Autoimmune Effects of Lung Cancer Immunotherapy Revealed by Data-Driven Analysis on a Nationwide Cohort

Shihao Yang[1,†], Kun-Hsing Yu[1,2,†], Nathan Palmer[2], Kathe Fox[3], S. C. Kou[1,*] and Isaac S. Kohane[2,*]

The autoimmune adverse effects of lung cancer immunotherapy are not fully understood at the population level. Using observational data from commercial health insurance claims, we compared autoimmune diseases risk of immune checkpoint inhibitors (including pembrolizumab and nivolumab) and that of chemotherapy using the matching method. By 6 months after treatment initialization, the cumulative incidence of new autoimmune diseases among patients receiving immunotherapy was 13.13% (95% confidence interval (CI), 10.79–15.50%) and that of the matched chemotherapy patients was 6.65% (95% CI, 5.79–7.50%), constituting a hazard ratio (HR) of 1.97 (95% CI, 1.58–2.48). Both pembrolizumab (HR = 2.06 (95% CI, 1.20–3.65), $P = 0.0032$) and nivolumab (HR = 1.76 (95% CI, 1.39–2.24), $P < 0.0001$) were associated with higher risks of developing autoimmune diseases, especially for hypothyroidism ($P < 0.0001$). Our findings suggest the need to monitor autoimmune side effects of immunotherapy.

Study Highlights

WHAT IS THE CURRENT KNOWLEDGE ON THE TOPIC?

☑ Side effects of immune checkpoint inhibitors, including diarrhea and skin reaction, were reported during clinical trials. However, no studies investigate the risk of autoimmune disorders at a population level.

WHAT QUESTION DID THIS STUDY ADDRESS?

☑ What are the autoimmune effects of lung cancer immunotherapy on a national cohort?

WHAT DOES THIS STUDY ADD TO OUR KNOWLEDGE?

☑ Our real-world data analysis showed that those receiving immunotherapy were 1.97 times more likely to develop autoimmune diseases during the first 6 months of treatment, compared with the matched chemotherapy group. The limited granularity of the International Classification of Diseases (ICD) codes may affect the risk estimates.

HOW MIGHT THIS CHANGE CLINICAL PHARMACOLOGY OR TRANSLATIONAL SCIENCE?

☑ Our results suggested the clinical need to monitor autoimmune diseases in patients receiving immunotherapy agents. The reported methods could be routinely used for postmarket adverse effect surveillance of US Food and Drug Administration (FDA)-approved medications.

Lung cancer is the leading cause of cancer death worldwide, accounting for 1.58 million deaths per year.[1] Recent advances in clinical trials involving immunotherapy drugs showed great promise in treating this deadly malignancy.[2–4] The mechanism of action of immunotherapy is to enhance immune surveillance against tumor cells.[5] Studies have shown that immunotherapy extends lung cancer patients' life expectancy and decreases mortality.[3,5] Many immunotherapy drugs have been approved by the US Food and Drug Administration (FDA) to treat lung cancer and have entered the market since 2015.[3,6]

The most commonly used immunotherapy agents in lung cancer are programmed cell death protein 1 (PD-1) and programmed cell death 1 ligand 1 (PD-L1) inhibitors.[7] PD-1 inhibitors include pembrolizumab and nivolumab, and PD-L1 inhibitors include atezolizumab, avelumab, and durvalumab.[8] PD-1 is a transmembrane protein expressed in many immune cells, including T cells, and binds to its ligand PD-L1. The PD-1:PD-L1 binding promotes the conversion of T effector cells to regulatory T cells and inhibits the apoptosis of tumor cells.[8] PD-1 and PD-L1 inhibitors block such binding, thereby mobilizing immune cells against the tumor.[8,9] Compared with radiotherapy and chemotherapy, this treatment strategy avoided the direct cytotoxic damage to non-tumor tissues and is considered a significant advancement in the management of advanced lung cancer.[3,6]

[1]Department of Statistics, Harvard University, Cambridge, Massachusetts, USA; [2]Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts, USA; [3]Aetna Inc., Hartford, Connecticut, USA. *Correspondence: S. C. Kou (kou@stat.harvard.edu) and Isaac S. Kohane (Isaac_Kohane@hms.harvard.edu)

[†]Co-first authors.

Received April 10, 2019; accepted July 14, 2019. doi:10.1002/cpt.1597

Due to the recency of the clinical introduction of checkpoint inhibitor immunotherapy, the full spectrum of its side effects has not been studied at the population level. Known side effects of immunotherapy include fatigue, decreased appetite,[10] skin reactions, endocrine disorders, arthralgia,[11] pyrexia,[2] and drug-induced hepatitis.[12] These adverse effects are reported by clinicians based on their observations during clinical trials[2,3] and in postmarket surveillance of immunotherapy agents. However, for adverse effects with lower incidence rates, it will require additional time to accumulate a sufficient number of cases for an adverse event to receive adequate attention. As PD-1 and PD-L1 play a substantial role in modulating the immune system and preventing autoimmune diseases,[13] immunotherapy agents have been associated with immune-related adverse events.[14] Sporadic cases of hyperthyroidism, hypothyroidism, diarrhea, colitis, and skin reaction have been reported in the clinical trials.[2,3] In addition, case reports have described the development of cerebral vasculitis in lung cancer patients receiving anti-PD-1 drugs.[15] The onsets of pneumonitis,[16] fulminant myocarditis,[17] and thyroiditis[14,18] have been implicated with the use of PD-1 checkpoint inhibitors in treating other cancer types. However, there are no studies that investigate and quantify the risk of autoimmune disorders at a population level.[19] Accurate estimates of these risks will help guide oncologists and their lung cancer patients in selecting immunotherapies or other treatment modalities.

The recent availability of electronic health data at the population scale has enabled large-scale studies on the adverse events of novel treatment modalities using real-world data.[19,20] Leveraging nationwide insurance data sets, researchers have identified the cancer risk of rheumatoid arthritis patients[21] and have detected drug adverse effects at the population level.[22] These studies indicated the potential of using large data sets to detect the adverse effects of new treatment strategies.

In this study, we utilized the real-world data from a nationwide health insurance data set to identify the risk of developing autoimmune diseases in lung cancer patients who have undergone immunotherapy. Using a de-identified nationwide cohort that covers 54 million insured members in North America, we identified 1,809 lung cancer patients receiving immunotherapy and extracted the associations between immunotherapy and the risk of autoimmune diseases systematically. Our results indicated that both pembrolizumab and nivolumab are associated with a significantly higher risk of developing autoimmune diseases (at the 0.05 significance level). The increased risk of autoimmune diseases in patients receiving immunotherapy should prompt clinicians to evaluate their patients carefully for evidence of autoimmune diseases.[19] Our methods are extensible to identifying the adverse effects of other treatment modalities.

RESULTS

Overview

Figure 1 shows the derivation of our study cohort, and **Table 1** shows patients' detailed demographics. During the period of January 1, 2008, to June 30, 2017, we identified 1,809 lung cancer patients without prior autoimmune diseases before treatment initiation who received immune checkpoint inhibitors, of which 374 were on pembrolizumab (20.67%) and 1,392 were on nivolumab (76.95%). In addition, we identified 24,186 of those who received chemotherapy but no immune checkpoint inhibitors. The ethnicity is known in about 13% of the patients. The sex and age distribution were not significantly different between patients receiving immunotherapy and those receiving chemotherapy ($P = 0.0532$ for sex and $P = 0.1145$ for age). However, the ethnicity distribution was different ($P < 0.0001$), demonstrating the necessity for matching to balance those baseline covariates. **Figure S1** suggests the covariate balance after matching was better than that before matching, as expected. About 80% of the patients with immunotherapy could find at least one matched chemotherapy patient (**Table 1**).

Regarding the therapy initiation date (**Figure 2**), almost all immunotherapies were initiated after 2015, with the number of patients on pembrolizumab increasing yearly, and the number of patients on nivolumab peaking in the first quarter of 2016 and remaining stable afterward.

Immunotherapy and the risk of developing autoimmune diseases

Figure 3 shows the time-to-event plot for immunotherapy compared with chemotherapy. For analyses based on both before-matching and post-matching samples, immunotherapy in lung cancer patients was associated with a higher risk of developing autoimmune diseases compared with those on chemotherapy. Such difference in autoimmune diseases rate is statistically significant ($P < 0.0001$; **Table 2**), with post-matching (against chemotherapy) hazard ratios for the cumulative incidence rates of 1.97 (95% confidence interval (CI), 1.58–2.48) and 1.88 (95% CI, 1.52–2.33) at 6 months and 15 months after treatment initiation for the at-risk patients (i.e., those still in the data set. Patients who dropped out due to death or switching of insurance plans were considered as censored). In addition, when we stratified the time-to-autoimmune-diseases analysis into 56 specific autoimmune disease categories, we found that patients who received immunotherapy had significantly higher risks of acquiring hypothyroidism compared with those receiving chemotherapy (**Figure 3**, Benjamini & Hochberg adjusted $P < 0.0001$).

In assessing the specific immune checkpoint inhibitors (pembrolizumab and nivolumab), we observed that patients who received pembrolizumab had a higher risk of developing autoimmune diseases compared with those treated by chemotherapy ($P = 0.0032$; **Figure 3** and **Table 2**), and the nivolumab group also had a higher risk of developing autoimmune diseases compared with the chemotherapy group ($P < 0.0001$). Patients treated with pembrolizumab had hazard ratios of 2.06 (95% CI, 1.20–3.65) and 2.22 (95% CI, 1.21–4.39) compared with the matched chemotherapy group at 6 and 15 months after therapy initiation for the at-risk patients, and those on nivolumab also had similar increase in hazard ratios (1.76 (95% CI, 1.39–2.24) and 1.70 (95% CI, 1.35–2.16) at 6 and 15 months, respectively; **Table 2**).

For the stratification of 56 different autoimmune disease categories, the differential risk of hypothyroidism was significant in both the nivolumab group (Benjamini & Hochberg adjusted $P < 0.0001$) and the pembrolizumab group (Benjamini & Hochberg adjusted $P = 0.0001$).

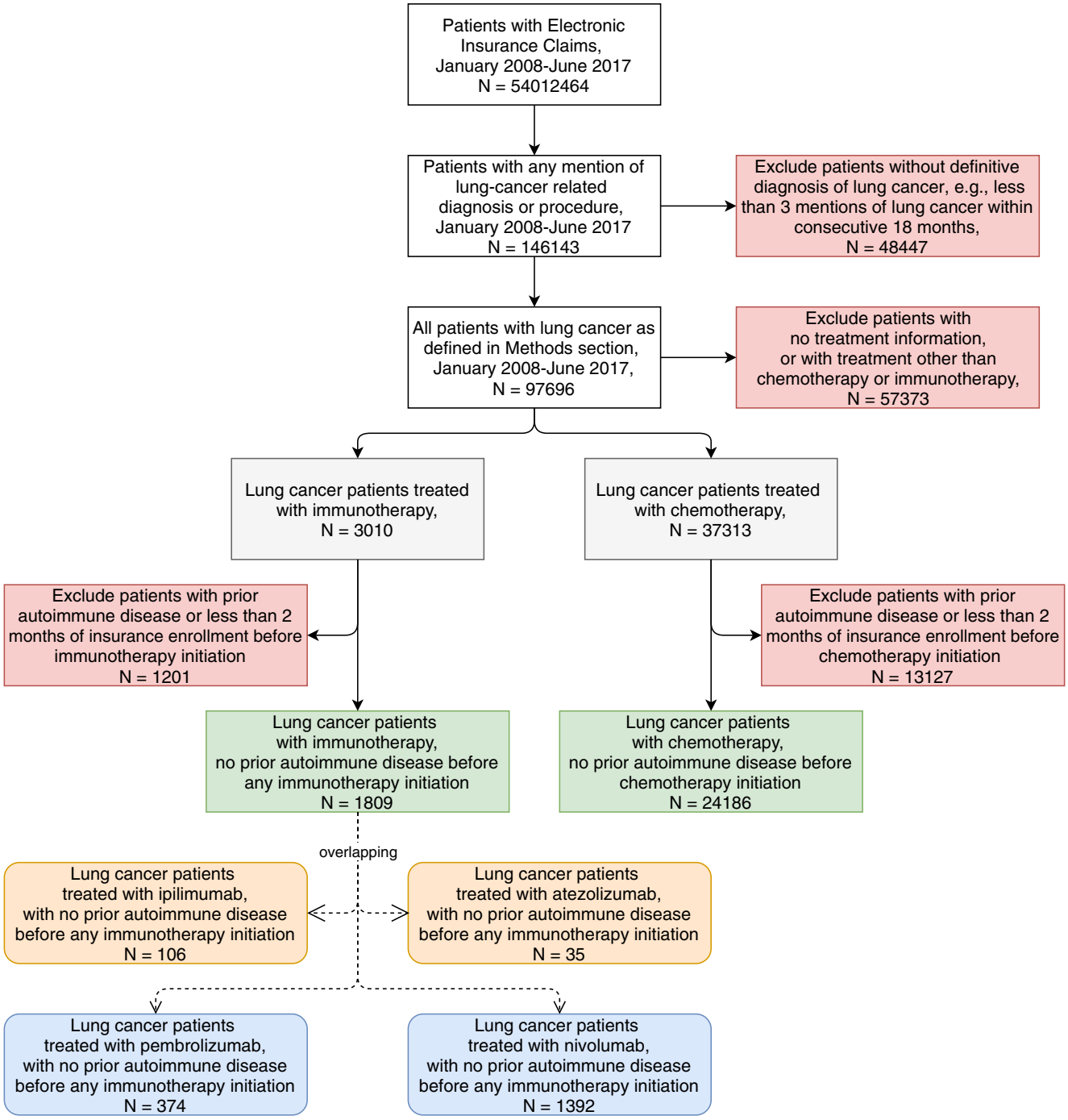

Figure 1 Derivation of the study cohort.

Sensitivity analyses. By varying different matching criteria, we demonstrated that the result was robust to the categorization of baseline hospital visits and the number of diagnoses. We also found results to be robust to the definition of the dropout criteria by looking only at claim activity and membership expiration and imposing a 1-month grace period after inactivity. We further verified that our results were insensitive to different quiescence period requirements. In addition, we also confirmed our results to be robust to time-frame specification where patients are required to have chemotherapy or immunotherapy initiation after 2015. To validate the results on hypothyroidism, we additionally required the identification of thyroid-related diseases to be having one disease International Classification of Diseases (ICD) code plus one or more relevant National Drug Code (NDC) drug codes or Current Procedural Terminology (CPT) codes and found that immunotherapy's effect on hypothyroidism risk persisted under the new identification procedure. For the patients receiving immunotherapy as a second-line treatment after chemotherapy,

Table 1 Patient characteristics

	Pembrolizumab		Nivolumab		All immunotherapy drugs		Chemotherapy	
	N	%	*N*	%	*N*	%	*N*	%
All	374	100.00	1,392	100.00	1,809	100.00	24,186	100.00
Sex								
P value of difference with chemotherapy	1.0000		0.0458		0.0532		Control group	
F	163	43.58	570	40.95	748	41.35	10,573	43.72
M	211	56.42	822	59.05	1,061	58.65	13,613	56.28
Ethnicity								
P value of difference with chemotherapy	0.3797		<0.0001		<0.0001		Control group	
African American/Black	<5	<1.34	17	1.22	19	1.05	170	0.70
American Indian/Alaskan Native	0	0.00	<5	<0.36	<5	<0.28	13	0.05
Asian	<5	<1.34	<5	<0.36	7	0.39	84	0.35
Black (non-Hispanic)	0	0.00	0	0.00	0	0.00	<5	<0.02
White (non-Hispanic)	<5	<1.34	<5	<0.36	<5	<0.28	24	0.10
Hispanic/Latino	<5	<1.34	10	0.72	11	0.61	68	0.28
Other	<5	<1.34	6	0.43	7	0.39	39	0.16
Pacific Islander	0	0.00	0	0.00	0	0.00	7	0.03
Two or more races	5	1.34	14	1.01	20	1.11	193	0.80
White	49	13.10	201	14.44	258	14.26	2,383	9.85
Unknown	313	83.69	1,138	81.75	1,484	82.03	21,201	87.66
Age								
P value of difference with chemotherapy	0.0018		0.0613		0.1145		Control group	
[0,40)	13	3.48	15	1.08	28	1.55	334	1.38
[40,50)	22	5.88	60	4.31	86	4.75	1,359	5.62
[50,60)	75	20.05	288	20.69	378	20.90	5,536	22.89
[60,70)	116	31.02	507	36.42	635	35.10	8,422	34.82
[70,80)	97	25.94	375	26.94	480	26.53	6,115	25.28
[80,120)	51	13.64	147	10.56	202	11.17	2,420	10.01
Post-matching size	298	79.68	1,098	78.88	1,424	78.72	Control group	

<5: Suppressed patient count to protect patient privacy.

we conducted additional sensitivity analysis for the definition of treatment initiation date by changing this index date to be the time of the first chemotherapy initiation. As such, we derived a conservative lower bound of autoimmune risk of the second-line immunotherapy, which confirmed our main results. Details are included in the Supporting Information (**Figures S3–S8**).

Subpopulation analyses. We further investigated our results in different subsets of study populations. We stratified the results by the lines of therapy in which immunotherapy was received and found a similar differential risk of immunotherapy to autoimmune disease regardless of the lines of the therapy. We also compared the rates of these autoimmune adverse events between the two sexes and found male and female patients having similar differential risk. However, when we compared across different age groups, we found the differential risk of immunotherapy to autoimmune disease to be less strong in the elderly group. Details are included in the Supporting Information (**Figures S9–S11**).

DISCUSSION

We conducted a large population-level analysis on the risk of autoimmune diseases in patients treated for lung cancer and found that patients who received immune checkpoint inhibitors were more likely to develop autoimmune diseases compared with those receiving conventional chemotherapy. Patients treated with nivolumab or pembrolizumab had significantly higher risks of developing hypothyroidism. These results further quantify the associations between commonly used immunotherapy drugs and the development of autoimmune diseases that have been anticipated on an epidemiological and mechanistic basis.[23,24]

Immunotherapy has received a great amount of scientific and media attention in the recent years, as it has significantly changed the outcomes of advanced-stage lung cancer patients, including in randomized controlled trials. In this nationwide cohort, we observe the subsequent paradigm shift in lung cancer treatment. In early 2015, there were fewer than 50 patients treated with any immunotherapy drugs in our cohort. Since Q3 2015, the use of

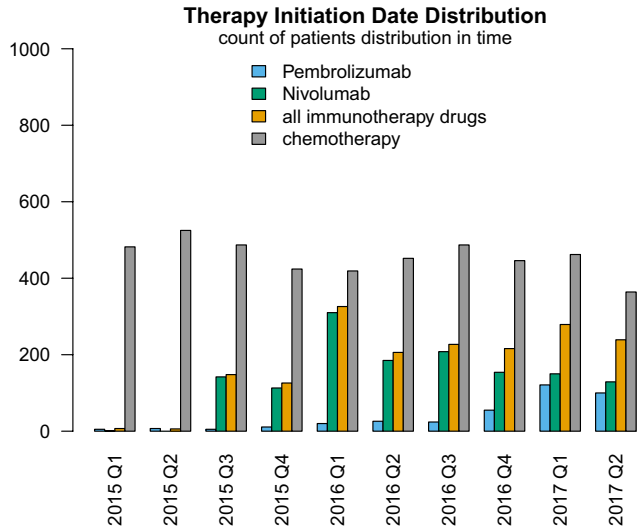

Figure 2 Treatment initiation date distribution. The histogram shows the number of patient counts for different treatments from 2015 to 2017. For better comparisons to recent immunotherapy, we only show treatment initiation that starts on or after 2015 in this graph.

immunotherapy agents increased substantially. The number of patients treated with nivolumab seemed to plateau in late 2016, while the number of patients receiving pembrolizumab continued to rise, although there were still fewer lung cancer patients treated with pembrolizumab than those treated with nivolumab in 2017. This might stem from the fact that nivolumab was approved for treating non-small cell lung cancer earlier than pembrolizumab.[25,26]

Data from the patients treated in the first 2–3 years of the introduction of checkpoint inhibitors allowed us to systematically estimate the potential adverse effects of immune checkpoint inhibitors. Since PD-1 and PD-L1 inhibitors block the PD-1:PD-L1 binding[8,9] and the PD-1:PD-L1 pathway can thwart self-reactive T cells and reduce autoimmune reactions, we hypothesized that PD-1 and PD-L1 inhibitors could have promoted the development of autoimmune diseases,[27] and we validated this hypothesis in our analysis. As lung cancer patients receiving immunotherapy demonstrated better survival outcomes,[3,6,28] longer-term adverse effects involving the immune system may become increasingly clinically relevant.

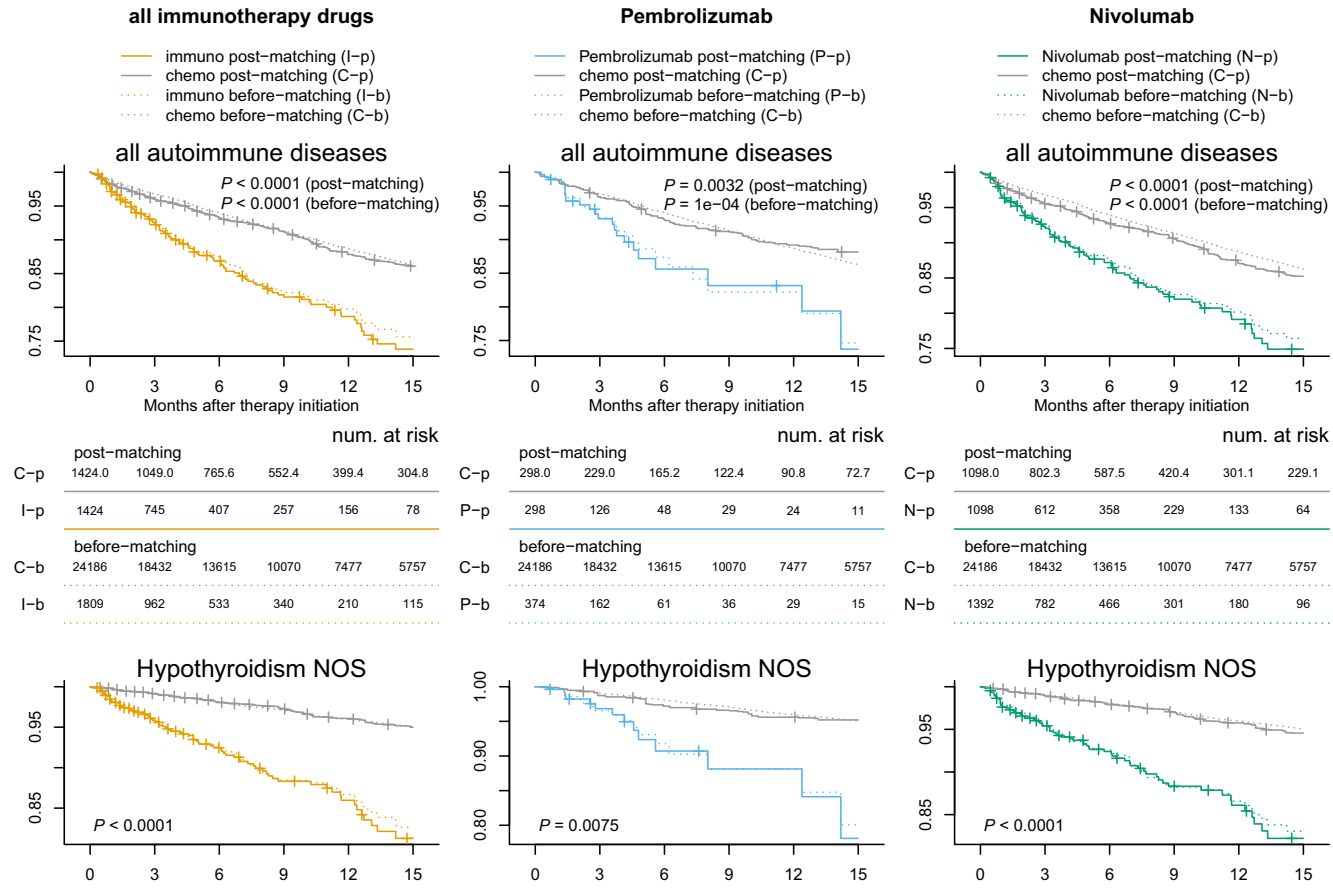

Figure 3 Time to autoimmune diseases in patients, comparing those receiving chemotherapy with (**a**) all immunotherapy drugs in the left column, (**b**) pembrolizumab in the middle column, and (**c**) nivolumab in the right column. Upper panels are time-to-event analysis for all autoimmune disease types combined, where log-rank test *P* values for unmatched samples and matched samples are reported in the upper right corner, respectively. Lower panels are time-to-event analysis for specific categories of autoimmune disease, where *P* values are based on the log-rank test on matched samples with Benjamini-Hochberg correction for multiple testing of all 56 autoimmune disease types. Only hypothyroidism is shown here. Each tick represents 50 censored patients without an autoimmune disease event. The curves are truncated at the 15th month after treatment initiation for better visualization; plot for other autoimmune disease categories and a longer time horizon can be found in **Figure S2** of the Supporting Information. NOS, not otherwise specified; num., number.

Table 2 The cumulative incidence rate (in %) at the end of each quarter after treatment initiation, comparing patients on any immunotherapy, patients on pembrolizumab, and patients on nivolumab with the matched chemotherapy groups

	Adjusted			Unadjusted		
Index	Immunotherapy	Chemotherapy	Hazard ratio	Immunotherapy	Chemotherapy	Hazard ratio
All immunotherapy						
Month = 3	7.51 [5.92, 9.10]	3.96 [3.29, 4.62]	1.90 [1.45, 2.50]	7.15 [5.76, 8.50]	3.21 [2.97, 3.44]	2.23 [1.82, 2.74]
Month = 6	13.13 [10.79, 15.50]	6.65 [5.79, 7.50]	1.97 [1.58, 2.48]	12.50 [10.47, 14.49]	6.06 [5.72, 6.40]	2.06 [1.74, 2.45]
Month = 9	18.15 [15.06, 21.27]	8.76 [7.71, 9.78]	2.07 [1.68, 2.58]	17.55 [14.82, 20.24]	8.67 [8.23, 9.10]	2.02 [1.72, 2.39]
Month = 12	21.34 [17.66, 25.06]	12.17 [10.79, 13.54]	1.75 [1.42, 2.18]	20.26 [17.09, 23.39]	11.33 [10.79, 11.86]	1.79 [1.52, 2.11]
Month = 15	26.18 [21.38, 30.97]	13.96 [12.45, 15.47]	1.88 [1.52, 2.33]	24.36 [20.33, 28.32]	13.71 [13.08, 14.34]	1.78 [1.50, 2.11]
Log-rank test		$P < 0.0001$			$P < 0.0001$	
Pembrolizumab						
Month = 3	6.91 [3.26, 10.50]	3.64 [2.13, 5.15]	1.90 [0.99, 3.70]	6.74 [3.51, 9.95]	3.21 [2.97, 3.44]	2.10 [1.30, 3.59]
Month = 6	14.39 [7.92, 20.80]	6.99 [4.88, 9.09]	2.06 [1.20, 3.65]	12.61 [7.32, 17.87]	6.06 [5.72, 6.40]	2.08 [1.38, 3.30]
Month = 9	16.84 [8.89, 24.78]	8.80 [6.37, 11.20]	1.91 [1.11, 3.45]	17.82 [10.11, 25.43]	8.67 [8.23, 9.10]	2.05 [1.34, 3.28]
Month = 12	16.84 [8.89, 24.78]	10.82 [8.10, 13.50]	1.56 [0.91, 2.78]	17.82 [10.11, 25.43]	11.33 [10.79, 11.86]	1.57 [1.03, 2.51]
Month = 15	26.29 [11.52, 41.09]	11.85 [9.01, 14.65]	2.22 [1.21, 4.39]	25.37 [12.75, 37.87]	13.71 [13.09, 14.33]	1.85 [1.14, 3.18]
Log-rank test		$P = 0.0032$			$P = 0.0001$	
Nivolumab						
Month = 3	7.81 [6.01, 9.59]	4.45 [3.68, 5.24]	1.76 [1.31, 2.36]	7.36 [5.81, 8.90]	3.21 [2.97, 3.44]	2.29 [1.84, 2.87]
Month = 6	12.81 [10.32, 15.29]	7.28 [6.30, 8.26]	1.76 [1.39, 2.24]	12.19 [10.03, 14.32]	6.06 [5.72, 6.40]	2.01 [1.68, 2.43]
Month = 9	17.65 [14.43, 20.90]	9.38 [8.24, 10.54]	1.88 [1.50, 2.37]	17.06 [14.20, 19.87]	8.67 [8.24, 9.10]	1.97 [1.66, 2.35]
Month = 12	20.89 [17.02, 24.79]	12.95 [11.39, 14.52]	1.61 [1.29, 2.03]	19.84 [16.49, 23.14]	11.33 [10.81, 11.86]	1.75 [1.47, 2.09]
Month = 15	25.10 [20.14, 30.11]	14.76 [13.03, 16.50]	1.70 [1.35, 2.16]	23.57 [19.37, 27.72]	13.71 [13.10, 14.33]	1.72 [1.43, 2.07]
Log-rank test		$P < 0.0001$			$P < 0.0001$	

Our study revealed that patients receiving immunotherapy are more likely to develop a plethora of autoimmune diseases, including autoimmune hypothyroidism that necessitated treatments. When comparing the two most commonly prescribed immunotherapy agents, both pembrolizumab and nivolumab are associated with a higher rate of autoimmune diseases overall. With regard to specific autoimmune diseases, patients receiving pembrolizumab or nivolumab had a higher risk of getting autoimmune-related hypothyroidism, and the effect size of the two immune checkpoint inhibitors are similar throughout our observational period. Our results suggest pembrolizumab and nivolumab may present similar autoimmune syndromes and overall autoimmune diseases risk profile, which may be expected due to their similar mechanism of action that targets PD-1.

Comparing with the adverse effects reported in randomized controlled trials,[2,3] our study assembled a significantly larger cohort and revealed autoimmune adverse effects in the real-world data. Results from clinical trials hinted at the risk of hypothyroidism in the immunotherapy group but reported fewer than 22 cases in either the pembrolizumab[2] or the nivolumab trial.[3] Harnessing the nationwide insurance data, we systematically investigated the risk of developing each autoimmune disorder in both the immunotherapy and the chemotherapy groups and demonstrated the effects with greater statistical power. In addition, although internal validity is expected for results from randomized controlled trials, their external validity has been challenged.[29,30] As an illustration, participants in drug trials generally have fewer comorbidities. However, once the new drugs are approved, they are applied to a

much larger group of people with different genetic make-ups, various comorbidities, and a myriad of treatments for other comorbidities.[30] Thus, the effects of immune checkpoint inhibitors in the real world may be different from what could be extrapolated from trial results. Utilizing the real-world data, our analyses showed that autoimmune hypothyroidism is more common among patients receiving immune checkpoint inhibitors and may require clinical attention. Our approach complements the conventional methods of investigating the adverse effects of cancer therapies.

It is also worth noting that our methods for identifying treatment adverse effects are easily extensible to other treatment modalities and disease phenotypes. Our approaches can accommodate any treatment of interest with specific drug codes and disease phenotypes with appropriate ICD codes. The sensitivity analyses presented in this study are instrumental for ascertaining the plausibility of the detected adverse effects, and the subgroup analyses identified potential effect modifications among different populations in the nationwide insurance claims data sets. The reported methods could be routinely used for postmarket adverse effect surveillance of FDA-approved medications.

One limitation of this study is that the analysis is restricted to the diagnostic codes of autoimmune diseases. The ICD codes recorded in the insurance claims database may not be able to identify patients who have started to develop low-level autoimmune responses but did not present clinically apparent autoimmune diseases, nor is it able to directly capture the molecular mechanisms underpinning the observed autoimmune diseases. For example, immunotherapy agents may promote the formation of autoantibodies or activate T cells before the patients develop clinical manifestations of autoimmune diseases, and these subclinical responses would not be evident from the claims data nor from many clinical health records. In addition, the ICD codes do not distinguish different histological subtypes of lung cancer, PD-L1 status, or aberrations in *EGFR* or *ALK* genes. The use of immunotherapy on different lung cancer subtypes was not approved simultaneously. For example, nivolumab was approved for treating metastatic squamous non-small cell lung cancer on March 4, 2015, and the approval was later extended to metastatic non-squamous non-small cell lung cancer on October 9, 2015. Moreover, PD-L1 expression levels and the genomic variation status of the patients impact the effectiveness of immunotherapy and thus affect clinicians' decisions of which patients would receive immune checkpoint inhibitors. Another limitation is that insurance claims did not routinely record certain disease and demographic information. For example, patients' stage information was not encoded in the ICD codes, and the majority of patients did not specify their ethnicity. Lastly, our data set did not capture all lung cancer patients in the United States, and the patient age in this commercial health insurance was relatively young since elder patients may be eligible for Medicare. Although we conducted rigorous sensitivity analyses to mitigate the effects of noise in this real-world data set, further studies are needed to validate our findings in other populations.

In conclusion, this study quantifies the increased risk of autoimmune diseases in lung cancer patients receiving immunotherapy, suggesting the clinical need to monitor autoimmune diseases in patients receiving immunotherapy agents. Further studies are needed to identify the mechanisms of the increased autoimmune risk in patients receiving immune checkpoint inhibitors, characterize agent-specific autoimmune vulnerabilities, and monitor other life-threatening adverse effects of such treatment strategy.

MATERIALS AND METHODS

Study cohort

Using de-identified member claims data from Aetna (Hartford, CT), lung cancer patients who received chemotherapy or immunotherapy were selected. The data set contains de-identified medical and pharmacy claims, enrollment, and demographics for about 54 million insured members in the United States from January 1, 2008, to June 30, 2017. This study was approved by the Harvard Medical School Institutional Review Board. The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Patients with lung cancer were identified based on three related criteria: Diagnosis Criterion (ICD-9 and ICD-10 codes in **Table S1 of Supplementary Material**), Procedure Criterion (CPT codes in **Table S2 of Supplementary Material**), and Pharmacy Criterion (NDC codes in **Table S3** and Healthcare Common Procedure Coding System (HCPCS) codes in **Table S4 of Supplementary Material**). To be identified as having lung cancer, a patient needed to have three independent claims records meeting the Diagnosis Criterion; or three independent claims records meeting the Procedure Criterion; or two independent claims records meeting any two of the Diagnosis Criterion, Procedure Criterion, or Pharmacy Criterion. The identifying claims records were required to be within a consecutive 18-month period. This process ruled out most patients without definitive diagnoses of lung cancer or with miscoded lung cancer–related claims. See **Supporting Information** for details.

Among the identified lung cancer patients, those with any HCPCS/NDC code corresponding to immune checkpoint inhibitors (nivolumab, pembrolizumab, ipilimumab, atezolizumab; **Table S5 of Supplementary Material**) were identified as patients on immunotherapy; those with any HCPCS/CPT/NDC code corresponding to lung cancer chemotherapy drugs (carboplatin, cisplatin, docetaxel, etoposide, gemcitabine, irinotecan, mechlorethamine, methotrexate, paclitaxel, pemetrexed, vinblastine, vinorelbine; **Table S6 of Supplementary Material**) but without any immune checkpoint inhibitors were identified as patients on chemotherapy. The treatment initiation date was defined to be the first date of administration of the corresponding drug.

Besides treatment types, additional independent variables (including age, sex, ethnicity (partially available), zip code, the annualized number of hospital visits prior to treatment initialization, and the annualized number of ICD code counts prior to treatment initialization) were retrieved for each patient. The median income and unemployment rate were identified for each zip code using 2010 US census data.

Eligibility and censoring

Patients with any autoimmune disease ICD codes (**Table S7 of Supplementary Material**) prior to treatment initiation were considered as having preexisting autoimmune diseases and were excluded from the study. To rule out patients with prior cancer treatment billed to other insurance policies, a 2-month quiescence period without any mention of lung cancer treatment was required upon insurance enrollment. That is, patients with mentions of immunotherapy or chemotherapy treatment within the first 2 months of insurance enrollment were excluded from the study. To ensure patients in the study cohort were continuously enrolled and thus having the vast majority of their health information reflected in the data set, patient's medical records were censored at the

first insurance membership expiration date after treatment initiation. Records after the censoring date were disregarded in this study.

Outcomes of the study

The primary outcome of the study is the time from treatment initiation to the development of autoimmune disease. For patients receiving chemotherapy only, the treatment initiation date is set to be the first chemotherapy date; for patients receiving immunotherapy, the treatment initiation date is set to be the first immunotherapy date, irrespective of the lines of the immunotherapy. Fifty-six categories of ICD codes associated with autoimmune diseases,[31,32] including those for rheumatoid arthritis, systemic lupus erythematosus, ankylosing spondylitis, type 1 diabetes mellitus, inflammatory bowel disease (including ulcerative colitis and Crohn's disease), autoimmune-related hypothyroidism, thyrotoxicosis, and thyroiditis, sicca syndrome, and systemic sclerosis (a complete list of autoimmune diseases and their ICD codes are shown in **Table S7 of Supplementary Material**), were examined. Patients who acquired any ICD code for these autoimmune diseases after lung cancer treatment were considered as having an event, and the time to event was defined as the number of days from lung cancer treatment initiation to the date of the first appearance of an autoimmune ICD code, and the corresponding autoimmune disease was defined as the acquired autoimmune disease. Patients without any autoimmune diseases ICD codes were regarded as having no event, and the at-risk time is defined as the number of days from lung cancer treatment initiation to the date of the last medical record or the end of continuous insurance enrollment, whichever is earlier.

Causal inference analysis on posttreatment autoimmune diseases

Matching was used to conduct causal inference analyses. The log-rank test on time-to-event outcome was employed to identify the significance of any observed difference.

Three related treatment groups were considered: lung cancer patients who received (i) any immune checkpoint inhibitors, (ii) pembrolizumab, and (iii) nivolumab. The control group consisted of lung cancer patients who received chemotherapy without any immune checkpoint inhibitors.

Matching was performed such that each patient in the immune checkpoint inhibitors group was matched to multiple patients in the control group[33] by sex (exactly the same), age (plus or minus 2 years), ethnicity (exactly the same whenever available), zip-code-defined median income (quintile bins), zip-code-defined unemployment rate (quintile bins), the annualized number of hospital visits prior to treatment initialization (quintile bins), and the annualized number of ICD code counts prior to treatment initialization (quintile bins). The first five matching conditions were proxies for socioeconomic status, whereas the last two were proxies for healthcare utilization and general sickness levels,[34] and they were calculated for each patient right before the lung cancer treatment initialization. Due to the limitation in sample size, it was not feasible to match for other covariates such as prior medical history or concurrent medications.

Once matching was completed, time-to-event analyses were conducted through inverse probability weighting. Each member of the treatment group was given a weight of 1, and the corresponding matched control group members were given weights inverse to the total number of matches for that treatment group member. For instance, if four participants in the control group were matched to one person in the treatment group, each of the four controls would receive a weight of 0.25. The Kaplan–Meier curves of time to the onset of autoimmune diseases were plotted for both post-matching samples (with inverse probability weighting adjustment) and before-matching samples. The statistical significance was assessed using the weighted log-rank test for both post-matching and before-matching samples. The 95% CIs for Kaplan–Meier estimate of cumulative incidence rates and hazard ratios were constructed using the bootstrap method.[35] Details are included in the **Supporting Information**.

Sensitivity analyses

To ensure the robustness of our results, extensive sensitivity analyses were conducted by varying definitions and settings in study design. Statistical models with different definitions of data censoring date (e.g., censor on the last claim activity, insurance policy cancellation, or after 1 month following the last claim activity) were fitted, the percentile specifications were varied when matching for continuous covariates, and the different lengths of quiescence period requirement (1 week, 1 month, 3 months, and 12 months) after insurance enrollment were tested. A sensitivity study was also conducted on the time-frame specification where the patients were additionally required to have chemotherapy or immunotherapy initiation date after 2015. To further validate the results on the risk of clinically significant thyroid diseases, we conducted sensitivity analysis where the identification of thyroid diseases, in addition to ICD code, must be accompanied by at least one subsequent NDC drug code or CPT code related to the type of thyroid disease. For the patients receiving immunotherapy as a second line after chemotherapy, we also investigated the effect of redefining the treatment initiation date to be the time of first chemotherapy initiation. The corresponding results from these models were examined and compared with those from the primary analyses.

Subpopulation analyses

We further conducted the same analyses stratified by different subpopulations. We first stratified the results by the line of therapy in which immunotherapy was received to evaluate the impact of immunotherapy as first-line or second-line treatments. We also compared the rates of these autoimmune adverse events across different populations by sex and age groups to determine whether there was any difference in their risk profiles.

SUPPORTING INFORMATION

Supplementary information accompanies this paper on the *Clinical Pharmacology & Therapeutic*s website (www.cpt-journal.com).

Figure S1. Covariate balance before and after matching.
Figure S2. Time-to-autoimmune diseases in patients, comparing those receiving chemotherapy with (**a**) all immunotherapy drugs in the left column, (**b**) pembrolizumab in the middle column, and (**c**) nivolumab in the right column.
Figure S3. Sensitivity to the categorization of continuous variables for matching.
Figure S4. Sensitivity analyses on the definitions of censoring.
Figure S5. Sensitivity analyses on the quiescence period requirement.
Figure S6. Analyses with additional time-frame requirement that patients must have chemotherapy or immunotherapy initiation date after 2015.
Figure S7. Sensitivity analyses on the clinically significant thyroid diseases with further evaluations or treatments.
Figure S8. Sensitivity studies on the impact of treatment initiation date definition for patients on second-line immunotherapy.
Figure S9. (**a**) Results stratified by the lines of immunotherapy. This panel shows results where the treatment group are all first-line immunotherapy patients. (**b**) Results stratified by lines of immunotherapy. This panel shows results where the treatment group are all second-line immunotherapy patients.
Figure S10. (**a**) Results stratified by sex. This panel shows results where both the treatment group and the control group are all female patients. (**b**) Results stratified by sex. This panel shows results where both the treatment group and the control group are all male patients.
Figure S11. (**a**) Results stratified by age groups. This panel shows results where both the treatment group and the control group are restricted to have age <60. (**b**) Results stratified by age group. This panel shows results where both the treatment group and the control group are restricted to have age greater or equal to 60 but less than 70. (**c**) Results stratified by age group. This panel shows results where both the treatment group and the control group are restricted to have age greater or equal to 70 but less than 80. (**d**) Results stratified by age group. This panel shows results where both the treatment group and the control group are restricted to have age greater or equal to 80.
Supplemental Material.

ACKNOWLEDGMENTS
We thank L. J. Wei and Tianxi Cai for their helpful suggestions on survival analysis methods, and Kenneth L. Kehl for helpful feedback on autoimmune diseases categorization. We thank the Nvidia Corporation Grant Program for supporting the computational infrastructure.

FUNDING
This work is supported in part by National Science Foundation (NSF) grant DMS-1510446, DMS-1810914 (S.C.K.), and Harvard Data Science Fellowship (K.-H. Y.).

CONFLICT OF INTEREST
K.F. is directly employed by Aetna. All other authors declared no competing interests for this work.

AUTHOR CONTRIBUTIONS
S.Y., K.-H.Y., S.C.K., and I.S.K. wrote the manuscript; S.Y., K.-H.Y., S.C.K., and I.S.K. designed the research; S.Y. and K.-H.Y. performed the research; S.Y. and K.-H.Y. analyzed the data; S.Y., K.-H.Y., N.P., K.F., S.C.K., and I.S.K. contributed new reagents/analytical tools.

ROLE OF THE FUNDER/SPONSOR
The funding sources had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

1. Torre, L.A., Bray, F., Siegel, R.L., Ferlay, J., Lortet-Tieulent, J. & Jemal, A. Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87–108 (2015).
2. Reck, M. *et al*. Pembrolizumab versus chemotherapy for PD-L1-positive non-small-cell lung cancer. *N. Engl. J. Med.* **375**, 1823–1833 (2016).
3. Borghaei, H. *et al*. Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* **373**, 1627–1639 (2015).
4. Yu, K.H. & Snyder, M. Omics profiling in precision oncology. *Mol. Cell Proteomics* **15**, 2525–2536 (2016).
5. Brahmer, J.R. & Pardoll, D.M. Immune checkpoint inhibitors: making immunotherapy a reality for the treatment of lung cancer. *Cancer Immunol. Res.* **1**, 85–91 (2013).
6. Herbst, R.S. *et al*. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet* **387**, 1540–1550 (2016).
7. Anagnostou, V.K. & Brahmer, J.R. Cancer immunotherapy: a future paradigm shift in the treatment of non-small cell lung cancer. *Clin. Cancer Res.* **21**, 976–984 (2015).
8. Hamanishi, J., Mandai, M., Matsumura, N., Abiko, K., Baba, T. & Konishi, I. PD-1/PD-L1 blockade in cancer treatment: perspectives and issues. *Int. J. Clin. Oncol.* **21**, 462–473 (2016).
9. Zou, W., Wolchok, J.D. & Chen, L. PD-L1 (B7-H1) and PD-1 pathway blockade for cancer therapy: mechanisms, response biomarkers, and combinations. *Sci. Transl. Med.* **8**, 328rv4 (2016).
10. Topalian, S.L. *et al*. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N. Engl. J. Med.* **366**, 2443–2454 (2012).
11. Brahmer, J.R. *et al*. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N. Engl. J. Med.* **366**, 2455–2465 (2012).
12. Cramer, P. & Bresalier, R.S. Gastrointestinal and hepatic complications of immune checkpoint inhibitors. *Curr. Gastroenterol. Rep.* **19**, 3 (2017).
13. Francisco, L.M., Sage, P.T. & Sharpe, A.H. The PD-1 pathway in tolerance and autoimmunity. *Immunol. Rev.* **236**, 219–242 (2010).
14. Postow, M.A., Sidlow, R. & Hellmann, M.D. Immune-related adverse events associated with immune checkpoint blockade. *N. Engl. J. Med.* **378**, 158–168 (2018).
15. Läubli, H. *et al*. Cerebral vasculitis mimicking intracranial metastatic progression of lung cancer during PD-1 blockade. *J. Immunother. Cancer* **5**, 46 (2017).
16. Nishino, M., Sholl, L.M., Hodi, F.S., Hatabu, H. & Ramaiya, N.H. Anti-PD-1-related pneumonitis during cancer immunotherapy. *N. Engl. J. Med.* **373**, 288–290 (2015).
17. Johnson, D.B. *et al*. Fulminant myocarditis with combination immune checkpoint blockade. *N. Engl. J. Med.* **375**, 1749–1755 (2016).
18. Morganstein, D.L. *et al*. Thyroid abnormalities following the use of cytotoxic T-lymphocyte antigen-4 and programmed death receptor protein-1 inhibitors in the treatment of melanoma. *Clin. Endocrinol. (Oxf)* **86**, 614–620 (2017).
19. Khan, S.A., Pruitt, S.L., Xuan, L. & Gerber, D.E. Prevalence of autoimmune disease among patients with lung cancer: implications for immunotherapy treatment options. *JAMA Oncol.* **2**, 1507–1508 (2016).
20. Yu, K.H., Beam, A.L. & Kohane, I.S. Artificial intelligence in healthcare. *Nat. Biomed. Eng.* **2**, 719–731 (2018).
21. Chen, Y.J., Chang, Y.T., Wang, C.B. & Wu, C.Y. The risk of cancer in patients with rheumatoid arthritis: a nationwide cohort study in Taiwan. *Arthritis Rheum.* **63**, 352–358 (2011).
22. Gibbons, R.D., Coca Perraillon, M., Hur, K., Conti, R.M., Valuck, R.J. & Brent, D.A. Antidepressant treatment and suicide attempts and self-inflicted injury in children and adolescents. *Pharmacoepidemiol. Drug Saf.* **24**, 208–214 (2015).
23. Mahoney, K.M., Freeman, G.J. & McDermott, D.F. The next immune-checkpoint inhibitors: PD-1/PD-L1 blockade in melanoma. *Clin. Ther.* **37**, 764–782 (2015).
24. Ansari, M.J. *et al*. The programmed death-1 (PD-1) pathway regulates autoimmune diabetes in nonobese diabetic (NOD) mice. *J. Exp. Med.* **198**, 63–69 (2003).
25. US Food and Drug Administration. Supplement Approval <https://www.accessdata.fda.gov/drugsatfda_docs/appletter/2018/125554Orig1s048s049s050s051s052s061s062s064s065s066ltr.pdf> (2015). Accessed June 13, 2018.
26. US Food and Drug Administration. Pembrolizumab (KEYTRUDA) Checkpoint Inhibitor <https://www.fda.gov/drugs/informationondrugs/approveddrugs/ucm526430.htm> (2015). Accessed June 13, 2018.
27. Kong, Y.C. & Flynn, J.C. Opportunistic autoimmune disorders potentiated by immune-checkpoint inhibitors anti-CTLA-4 and anti-PD-1. *Front Immunol.* **5**, 206 (2014).
28. Brahmer, J. *et al*. Nivolumab versus docetaxel in advanced squamous-cell non-small-cell lung cancer. *N. Engl. J. Med.* **373**, 123–135 (2015).
29. Kennedy-Martin, T., Curtis, S., Faries, D., Robinson, S. & Johnston, J. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* **16**, 495 (2015).
30. Makady, A., de Boer, A., Hillege, H., Klungel, O. & Goettsch, W. What is real-world data? A review of definitions based on literature and stakeholder interviews. *Value Health* **20**, 858–865 (2017).
31. Song, H. *et al*. Association of stress-related disorders with subsequent autoimmune disease. *JAMA* **319**, 2388–2400 (2018).
32. Quest Diagnostics. ICD-9 CM to ICD-10 Common Codes Related to Autoimmune <https://www.questdiagnostics.com/dms/Documents/Other/CPT-2015/ICD_9-10_Codes_Autoimmune-MI4954.pdf>(2015). Accessed September 19, 2018.
33. Yu, K.H. *et al*. Data-driven analyses revealed the comorbidity landscape of tuberous sclerosis complex. *Neurology* **91**, 974–976 (2018).
34. Castro, V.M. *et al*. Evaluation of matched control algorithms in EHR-based phenotyping studies: a case study of inflammatory bowel disease comorbidities. *J. Biomed. Inform.* **52**, 105–111 (2014).
35. Field, C.A. & Welsh, A.H. Bootstrapping clustered data. *J. R. Stat. Soc. Series B Stat. Methodol.* **69**, 369–390 (2007).