presentation, and morbidity. Ophthalmology **2008**; 115:S3–12.
3. Weinberg JM. Herpes zoster: epidemiology, natural history, and common complications. J Am Acad Dermatol **2007**; 57:S130–5.
4. Kawai K, Gebremeskel BG, Acosta CJ. Systematic review of incidence and complications of herpes zoster: towards a global perspective. BMJ Open **2014**; 4:e004833.
5. O'Connor KM, Paauw DS. Herpes zoster. Med Clin North Am **2013**; 97:503–22, ix.
6. Shaikh S, Ta CN. Evaluation and management of herpes zoster ophthalmicus. Am Fam Physician **2002**; 66:1723–30.
7. Yawn BP, Wollan PC, St Sauver JL, Butterfield LC. Herpes zoster eye complications: rates and trends [published correction appears in Mayo Clin. Proc. Sep 2017;92(9):1458]. Mayo Clin Proc **2013**; 88:562–70.
8. Lin HC, Chien CW, Ho JD. Herpes zoster ophthalmicus and the risk of stroke: a population-based follow-up study. Neurology **2010**; 74:792–7.
9. Yawn BP, Wollan P, St Sauver J. Comparing shingles incidence and complication rates from medical record review and administrative database estimates: how close are they? Am J Epidemiol **2011**; 174:1054–61.
10. Pimentel MA, Browne EN, Janardhana PM, et al. Assessment of the accuracy of using ICD-9 codes to identify uveitis, herpes zoster ophthalmicus, scleritis, and episcleritis. JAMA Ophthalmol **2016**; 134:1001–6.
11. Zheng C, Luo Y, Mercado C, et al. Using natural language processing for identification of herpes zoster ophthalmicus cases to support population-based study. Clin Exp Ophthalmol **2019**; 47:7–14.
12. Opstelten W, Mauritz JW, de Wit NJ, et al. Herpes zoster and postherpetic neuralgia: incidence and risk indicators using a general practice research database. Fam Pract **2002**; 19:471–5.
13. Borkar DS, Gonzales JA, Tham VM, et al. Association between atopy and herpetic eye disease: results from the pacific ocular inflammation study. JAMA Ophthalmol **2014**; 132:326–31.
14. Tseng HF, Bruxvoort K, Ackerson B, et al. The epidemiology of herpes zoster in immunocompetent, unvaccinated adults ≥50 years old: incidence, complications, hospitalization, mortality, and recurrence. J Infect Dis **2020**; 222:798–806.
15. Davies EC, Pavan-Langston D, Chodosh J. Herpes zoster ophthalmicus: declining age at presentation. Br J Ophthalmol **2016**; 100:312–4.
16. Koebnick C, Langer-Gould AM, Gould MK, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. Perm J **2012**; 16:37–41.
17. Womack LW, Liesegang TJ. Complications of herpes zoster ophthalmicus. Arch Ophthalmol **1983**; 101:42–5.
18. Ting DSJ, Ghosh N, Ghosh S. Herpes zoster ophthalmicus. BMJ **2019**; 364:k5234.
19. Vrcek I, Choudhury E, Durairaj V. Herpes zoster ophthalmicus: a review for the internist. Am J Med **2017**; 130:21–6.
20. Zheng C, Rashid N, Wu YL, et al. Using natural language processing and machine learning to identify gout flares from electronic clinical notes. Arthritis Care Res (Hoboken) **2014**; 66:1740–8.
21. Jackson LW, Kershner PRM. Herpes zoster ophthalmicus: a case of reactivated varicella. Hosp Physician **1999**; 35:45.
22. Kawai K, Yawn BP. Risk factors for herpes zoster: a systematic review and meta-analysis. Mayo Clin Proc **2017**; 92:1806–21.
23. Tran KD, Falcone MM, Choi DS, et al. Epidemiology of herpes zoster ophthalmicus: recurrence and chronicity. Ophthalmology **2016**; 123:1469–75.
24. Kong CL, Thompson RR, Porco TC, et al. Incidence rate of herpes zoster ophthalmicus: a retrospective cohort study from 1994 through 2018. Ophthalmology **2020**; 127:324–30.
25. Borkar DS, Tham VM, Esterberg E, et al. Incidence of herpes zoster ophthalmicus: results from the Pacific Ocular Inflammation Study. Ophthalmology **2013**; 120:451–6.
26. Thomas SL, Hall AJ. What does epidemiology tell us about risk factors for herpes zoster? Lancet Infect Dis **2004**; 4:26–33.
27. Forbes HJ, Bhaskaran K, Thomas SL, et al. Quantification of risk factors for herpes zoster: population based case-control study. BMJ **2014**; 348:g2911.
28. Jastrzebski A, Brownstein S, Ziai S, et al. Reactivation of herpes zoster keratitis with corneal perforation after zoster vaccination. Cornea **2017**; 36:740–2.
29. Marra F, Parhar K, Huang B, Vadlamudi N. Risk factors for herpes zoster infection: a meta-analysis. Open Forum Infect Dis **2020**; 7:ofaa005.