Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Comparative Real-World Safety Profiles of Pembrolizumab Plus Axitinib Versus Pembrolizumab Plus Lenvatinib in Advanced Renal Cell Carcinoma: A Pharmacovigilance Study

zihan Guo
First Affiliated Hospital of Dalian Medical University
Dandan Guo
First Affiliated Hospital of Dalian Medical University
Bufu Tang
Fudan University
Na Hang
First Affiliated Hospital of Dalian Medical University
Chenyu Gao
First Affiliated Hospital of Dalian Medical University
Xiaonan Rong
First Affiliated Hospital of Dalian Medical University
Jiadi Han
Dalian Medical University
Peng Luo
Southern Medical University
Ruiqing Gao
First Affiliated Hospital of Dalian Medical University
Zhijun Shen
First Affiliated Hospital of Dalian Medical University
Qing Li
First Affiliated Hospital of Dalian Medical University
Chun Wang
First Affiliated Hospital of Dalian Medical University
Lei Li
First Affiliated Hospital of Dalian Medical University
Zhao Xie
First Affiliated Hospital of Dalian Medical University
Bin Zhang
First Affiliated Hospital of Dalian Medical University

Ling Wang

wangling@firsthosp-dmu.com

First Affiliated Hospital of Dalian Medical University

Research Article

Keywords: Advanced renal cell carcinoma, Pembrolizumab, Axitinib, Lenvatinib, Pharmacovigilance, Adverse events, Immune-related adverse events, Disproportionality analysis

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11307010/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Comparative Real-World Safety Profiles of Pembrolizumab Plus Axitinib Versus Pembrolizumab Plus Lenvatinib in Advanced Renal Cell Carcinoma: A Pharmacovigilance Study

Zihan Guo[1#], Dandan Guo[1#], Bufu Tang[2], Na Hang[1], Chenyu Gao[1], Xiaonan Rong[1], Jiadi Han[1], Peng Luo[3], Ruiqing Gao[1], Zhijun Shen[1], Qing Li[1], Chun Wang[1], Lei Li[1], Zhao Xie[1], Bin Zhang[1*], Ling Wang[1*]

1.Department of Oncology, First Affiliated Hospital of Dalian Medical University, Dalian 116011, China.
2.Department of Interventional Radiology, Zhongshan Hospital, Shanghai Institute of Medical Imaging, National Clinical Research Center of Interventional Medicine, Fudan University, Shanghai, China.
3.Department of Oncology, Zhujiang Hospital, Southern Medical University, Guangzhou, Guangdong, 510280, China.

#These authors contribute equally to this work.
***Corresponding Author:**
Bin Zhang, Department of Oncology, First Affiliated Hospital of Dalian Medical University, Dalian 116011, China. E-mail: zhangbin_dlmu@163.com
Ling Wang, Department of Oncology, First Affiliated Hospital of Dalian Medical University, Dalian 116011, China. E-mail: wangling@firsthosp-dmu.com

Abstract

Objective: Pembrolizumab plus axitinib (P+A) and pembrolizumab plus lenvatinib (P+L) are already used as first-line standard treatments for advanced renal cell carcinoma (aRCC); however, their real-world safety and the toxicity patterns that each regimen tends to cause have not been clearly and systematically explained. This study aimed to compare the adverse event (AE) spectra of these two immune checkpoint inhibitor-tyrosine kinase inhibitor (ICI-TKI) combinations; the main focus was on toxicity features specific to each regimen, differences in immune-related adverse events (irAEs), and time-to-onset patterns.
Methods: This study collected individual case safety reports of aRCC patients who received P+A or P+L from two independent pharmacovigilance cohorts; it used three methods for imbalance analysis, namely the reporting odds ratio (ROR), the proportional reporting ratio (PRR), and the Bayesian confidence propagation neural network (BCPNN); the rules for finding signals were set as follows: the lower limit of the 95% confidence interval (CI) of ROR > 1, PRR ≥ 2 with $\chi^2 \geq 4$, and $IC_{025} > 0$; the analysis looked at both the System Organ Class (SOC) level and the Preferred Term (PT) level, with special attention to ten clinically relevant irAEs; these were also divided by sex, age group, and time of onset.
Results: A total of 897 and 648 reports were included in the P+A and P+L groups, respectively; both regimens covered a wide range of SOCs, but their safety profiles showed clear differences; P+A had more obvious signals in hepatobiliary disorders and certain immune-related toxicities, especially stronger signals for Immune-mediated hepatic disorder and Immune-mediated myocarditis; in contrast, P+L was mainly linked to vascular,

renal, metabolic, and endocrine toxicities, with higher signals for Hypertension, Proteinuria, Palmar-plantar erythrodysaesthesia syndrome, and Immune-mediated hypothyroidism; subgroup analyses showed that some immune toxicities changed with age, for example, P+A-related Immune-mediated hepatic disorder had stronger signals in younger patients, while the P+L-related Hypothyroidism signals was more obvious in older patients; analysis of temporal onset showed different paths for different irAEs, such as very early start, middle-stage peak, and late build-up patterns. Several potential novel signals, including Tumor necrosis and Immune-mediated hyperthyroidism, were identified.
Conclusion: Pembrolizumab plus axitinib and pembrolizumab plus lenvatinib exhibited markedly distinct safety profiles in the treatment of aRCC. The P+A regimen was primarily characterized by hepatobiliary and cardiac-specific immunotoxicities, whereas the P+L regimen was more closely associated with vascular, renal, metabolic, and endocrine toxicities. Age stratification and temporal profiling provided novel evidence to inform personalized monitoring strategies. Because spontaneous reporting databases have intrinsic constraints, our observations should be confirmed in prospective clinical investigations.
Keywords: Advanced renal cell carcinoma; Pembrolizumab; Axitinib; Lenvatinib; Pharmacovigilance; Adverse events; Immune-related adverse events; Disproportionality analysis

1. Introduction

Renal cell carcinoma (RCC) exhibits intrinsic sensitivity to immunotherapy and anti-angiogenic therapy while displaying resistance to cytotoxic chemotherapy. Accordingly, first-line management of aRCC has evolved from VEGF-TKI monotherapy to ICI-based combinations, supported by evidence that VEGF signaling contributes to an immunosuppressive tumor microenvironment[1,2]. The KEYNOTE-426 trial demonstrated that pembrolizumab plus axitinib (P+A) significantly prolonged overall survival (OS) and progression-free survival (PFS) while improving the objective response rate (ORR) compared with sunitinib monotherapy in treatment-naive aRCC patients[3,4]. Similarly, pembrolizumab plus lenvatinib (P+L) has emerged as a cornerstone first-line treatment option. The CLEAR trial demonstrated that this combination significantly improved PFS, along with OS and ORR, compared with sunitinib. European Association of Urology (EAU) guidelines have incorporated both P+A and P+L as preferred first-line treatment options[5,6]. Extended follow-up data further support the sustained clinical benefit of these two combinations in terms of PFS and ORR[3,7].

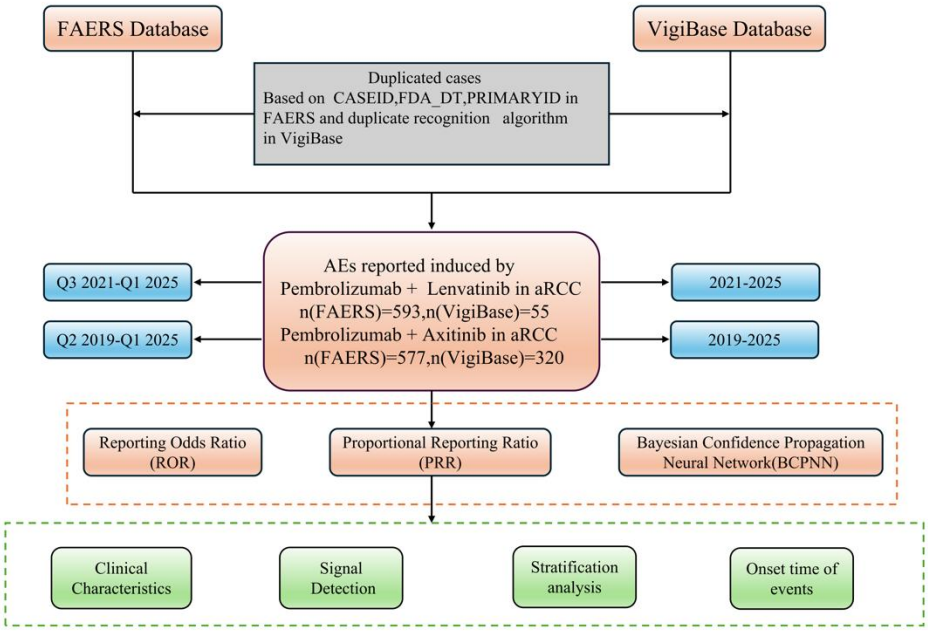
ICI-TKI combinations show notable efficacy; however, grade 3–5 toxicities occur fairly often with these treatments, which makes safety management a key issue in clinical practice[8,9]. Take pembrolizumab plus lenvatinib as an example: during long-term follow-up, treatment-emergent adverse events (TEAEs) appear quite frequently, and diarrhoea tends to be reported most often among them. Axitinib and lenvatinib are both classified as VEGFR-TKIs, yet their target profiles are not the same; axitinib mainly acts on VEGFR1–3, while lenvatinib, besides VEGFR1–3, also targets FGFR1–4, PDGFRα, RET, and KIT[10]. Pembrolizumab serves as the shared immunotherapy backbone in both regimens; still, because axitinib and lenvatinib bind to quite different targets, the toxicities linked to each regimen may differ accordingly.

While RCTs establish efficacy and characterize frequent toxicities, they often lack sufficient power to reliably detect uncommon, late-onset, or population-specific AEs[11]. In brief, Pharmacovigilance encompasses the systematic detection and evaluation of ADRs, with the aim of understanding and preventing medication-related harm[12]. The READUS guidelines for pharmacovigilance (READUS-PV) explicitly state that disproportionality analysis of individual case safety report (ICSR) data is primarily utilized to generate signals of disproportionate reporting[13,14]. A previous pharmacovigilance study based on the FDA Adverse Event Reporting System (FAERS) database reported post-marketing safety differences between P+A and P+L. That study performed comparisons at the System Organ Class (SOC) level and calculated reporting odds ratios (RORs), revealing differences between the two regimens in cardiac, hepatobiliary, blood and lymphatic, metabolic and nutritional, and vascular disorders[15]. Nevertheless, existing studies are constrained by reliance on a single database and broad adverse event classifications, lacking systematic, multi-dimensional comparisons. Furthermore, given that real-world databases vary in patient demographics, reporting mechanisms, and geographic coverage, cross-validation across multiple databases can substantially enhance the reliability of drug safety signal detection. Therefore, using FAERS together with VigiBase, we applied three complementary approaches (ROR, PRR, and BCPNN)[16,17] to compare P+A and P+L across SOC and PT-level patterns, irAEs, subgroup heterogeneity, time-to-onset features, and candidate emerging signals. We sought to define how AE profiles differ between these combinations, highlight regimen enriched safety signals, and examining potential differences in toxicity phenotypes, thereby providing real-world evidence to guide individualized regimen selection and monitoring in advanced RCC.

2. Methods

2.1 Data Sources

Data for patients with aRCC treated with pembrolizumab plus axitinib or pembrolizumab plus lenvatinib were retrieved from the FAERS and VigiBase databases. FAERS is an open FDA repository of ICSRs that includes report-level demographic and clinical information and can be used for large-scale safety signal analyses[18]. VigiBase is the WHO global ICSR database maintained by the Uppsala Monitoring Centre (UMC) and aggregates ADR reports submitted by more than 150 countries[19]. The data extraction timeframes were determined based on the respective U.S. Food and Drug Administration (FDA) approval dates for the first-line indication in aRCC: pembrolizumab plus axitinib (P+A) was approved in April 2019, and reports from Q2 2019 to Q1 2025 were extracted; pembrolizumab plus lenvatinib (P+L) was approved in August 2021, and reports from Q3 2021 to Q1 2025 were extracted. Following data screening and deduplication, the final analysis included: 577 cases in FAERS and 320 cases in VigiBase for the P+A cohort; 593 cases in FAERS and 55 cases in VigiBase for the P+L cohort. All comparisons between the two regimens used every eligible case that appeared after each regimen was approved for marketing, and the analyses did not restrict cases to a common time window; because P+A had a longer post-marketing observation period, this approach may introduce reporting bias related to time, and this limitation is addressed in the Discussion section.



2.2 Case Identification

The workflow for case screening and analysis is shown in the Flowchart; in this study, the target indication was advanced renal cell carcinoma, and it was defined by Preferred Terms (PTs) from the Medical Dictionary for Regulatory Activities (MedDRA), namely "metastatic renal cell carcinoma" and "renal cell carcinoma stage IV," while patients received either pembrolizumab plus axitinib or pembrolizumab plus lenvatinib; to make sure all relevant data were captured, the search query used both generic names and brand names, including pembrolizumab (Keytruda®), axitinib (Inlyta®), and lenvatinib (Lenvima®); each database was deduplicated separately; for the FAERS database, the deduplication algorithm recommended by the FDA was used to find and remove duplicate ICSRs based on case identification number (CASE ID), FDA receipt date (FDA_DT), and manufacturer report identification number (MFRID), and when multiple versions of the same case existed, the most recent version with the highest CASE ID was retained. For VigiBase, deduplication was performed using the platform's built-in duplicate-detection algorithm. Following screening and deduplication, eligible cases were analyzed for clinical baseline characteristics, disproportionality signal detection, stratified subgroup analyses, and time-to-onset patterns of adverse events.

Adverse events were coded using MedDRA PTs. The hierarchical structure of MedDRA comprises five levels, with the PT serving as a distinct descriptor for a single medical concept (e.g., symptom, sign, or diagnosis). This hierarchical structure also incorporates High-Level Terms (HLTs) and High-Level Group Terms (HLGTs), which are further aggregated into System Organ Classes (SOCs). Adverse event data recorded in the VigiBase database were classified according to MedDRA version 25.1 or the International Classification of Diseases (ICD-8, ICD-9, or ICD-10). Disproportionality analysis was utilized to evaluate all adverse event reports and identify severe adverse events to assess their clinical burden. We classified seriousness according to FDA definitions, including outcomes such as death, life-threatening events, hospitalization, disability, or congenital anomalies[20].

2.3 Disproportionality Analysis

To identify drug safety signals, three established disproportionality algorithms were employed[16,17]: the reporting odds ratio (ROR), the proportional reporting ratio (PRR), and the Bayesian confidence propagation neural network (BCPNN). Among these, the ROR is particularly effective in mitigating bias associated with low-frequency events, the PRR exhibits higher specificity in signal detection, and the BCPNN effectively identifies signals with limited data[21,22]. Signal detection thresholds were defined as follows: the lower limit of the 95% confidence interval (CI) of the ROR > 1 with at least three cases ($n \geq 3$); PRR ≥ 2, chi-square (χ^2) ≥ 4, and $n \geq 3$; and the lower limit of the information component (IC_{025}) > 0 for the BCPNN. Disproportionality analyses were conducted using 2 × 2 contingency tables, and the corresponding formulas are detailed in Supplementary Table 1. All data were processed and analyzed using Microsoft Excel 2024 (version 16.90; Microsoft Corp, Redmond, WA, USA) and R software (version 4.5.1; R Foundation for Statistical Computing, Vienna, Austria). Visualizations were generated using R software (version 4.5.1).

2.4 Temporal Analysis

The time-to-onset of adverse events was defined as the duration in days from the initiation of combination therapy to the onset of the adverse event. Reports with negative time intervals were excluded from the temporal analyses. Due to insufficient valid onset-time records for individual categories of immune-related adverse events, it was not feasible to construct a stable restricted cubic spline (RCS) logistic regression model; therefore, RCS modeling was not performed. Descriptive temporal analyses of all adverse events were conducted using time-stratified stacked bar charts and box plots[23]. To further visualize the temporal dynamics of immune-related adverse events, the two treatment cohorts (P+A and P+L) were pooled to achieve adequate sample size. Smoothed onset-probability curves with 95% confidence intervals were then generated for four key irAEs with sufficient valid onset-time records[24]. The remaining six predefined immune-related adverse events did not yield stable fitted models due to inadequate sample sizes and were thus excluded from curve fitting. Direct comparisons between the two regimens were not performed in the pooled analysis.

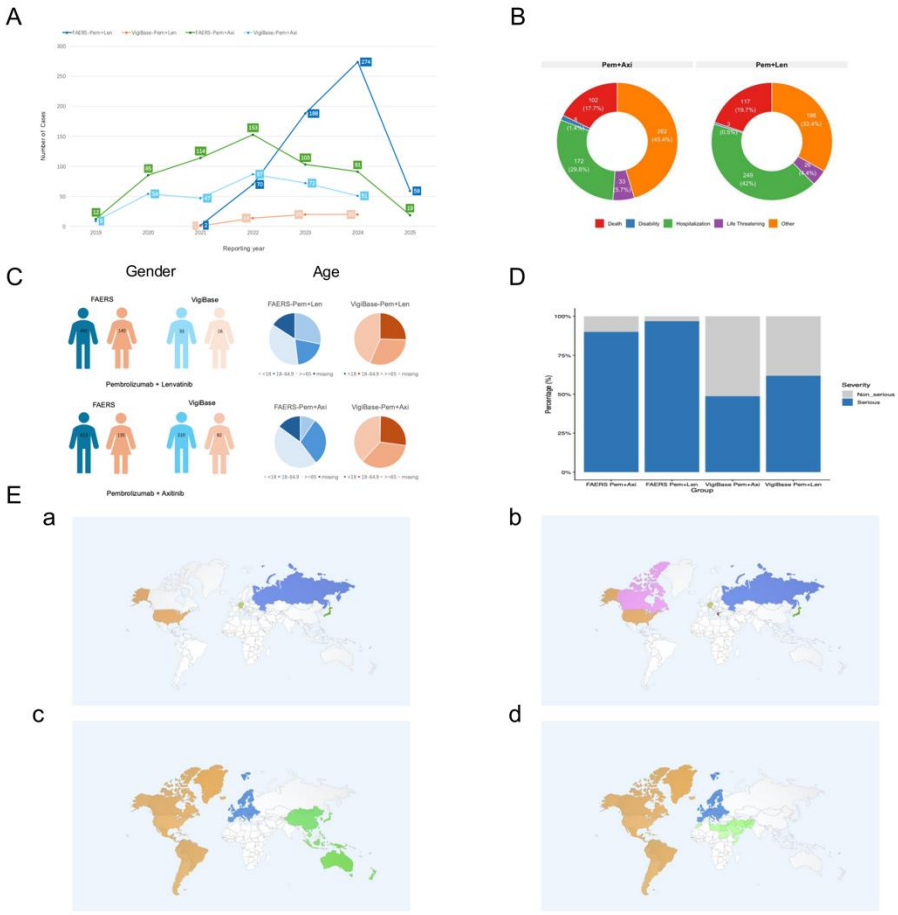
3. Results

3.1 Study Population and Reporting Characteristics

The demographic characteristics, reporting sources, geographic distributions, and temporal reporting patterns for both regimens are presented in Figure 1, with comprehensive baseline characteristics detailed in Table 1.

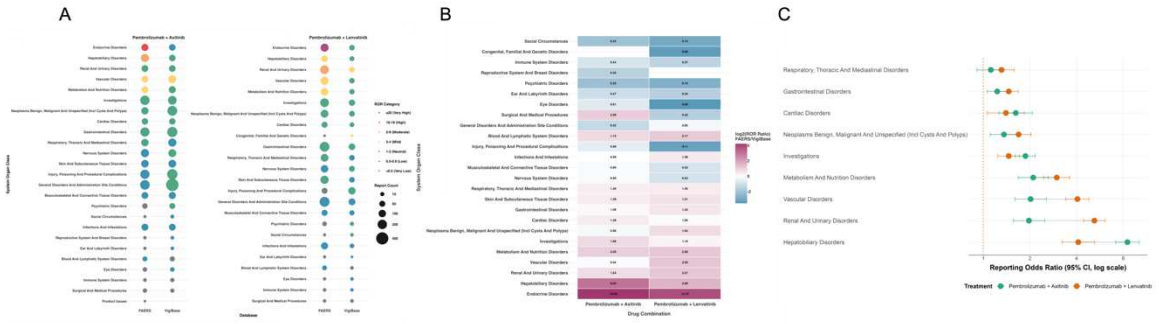
	Pembrolizumab + Lenvatinib		**Pembrolizumab + Axitinib**	
	FAERS	VigiBase	FAERS	VigiBase
	(N=593)	(N=55)	(N=577)	(N=320)
Gender				
Female	140 (23.6%)	16 (29%)	135 (23.4%)	82 (25.6%)
Male	442 (74.5%)	33 (60%)	413 (71.6%)	210 (65.6%)
missing	11 (1.9%)	6 (10.9%)	29 (5.0%)	28 (8.7%)
Age(year)				
< 18	167 (28.2%)	NA	55 (9.5%)	1 (0.3%)

18 ~ 64.9	118 (19.9%)	14 (25.4%)	174 (30.2%)	86 (26.8%)
>= 65	214 (36.1%)	17 (30.8%)	262 (45.4%)	111 (34.6%)
missing	94 (15.9%)	24 (23.6%)	86 (14.9%)	122 (38.1%)
REPORTER_COUNTRY(top 3)				
	JP 450 (75.9%)	EUR 37 (67.2%)	JP 183 (31.7%)	AMR 126 (39.3%)
	US 43 (7.3%)	AMR 15 (27.2%)	US 134 (23.2%)	EUR 121 (37.8%)
	DE 27 (4.6%)	WPR 3 (5.4%)	CA 86 (14.9%)	EMR 46 (14.3%)
Outcome				
Death	117 (19.7%)	Not recovered/not resolved 7 (12.7%)	102 (17.7%)	Not recovered/not resolved 23 (7.1%)
Disability	3 (0.5%)	Recovered/resolved 8 (14.5%)	8 (1.4%)	Recovered/resolved 57 (17.8%)
Hospitalization	249 (42.0%)	Recovering/resolving 8 (14.5%)	172 (29.8%)	Recovering/resolving 23 (7.1%)
Life-Threatening	26 (4.4%)	Fatal 2 (3.6%)	33 (5.7%)	Fatal 11 (3.4%)
Other	198 (33.4%)	Other 27 (49%)	262 (45.4%)	Other 206 (64.2%)
Seriousness				
NO	18 (3.0%)	21 (38.1%)	57 (9.9%)	164 (51.2%)
YES	575 (97.0%)	34 (61.8%)	520 (90.1%)	156 (48.7%)
Reportertype				
Consumer	14 (2.4%)	Report from study 14 (26.9%)	118 (20.5%)	Report from study 14 (4.3%)
Health Professional	23 (3.9%)	Spontaneous 38(73%)	108 (18.7%)	Spontaneous 306(95.6%)
Pharmacist	23 (3.9%)		29 (5.0%)	
Physician	532 (89.7%)		318 (55.1%)	
Missing	1 (0.2%)		4 (0.7%)	
GetDataYear				
2019			12 (2.1%) Q2	9 (2.8%)
2020			85 (14.7%)	54 (16.8%)
2021	2 (0.3%) Q3	1 (1.8%)	114 (19.8%)	47 (14.6%)
2022	70 (11.8%)	14 (25.4%)	153 (26.5%)	87 (27.1%)
2023	188 (31.7%)	20 (36.3%)	103 (17.9%)	72 (22.5%)
2024	274 (46.2%)	20 (36.3%)	91 (15.8%)	51 (15.9%)
2025	59 (9.9%) Q1	NA	19 (3.3%) Q1	NA



After pembrolizumab plus axitinib (P+A) was approved in 2019, the number of reports for this regimen went up each year, and it reached its highest point in 2022 (n_{FAERS} = 153 [26.5%]; $n_{VigiBase}$ = 87 [27.1%]); in a similar way, after pembrolizumab plus lenvatinib (P+L) was approved in 2021, the number of reports for this regimen rose steadily, then it peaked in 2024 (n_{FAERS} = 274 [46.2%]; $n_{VigiBase}$ = 20 [36.3%]), and after that it fell in 2025. The reporting trends for both regimens aligned with their respective post-marketing time windows (Figure 1A). Figures 1B illustrates the severity distribution for the P+A and P+L cohorts, respectively. In the P+L cohort, adverse events requiring hospitalization accounted for the highest proportion (n_{FAERS} = 249 [42.0%]), followed by other non-serious events (n_{FAERS} = 198 [33.4%]) and deaths (n_{FAERS} = 117 [19.7%]). In contrast, the P+A cohort was predominantly composed of other non-serious events (n_{FAERS} = 262 [45.4%]), followed by adverse events requiring hospitalization (n_{FAERS} = 172 [29.8%]) and deaths (n_{FAERS} = 102 [17.7%]). Serious adverse events, such as deaths and hospitalization, constituted a higher proportion of reports in the P+L cohort. Both the P+L and P+A cohorts were predominantly male across both databases (n_{FAERS} = 442 [74.5%], $n_{VigiBase}$ = 33 [60.0%]; and n_{FAERS} = 413 [71.4%], $n_{VigiBase}$ = 210 [65.6%], respectively). The majority of patients were ≥65 years of age (P+L: n_{FAERS} = 214 [36.1%], $n_{VigiBase}$ = 17 [30.8%]; P+A: n_{FAERS} = 262 [45.4%], $n_{VigiBase}$ = 111 [34.6%]), with minor variations in age-stratified composition across databases and regimens (Figure 1C). In the FAERS database, the proportion of serious adverse events substantially exceeded that of non-serious events in both cohorts; conversely, in VigiBase, the proportion of non-serious adverse events increased, narrowing the gap between serious

and non-serious events. This suggests a shift in reported severity composition between databases (Figure 1D). Figure 1E demonstrates that P+A reports originated primarily from North America and Europe. In contrast, P+L reports in the FAERS database were predominantly from Japan (n_{FAERS} = 450 [75.9%]), consistent with the high enrollment of Japanese patients in the CLEAR trial and the widespread clinical use of lenvatinib in Japan[25]. Meanwhile, P+L reports in VigiBase spanned a broader geographical distribution across Asia, Europe, and Oceania.



3.2 SOC-Level Safety Profile

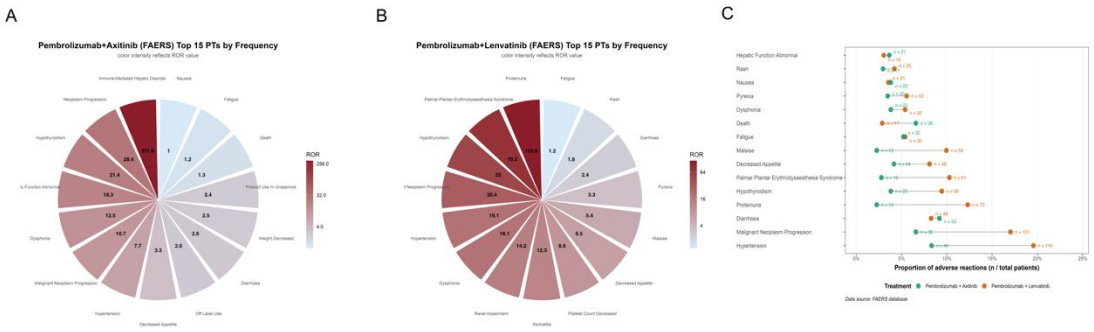
The overall adverse event profiles of pembrolizumab plus axitinib (P+A) and pembrolizumab plus lenvatinib (P+L) were analyzed at the System Organ Class (SOC) level (Figure 2). Both regimens were associated with a wide spectrum of adverse events. In the FAERS database, the most frequently reported SOC was General disorders and administration site conditions (n_{P+A} = 211, ROR = 0.55; n_{P+L} = 234, ROR = 0.62), whereas the strongest signal was detected for Endocrine disorders (P+A: ROR = 16.17 [12.99–20.12]; P+L: ROR = 24.84 [20.70–29.70]). In the VigiBase database, the most frequently reported SOC was Injury, poisoning and procedural complications (n_{P+A} = 266, ROR = 1.08; n_{P+L} = 266, ROR = 1.08), and the strongest signal was detected for Congenital, familial and genetic disorders (P+A: ROR = 2.69 [0.98–7.37]; P+L: ROR = 2.43 [0.33–17.67]). Both regimens also exhibited high reporting frequencies and elevated ROR signals for SOCs such as Respiratory, thoracic and mediastinal disorders; Gastrointestinal disorders; and Neoplasms benign, malignant and unspecified (incl cysts and polyps).

Despite substantial overlap in their adverse event profiles, distinct regimen-specific differences were observed. In the FAERS database, the P+L regimen demonstrated significantly stronger disproportionality signals for the following SOCs: Endocrine disorders (ROR = 24.84 vs. 16.17), Renal and urinary disorders (ROR = 5.21 vs. 1.97), Vascular disorders (ROR = 4.05 vs. 2.02), and Metabolism and nutrition disorders (ROR = 2.98 vs. 2.10). Additionally, Respiratory, thoracic and mediastinal disorders and Skin and subcutaneous tissue disorders reached positive signals exclusively in the P+L cohort, whereas these SOCs did not reach the positive threshold in the P+A cohort. Conversely, the P+A regimen exhibited a higher signal intensity for Hepatobiliary disorders (ROR = 8.51 vs. 4.10). In the VigiBase database, the number of SOCs reaching positive signals decreased markedly, consistent with the smaller sample size (P+L: n = 55; P+A: n = 320). For the P+L regimen, only Renal and urinary disorders (ROR = 2.30 [1.63–3.25]), Vascular

disorders (ROR = 1.76 [1.24–2.49]), Endocrine disorders (ROR = 1.81 [1.28–2.57]), and Injury, poisoning and procedural complications (ROR = 1.50 [1.18–1.90]) achieved statistical significance. For the P+A regimen, the signals that were positive included Vascular disorders (ROR = 2.16 [1.79–2.61]), Neoplasms benign, malignant and unspecified (incl cysts and polyps) (ROR = 1.54 [1.33–1.78]), General disorders and administration site conditions (ROR = 1.32 [1.17–1.49]), and Psychiatric disorders (ROR = 1.38 [1.03–1.86]); Vascular disorders stayed positive in both regimens and both databases, which suggests a strong signal; for the P+L regimen, Endocrine disorders and Renal and urinary disorders stayed positive in both databases; however, the strong Hepatobiliary disorders signal seen for P+A in FAERS (ROR = 8.51) was not found again in VigiBase (ROR = 1.06), probably because the sample size was limited; in general, the direction of SOC signals was mostly the same between FAERS and VigiBase, but signal sizes in VigiBase were weaker, because sample sizes were smaller and confidence intervals were wider.

3.3 PT-Level Adverse Event Profile

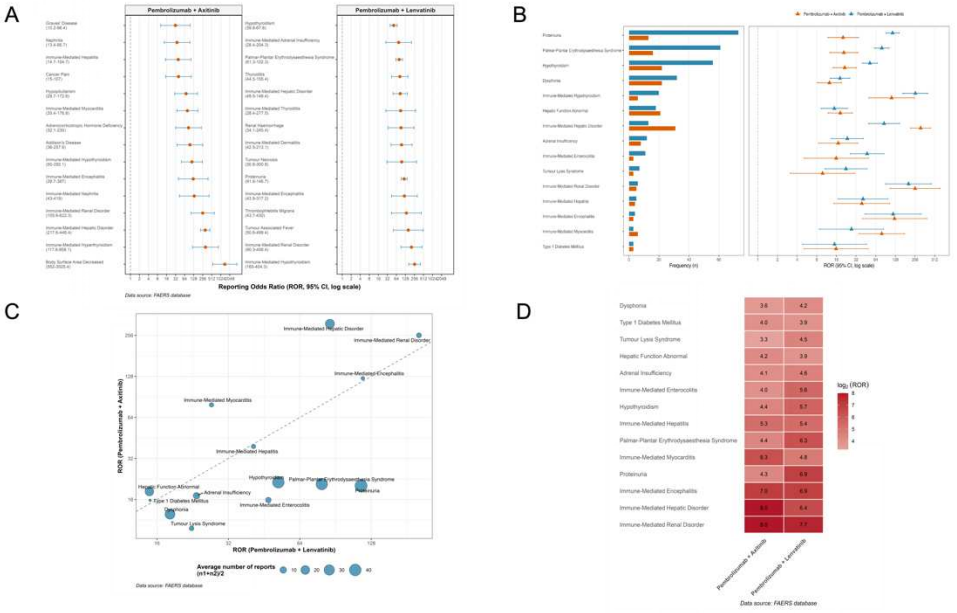
All PT-level analyses and figures used data only from the FAERS database; even though two-database checking was done at the SOC level in Section 3.2, the VigiBase sample size was not enough for PT-level analysis. When disaggregated to the PT level, the number of cases per PT was too small to yield robust positive disproportionality signals. This sample size constraint also applied to all subsequent PT-level analyses (including immune-related adverse event analyses and stratified subgroup analyses); consequently, these analyses were conducted using solely the FAERS database, with cross-database validation achievable only at the SOC level. At the PT level, analyses examined two dimensions: reporting frequency and disproportionality signal strength.



3.3.1 High-Frequency Reported PTs

Figures 3A–B display the top 15 most frequently reported PTs for both regimens. For the P+A regimen, the most frequently reported PT was Off label use (n = 96), followed by Diarrhoea (n = 53), Hypertension (n = 48), Neoplasm progression (n = 45), and Death (n = 38). The relatively high reporting counts for Off label use and Product use in unapproved indication (n = 30) suggest a notable off-label signal for this regimen. Among toxicity-related PTs, Immune-mediated hepatic disorder (n = 31), Hepatic function abnormal (n = 21), and Hypothyroidism (n = 22) were the most frequently reported. For the P+L regimen,

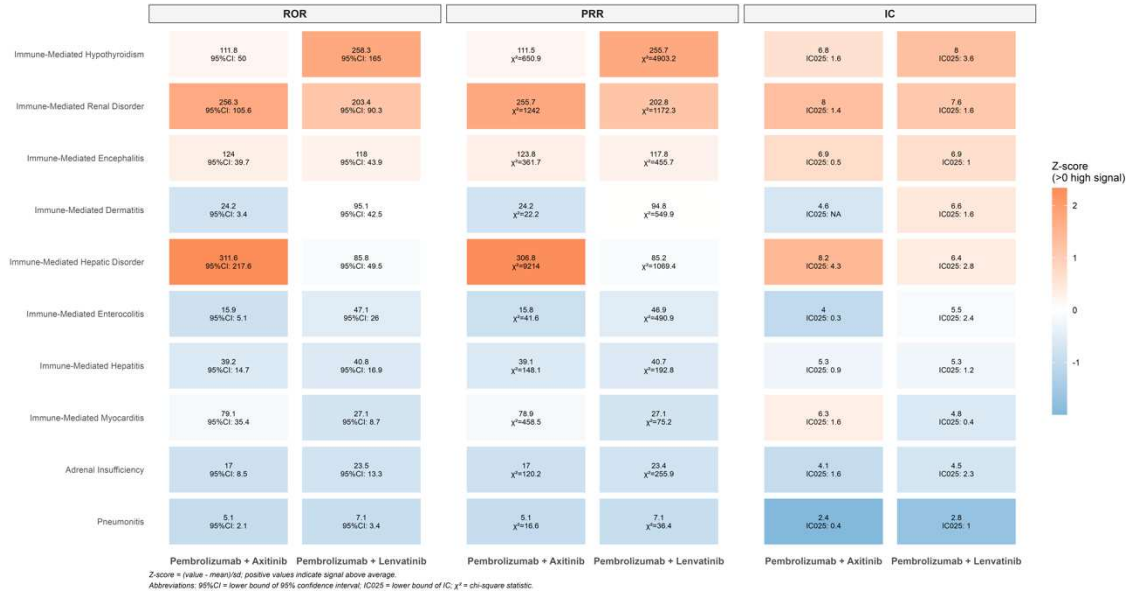
Hypertension was the most frequently reported PT (n = 116), followed by Malignant neoplasm progression (n = 101), Proteinuria (n = 73), and Palmar-plantar erythrodysaesthesia syndrome (n = 61). Among shared PTs, the reporting frequency for Diarrhoea was comparable (P+A: n = 53; P+L: n = 49). However, several PTs exhibited marked differences in reporting frequency: Hypertension (48 vs. 116), Proteinuria (13 vs. 73), Palmar-plantar erythrodysaesthesia syndrome (16 vs. 61), and Malaise (13 vs. 59) were reported more frequently in the P+L regimen than in the P+A regimen. Conversely, Death (38 vs. 17) and Hepatic function abnormal (21 vs. 18) were more frequent in the P+A cohort (Figure 3C). Reporting frequency and disproportionality signal strength were not equivalent (Figures 3A–B). Some PTs with high reporting frequencies did not reach statistical significance; for example, Death in the P+A regimen (n = 38, ROR = 1.34, 95% CI: 0.97–1.84) and Fatigue in the P+L regimen (n = 32, ROR = 1.25, 95% CI: 0.88–1.78). Conversely, Immune-mediated hepatic disorder in the P+A cohort ranked only 7th by reporting frequency (n = 31), yet demonstrated an extraordinarily high ROR of 311.65 (217.60–446.36). Meanwhile, PTs such as Proteinuria (115.94 [91.62–146.72]) and Hypertension (19.09 [15.82–23.03]) in the P+L regimen exhibited both high reporting frequencies and strong signals.



3.3.2 PT-Level Disproportionality Signals

This analysis examined PTs that tested positive and ranked them by ROR for both regimens, along with shared PTs detected in both cohorts (Figures 4A–D). The results demonstrate that high-signal adverse events in the P+A and P+L regimens exhibit distinct distributions dominated by immunotoxicity and targeted therapy-induced toxicity, respectively. A detailed comparison of immune-related adverse event PTs is presented in Section 3.4. The P+A cohort was significantly enriched for liver-related adverse events, among which Immune-mediated hepatic disorder was the most prominent signal (ROR = 311.65), far exceeding that of the P+L cohort (ROR = 85.75). Similarly, Immune-mediated myocarditis exhibited a stronger disproportionality signal in the P+A regimen (ROR = 79.13 vs. 27.15). Additionally, the strongest signals for the P+A regimen were Body surface area decreased

(ROR = 1394.93 [551.95–3525.36]), followed by Immune-mediated hyperthyroidism (ROR = 317.62 [117.57–858.11]). Consistent with the known toxicity profile of lenvatinib, the P+L cohort demonstrated higher signal intensities for Proteinuria (ROR = 115.94 vs. 20.19), Palmar-plantar erythrodysaesthesia syndrome (ROR = 79.22 vs. 20.77), and Immune-mediated hypothyroidism (ROR = 258.31 vs. 111.82) compared with the P+A cohort. Multiple positive disproportionality signals were detected at the PT level for both regimens. Hypothyroidism was a prominent shared signal (P+A: ROR = 21.42 [14.07–32.62]; P+L: ROR = 51.96 [39.81–67.84]). Other shared immunotoxicity signals included Immune-mediated enterocolitis, Adrenal insufficiency, and Immune-mediated encephalitis. For non-immunotoxicity signals, Dysphonia demonstrated stable positive disproportionality signals in both cohorts.

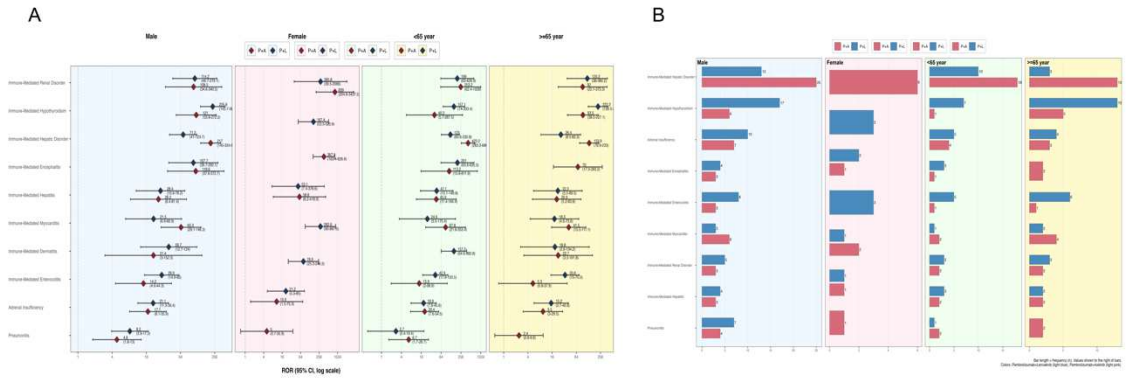


3.4 Immune-Related Adverse Events

Both regimens exhibited strong signals for immune-related adverse events (irAEs), which represent the core safety concern of anti-PD-1 therapy combined with anti-angiogenic tyrosine kinase inhibitors. For irAEs, we cross-checked signals using three indices (ROR, PRR, and IC) to improve robustness (Figure 5). We labeled an irAE as a positive signal only when it satisfied the prespecified criteria for all three metrics. Immune-mediated renal disorder and Immune-mediated encephalitis were shared irAEs with the strongest signals and similar magnitudes between both regimens. For Immune-mediated renal disorder, the signals in the P+A and P+L cohorts were ROR = 256.33 [95% CI: 105.59–622.29], PRR (χ^2) = 255.69 (1241.95), IC (IC_{025}) = 7.97 (1.36); and ROR = 203.42 [95% CI: 90.28–458.36], PRR (χ^2) = 202.81 (1172.34), IC (IC_{025}) = 7.62 (1.65), respectively, demonstrating significantly elevated values across all three indices. The signals for Immune-mediated encephalitis were nearly identical between both cohorts (ROR = 118.00 vs. 123.97), demonstrating highly consistent positive signals. Strong signals for Hypothyroidism and liver-related immunotoxicity were also detected in both cohorts, though their magnitudes differed substantially between the regimens.

Several irAEs exhibited prominent regimen-specific differences. Immune-mediated

hepatic disorder was the most prominent regimen-specific irAE for the P+A regimen: its ROR was 311.65 (95% CI: 217.60–446.36, n = 31), approximately 3.6 times that of the P+L cohort (ROR = 85.75, 95% CI: 49.55–148.42, n = 13). The PRR (306.83 vs. 85.20) and IC (8.22 vs. 6.40) were substantially elevated, with cross-validation results across all three indices showing high consistency, corroborating the PT-level hepatic toxicity findings in Section 3.3.2. The signal for Immune-mediated myocarditis was also stronger in the P+A cohort (ROR = 79.13 vs. 27.15, approximately 2.9-fold). Immune-mediated hypothyroidism was the most prominent irAE for the P+L regimen: its ROR was 258.31 (95% CI: 165.03–404.33, n = 20), approximately 2.3 times that of the P+A cohort (ROR = 111.82, 95% CI: 49.99–250.13, n = 6). The PRR (255.73 vs. 111.49) and IC (7.95 vs. 6.79) consistently supported a stronger signal for the P+L regimen. Additionally, the signals for Immune-mediated enterocolitis and Immune-mediated dermatitis were significantly higher in the P+L cohort (ROR = 47.14 vs. 15.86 and ROR = 95.10 vs. 24.19, respectively). Overall, irAE signals in the P+L regimen were dominated by thyroid and gastrointestinal immunotoxicity. For several classic anti-PD-1-related irAEs, no clear differences were observed between both cohorts. For instance, the signals for Immune-mediated hepatitis were nearly identical between both regimens (ROR = 40.84 vs. 39.18; IC = 5.34 vs. 5.28); the signal strength for Adrenal insufficiency was moderate (ROR = 23.48 vs. 17.05); and Pneumonitis exhibited the lowest signal strength in both cohorts (ROR = 7.08 vs. 5.13). This suggests that these events may be primarily associated with PD-1 inhibition rather than with the specific VEGFR-TKI partner.



3.5 Subgroup Analyses of Selected irAEs

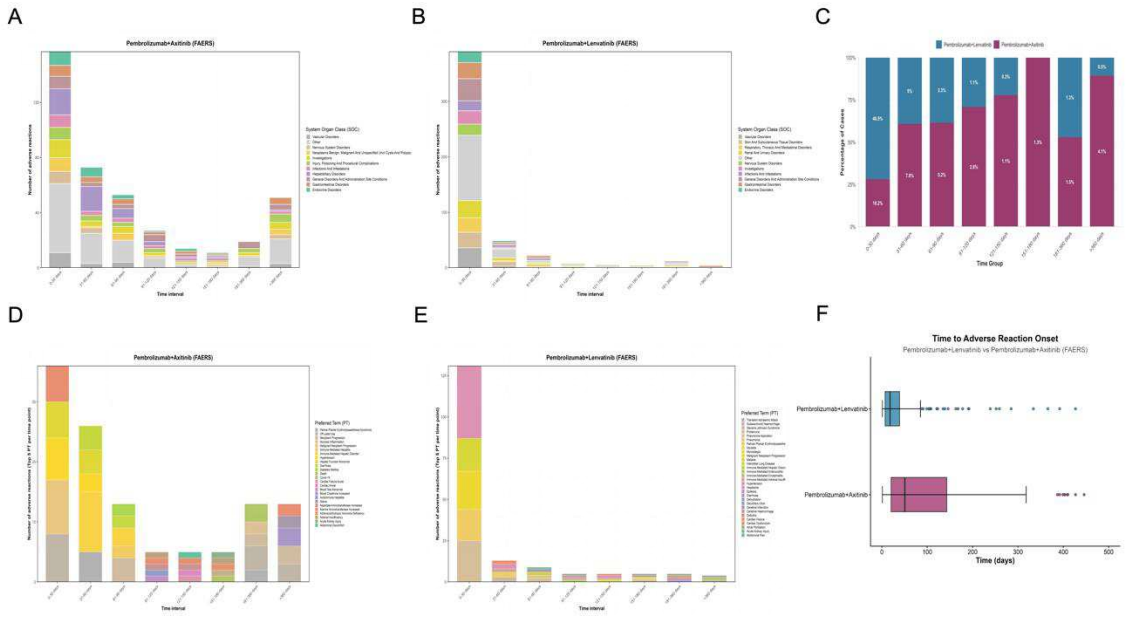
To systematically elucidate the population distribution heterogeneity of positive irAEs, stratified subgroup analyses were conducted according to age (<65 years / ≥65 years) and sex (male / female). To avoid case dilution and increased false-positive rates caused by subgroup disaggregation, analyses were not performed indiscriminately on all irAE PTs. Instead, the analysis focused on clinically relevant irAEs validated as positive by all three indices in Section 3.4, systematically comparing signal strength and distribution characteristics across demographic subgroups, and summarizing common patterns and regimen-specific differences. Subgroup signals were evaluated using the information component (IC) and its lower limit of the 95% confidence interval (IC_{025}), with $IC_{025} > 0$ defined as a positive signal. Subgroup signal comparisons are shown in Figure 6. Common features included positive signals ($IC_{025} > 0$) in both male and female populations for the

vast majority of irAEs across both regimens, revealing no sex-specific immunotoxicity. Endocrine-related irAEs were more prominent in the elderly population (≥65 years), whereas organ-specific irAEs were more prominent in patients <65 years across both regimens. Positive signals across all subgroups were consistent with overall trends, further verifying the robustness of immunotoxicity signals associated with both regimens.

Sex-stratified characteristics: Among the positive irAEs, endocrine-related irAEs (primarily Hypothyroidism) exhibited strong signals across both sexes in the P+L regimen, with signal strengths in males (IC = 5.67, IC_{025} = 4.11, n = 42) and females (IC = 5.65, IC_{025} = 2.69, n = 13) largely comparable, showing no significant sex differences. Furthermore, signals were markedly higher than those in the corresponding P+A subgroups. Conversely, liver-related irAEs in the P+A regimen showed no significant sex differences, though their signal strengths were lower than those of endocrine irAEs in the P+L regimen.
Age-stratified characteristics: The positive irAEs exhibited distinct age-related patterns between regimens. Endocrine-related irAEs in the P+L regimen were more prominent in the elderly population (≥65 years), with Hypothyroidism showing an IC of 5.23 (IC_{025} = 3.30, n = 23), significantly higher than in the corresponding P+A subgroup. This was a core irAE signal in the elderly population, whereas no significant enrichment was observed in patients <65 years. In contrast, Immune-mediated hepatic disorder was the most prominent regimen-specific irAE for the P+A regimen, demonstrating clear age-dependent characteristics. In patients <65 years, the IC for Immune-mediated hepatic disorder in the P+A cohort was 8.64 (IC_{025} = 3.50, n = 18), significantly higher than that in the P+L cohort (IC = 6.90, IC_{025} = 2.46, n = 10), with both regimens showing strong positive signals. This irAE did not rank among the top 10 for patients ≥65 years.

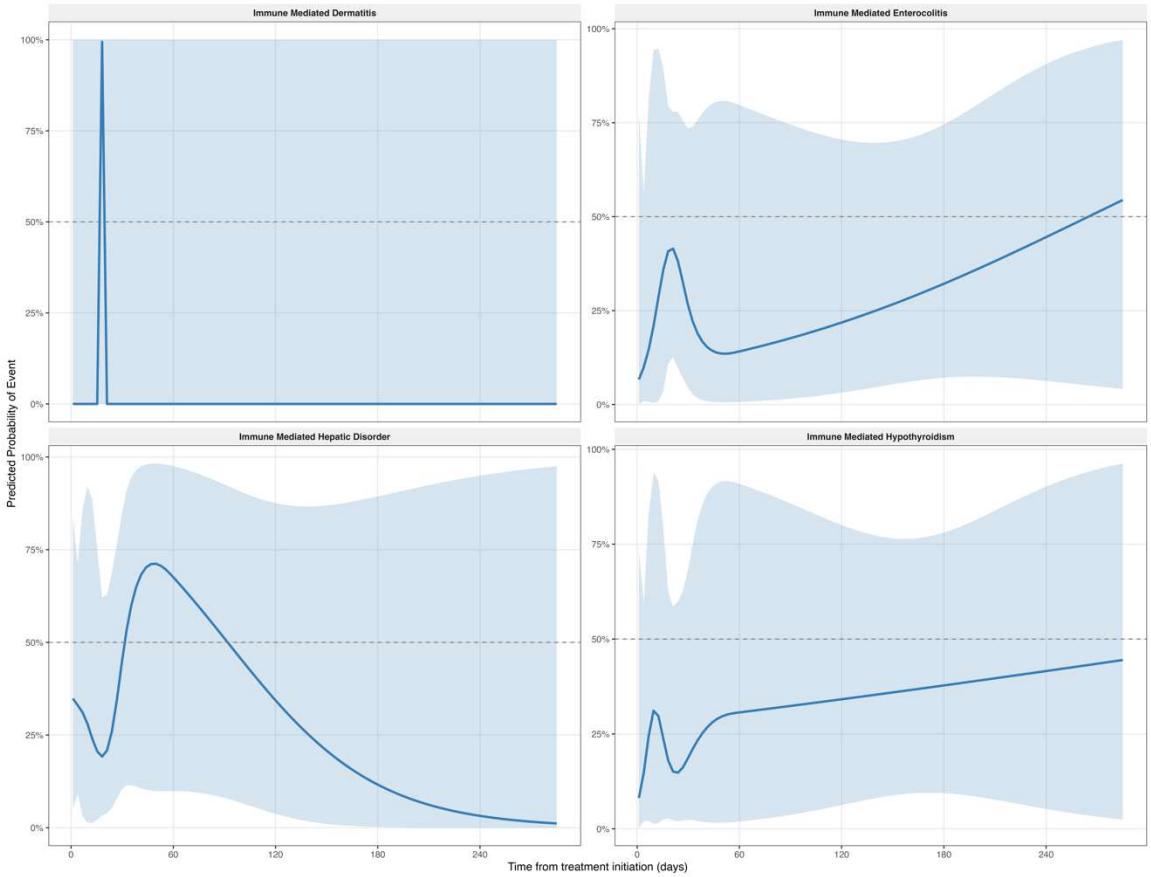
3.6 Time-to-Onset Analysis of Adverse Events

A time-to-onset (TTO) analysis was performed to describe the temporal distribution of adverse events associated with the P+A and P+L regimens. Reports with missing or discrepant dates were excluded. All temporal analyses were based exclusively on the FAERS database, as VigiBase lacks complete time data. Spontaneous reporting data describe the temporal distribution pattern and cannot establish causality.



3.6.1 Overall Temporal Distribution Characteristics

The overall time-to-onset distribution for adverse events is shown in Figures 7A–F. The majority of adverse event reports were concentrated within 0–30 days of treatment initiation (Figure 7C). The proportion of early-onset events (0–30 days) was higher in the P+L regimen, whereas adverse events in the P+A regimen showed a more dispersed temporal distribution, with a substantial proportion reported after 180 days. Box plots (Figure 7F) indicated that the median time-to-onset was longer and the interquartile range broader for the P+A regimen compared with the P+L regimen. At the SOC level (Figures 7A–B), reporting peaks were predominantly concentrated within 0–60 days post-initiation. For the P+A regimen, hepatobiliary adverse events persisted across multiple time windows, with substantial reports in Hepatobiliary disorders, Gastrointestinal disorders, and General disorders and administration site conditions still observed at 180 and >360 days. For the P+L regimen, General disorders and administration site conditions, Vascular disorders, Renal and urinary disorders, and Gastrointestinal disorders were the most frequently reported within 0–30 days. Reports declined rapidly after 60 days, with very few after 180 days. At the PT level (Figures 7D–E), high-frequency adverse events exhibited early-onset clustering within the first 60 days. For the P+A regimen, reports of Immune-mediated hepatic disorder continued to accumulate during 31–90 days, consistent with subacute-phase onset of immune-mediated hepatotoxicity observed clinically. For the P+L regimen, classic targeted toxicities such as Hypertension, Proteinuria, and Palmar-plantar erythrodysaesthesia syndrome were primarily concentrated within 0–30 days. A small fraction of adverse events was reported >180 days post-initiation in both regimens.



3.6.2 Temporal Risk Patterns of Selected irAEs

Due to limited sample sizes, time-to-event prediction models were constructed for four irAEs: Immune-mediated dermatitis, Immune-mediated enterocolitis, Immune-mediated hepatic disorder, and Immune-mediated hypothyroidism. Predicted event probability curves with 95% confidence intervals for these four irAEs within 250 days of treatment initiation are shown in Figure 8. The four irAEs exhibited distinct temporal risk patterns. Immune-mediated dermatitis presented a sharp, very early risk peak, with predicted probability surging within the first few days and then declining rapidly. Event risk then remained low, suggesting high early incidence of skin-related irAEs. Immune-mediated enterocolitis exhibited a biphasic risk profile: an early risk peak within 30 days, followed by decline and then steady increase, with an upward trajectory persisting after 180 days, indicating continuous risk accumulation during long-term treatment. The risk peak for Immune-mediated hepatic disorder occurred at approximately 60 days post-treatment, after which predicted probability gradually declined, reaching low levels by approximately 240 days. This single-peak pattern suggests subacute onset, with the high-risk window in the second month of treatment. Event risk for Immune-mediated hypothyroidism showed a slow, continuous increase. Following a small transient peak in the first month, the risk curve climbed steadily throughout the observation period with no decline, suggesting that risk gradually increases with prolonged treatment duration, representing delayed-onset endocrine immunotoxicity.The four irAEs exhibited distinct temporal risk patterns: very early onset (Immune-mediated dermatitis), subacute peak (Immune-mediated hepatic disorder), biphasic risk (Immune-mediated enterocolitis), and delayed cumulative risk (Immune-mediated hypothyroidism).

4. Discussion

In this drug safety study, this study used two databases and three imbalance analysis methods to broadly assess the safety of pembrolizumab plus axitinib (P+A) and pembrolizumab plus lenvatinib (P+L) as first-line treatment for advanced renal cell carcinoma (aRCC), and obtained several key findings; first, because both regimens share a PD-1 inhibitor backbone and a VEGFR inhibition mechanism, their toxicity profiles overlap a lot, but their adverse event features were still clearly different at both the System Organ Class (SOC) and Preferred Term (PT) levels; second, P+A showed stronger imbalance signals for hepatobiliary toxicity, especially Immune-mediated hepatic disorder, while P+L showed stronger signals for vascular, renal, and metabolic events, which matches the wider tyrosine kinase inhibition range of Lenvatinib[6,10]; third, irAE analysis showed risk patterns that differed by organ, as P+A had much stronger signals for Immune-mediated hepatitis and Immune-mediated myocarditis, whereas P+L was more clearly linked to Hypothyroidism, Immune-mediated enterocolitis, and Immune-mediated dermatitis[26-28]; fourth, subgroup and time-based analyses showed differences related to age and sex, and they also showed different patterns for when major irAEs started; overall, these findings go beyond existing clinical trials and drug safety studies, offer new information for patients who receive these regimens, and can support risk grouping and clinical care.

The clear overlap in side effect profiles comes from the shared treatment backbone; both regimens pair pembrolizumab, an immune checkpoint inhibitor (ICI) that acts on the PD-1/PD-L1 axis, with a strong VEGFR-TKI; when VEGF-mediated immune suppression and the PD-1 signaling pathway are blocked together, T-cell activity goes up, but class-effect toxicities are also amplified in a synergistic way, including Fatigue, Diarrhoea, Hypothyroidism, and different organ-specific irAEs[8,9,29]; this common drug principle explains the shared signals seen at the SOC level, such as Gastrointestinal disorders, General disorders and administration site conditions, and Metabolism and nutrition disorders; however, the different TKI parts drive clear differences in toxicity profiles; axitinib is a highly selective inhibitor of VEGFR-1, -2, and -3, so its anti-angiogenic effect is more focused; by contrast, lenvatinib has a broader multi-kinase profile, and it strongly inhibits FGFR1–4, PDGFRα, RET, and KIT in addition to VEGFR, so it brings in a wider range of toxicities[10,30]; this drug difference explains the distinct safety features seen in this study—the P+L regimen showed a greater burden of classic TKI-related toxicities, such as Hypertension, Proteinuria, and Palmar-plantar erythrodysaesthesia syndrome, which reflects its multifaceted effect on vascular and tissue homeostasis[1,31].

At the PT level, the analysis did not only repeat known safety signals from key trials and drug labels, but it also showed some possible differences in toxicity types between the two regimens; in the P+L group, strong signals appeared for Hypertension (ROR = 19.09), Proteinuria (ROR = 115.94), and Palmar-plantar erythrodysaesthesia syndrome (ROR = 79.22), and these findings are in line with what the CLEAR trial reported, where incidence rates were clearly higher than with sunitinib[6,9,32]; the P+A regimen, in contrast, showed more obvious immune-related organ toxicity, especially Immune-mediated hepatic disorder (ROR = 311.65 vs. 85.75), and this signal was clearly higher than the one seen in

the P+L group; this finding supports what came from the KEYNOTE-426 trial and earlier drug safety studies[3,4,15], and it suggests that hepatobiliary toxicity is a characteristic feature of P+A that may need priority in management; beyond these already known safety signals, several possible signals appeared that have not been prominently described for these combination therapies before; in the P+L regimen, Tumor-associated fever (ROR = 158.96, 95% CI: 50.60–499.36), Thrombophlebitis migrans (ROR = 137.10, 95% CI: 43.72–429.98), and Tumor necrosis (ROR = 96.21, 95% CI: 30.77–300.79) all had quite high ROR values; although case numbers here were fairly small, and both Tumor-associated fever and Tumor necrosis can also occur with cancer progression itself, these signals still seem to deserve some attention; for the P+A regimen, the strong signal for Body surface area decreased (ROR = 1394.93, 95% CI: 551.95–3525.36) is a rare and somewhat non-specific finding; it may be confounded by malignancy and disease progression itself, but it still points to the need for clinical vigilance; both regimens also showed signals for rare Immune-mediated endocrinopathies, such as Addison's disease, Adrenocorticotropic hormone deficiency, and Graves' disease; although these endocrine manifestations are already known to occur with anti-PD-1 monotherapy, they are rarely discussed in the setting of TKI-ICI combination therapies; taken together, the PT-level analysis helps outline the broader safety picture here: P+L appears more closely tied to the typical, often acute, toxicities expected from potent multi-targeted TKI therapy, while P+A shows more pronounced immune-mediated organ toxicities, especially in the liver; these findings provide real-world evidence that could help with risk identification and individualized toxicity monitoring for ICI-TKI combination regimens in the future, though the possible signals found here still need further validation before firm conclusions can be drawn.

The analysis that focused on irAEs found immune toxicities that were specific to each regimen, and these findings are quite important for clinical decisions; both regimens shared strong signals for Immune-mediated renal disorder and Immune-mediated encephalitis, which suggests that these may be class-effect irAEs linked to PD-1 inhibition together with TKI co-administration, and they may not depend much on which specific TKI partner is used[28]; even so, differences also appeared in other irAEs, and these differences had real clinical meaning; in the P+A regimen, Immune-mediated hepatitis and Immune-mediated myocarditis were notably enriched, which points to a synergistic induction of liver and heart autoimmunity; for the mechanism behind this, it may involve selective VEGFR blockade caused by axitinib, and this could change the liver microenvironment in ways that make it more open to T-cell infiltration caused by PD-1 inhibitors[33,34]; on the other hand, the P+L regimen showed stronger signals for endocrine (Hypothyroidism), gastrointestinal (Immune-mediated enterocolitis), and skin (Immune-mediated dermatitis) irAEs; in the P+L group, the signal for Hypothyroidism was especially high (ROR = 258.31), and it seems to stand out more in real-world data than it did in clinical trials; this may reflect the direct effects of lenvatinib on thyroid physiology, acting through RET and FGFR inhibition, which may worsen immune-mediated thyroiditis.

Subgroup analysis also showed that age and sex may lead to some differences in how likely patients are to have irAEs; consistent with what is already known about RCC epidemiology, male patients made up most of the cohort (71%–74% in FAERS), which reflects the higher

prevalence generally seen in men; across both sexes, most irAE signals stayed fairly strong, and this suggests that irAEs caused by ICI-TKI combinations are mostly not dependent on sex, so routine monitoring seems reasonable for both sexes; age-related patterns still appeared, though: for P+A, Immune-mediated hepatitis had a stronger signal in patients under 65, while for P+L, endocrine irAEs, especially Hypothyroidism, had a stronger signal in patients 65 and older; the strong signal for P+A-associated Immune-mediated hepatitis in patients under 65 (IC_{025} = 3.50) is a fairly new finding; younger age is often linked to stronger immune function, and that could plausibly increase the risk of ICI-related irAEs, but this specific interaction with axitinib still needs more study; even though these findings are based on a limited number of cases, they do suggest that monitoring intensity for specific irAEs could be tailored by age, which would add a layer of individualization to toxicity management overall.

Temporal patterns can give an important basis for clinical monitoring; most side effects tended to gather within the first 60 days of treatment, and this shows that closer monitoring is really needed during this early period; the P+L group showed a more noticeable early side effect peak, and this fits with its strong ability to block many kinases; by contrast, the P+A treatment plan had a somewhat more spread-out pattern, and side effects were still reported even after 180 days, which may be linked to delayed immune-related side effects, such as Immune-mediated hepatitis. Curves that show when events happen are quite useful here; Immune-mediated dermatitis reached a fast peak within only the first few days, which shows why skin monitoring right away matters; Immune-mediated enterocolitis, meanwhile, showed a two-phase risk pattern, which suggests that watchfulness may need to be kept both early on and later during continuing treatment; Immune-mediated hepatic disorder showed a middle-stage peak around day 60 or so, which indicates that liver function should be watched closely during the first few months of treatment; lastly, Hypothyroidism risk went up slowly but steadily over time, which suggests that initial testing plus ongoing thyroid tests would be worthwhile. Although these curves are mainly descriptive and depend on a limited number of events, they still offer some early evidence that supports monitoring plans which change over time.

These findings have real meaning for clinical work; because the safety profiles of the two regimens are clearly different, when doctors choose a regimen, they probably should not decide only by how well the drug works, but should also think about the patient's own risk factors and how well the patient is likely to tolerate treatment; for patients who already have liver problems or a higher risk of autoimmune hepatitis, P+L may be the better choice, because it can help avoid serious liver-related immune side effects; on the other hand, for patients with heart and blood vessel risk, clear high blood pressure, or protein in urine, P+A may be the better choice, since its signs of blood vessel and kidney toxicity are relatively lower.Monitoring plans that are adjusted according to risk seem needed here; for patients taking P+A, close checking of liver function and heart symptoms is important, especially in the first several months; for patients taking P+L, careful control of blood pressure, kidney function, and hormone-related indicators becomes important, together with early attention to skin and stomach or bowel symptoms; patient education about side effects specific to each regimen and side effects shared by both also plays a key role, because it helps patients notice and report problems early, supports them in managing their own care,

and allows timely help from different medical teams when needed.

This study has several strengths that are worth mentioning; first, it is the most complete drug safety comparison between P+A and P+L so far; the researchers examined safety signals in a systematic way from several angles, including different levels (SOC, PT, and irAEs), demographic subgroups, and time patterns; second, the study used two databases (FAERS and VigiBase) to cross-check results at the SOC level; this helped make the findings more solid and more likely to apply more widely; although the two databases had different reporting groups and sources, the signals still matched, and this made the results more believable[13,35]; third, the study used three methods that work together—ROR, PRR, and BCPNN—and it also used strict cutoffs; this helped reduce the risk of false positives as much as possible; fourth, the study followed the READUS-PV reporting guidelines closely; this made the methods rigorous and reproducible, and it also set a kind of standard for future drug safety studies; finally, the detailed time analysis for irAEs provided fairly new data, and these data could help guide clinical practice.

Still, there are also some limits worth noting; first, like any system that depends on voluntary reports, the findings here may still be affected by cases that are not reported and by reporting bias[22]; second, data quality is somewhat limited, because some safety reports for single cases are not complete, such as missing basic patient information or information on when side effects started, and there may also be mixed effects from the condition being treated, disease stage, and other medicines taken at the same time; third, the imbalance analysis mainly helps produce ideas, since it can find statistical links, but it cannot prove cause and effect by itself. Fourth, because approval dates differed between the two regimens, observation periods and sample sizes ended up unequal, which could introduce some time-related reporting bias. The smaller P+L sample size within VigiBase, in particular, limited how much PT-level cross-validation and robust subgroup analysis could be done. Sample size constraints also meant that temporal analysis had to be restricted to just a subset of irAEs. Finally, the subgroup findings and exploratory signals identified here—especially those built on a limited number of cases—still need validation through independent real-world cohorts before they can be fully trusted.

5. Conclusions

Both regimens share certain common toxicities, which come from combining a PD-1 inhibitor with anti-angiogenic therapy; even so, their safety profiles still differ quite a bit. For P+A, hepatobiliary and immune-mediated toxicity signals stand out more; for P+L, vascular, renal, metabolic, and endocrine toxicities are more dominant instead. Immune-mediated renal disorder and encephalitis appear to be shared immunotoxicities across these ICI-TKI combinations, while hepatic, thyroid, and gastrointestinal immunotoxicities seem to be regimen-specific. Different irAEs follow different timing patterns; this suggests that clinical monitoring should probably be adjusted according to each specific toxicity. When choosing an ICI-TKI regimen for aRCC, clinicians need to think about patient comorbidities, organ function, and susceptibility to toxicity, not just efficacy alone. Going forward, prospective studies along with mechanistic research are still needed, so as to confirm these safety patterns and help move toward more precise risk management in aRCC.

Abbreviations

AEs Adverse Events
FAERS Adverse Event Reporting System
FDA Food and Drug Administration
HLGT High Level Group Term
HLT High Level Term
MedDRA Medical Dictionary for Regulatory Activities
PT Preferred Term
RCS Restricted Cubic Spline
SOC System Organ Classification
aRCC Advanced Renal Cell Carcinoma
P+A Pembrolizumab Plus Axitinib
P+L Pembrolizumab Plus Lenvatinib

Figure Legends

Flowchart.
The flowchart illustrates data extraction and processing from the FAERS and VigiBase databases, including screening, deduplication, exclusion criteria, and final case numbers. Abbreviations: FAERS, FDA Adverse Event Reporting System.

Figure 1. Baseline characteristics of included individual case safety reports.
(A) This panel shows annual reporting trends for P+L and P+A, drawn from both FAERS and VigiBase.
(B) Adverse event seriousness is shown here for both regimens; seriousness categories cover death, life-threatening events, disability, hospitalization, and other non-serious outcomes.
(C) Sex and age distribution are presented; age groups were split into <18 years, 18–64 years, ≥65 years, plus a missing-data category.
(D) The proportion of serious events (blue) versus non-serious events (gray) is compared across four cohorts.
(E) Geographic distribution of reports is shown for four cohorts, namely FAERS P+L, FAERS P+A, VigiBase P+L, and VigiBase P+A.

Figure 2. Disproportionality analysis at the System Organ Class (SOC) level.
(A) A bubble plot compares ROR values alongside case numbers for different SOCs, across FAERS and VigiBase; bubble size reflects case numbers, while color indicates ROR signal strength.
(B) Heatmap of $\log_2$(ROR) signal intensity for each SOC in both regimens. Red represents signal enhancement and blue represents signal attenuation.
(C) Forest plot comparing RORs and 95% CIs between regimens(logarithmic scale).

Figure 3. High-frequency Preferred Terms (PTs) in FAERS.
(A-B) Round heat maps show the top 15 PTs by report frequency for P+A and P+L; color strength shows ROR signal strength.
(C) A dumbbell plot shows PTs shared between the two regimens, with case numbers(n)

marked.

Figure 4. Disproportionality signal analysis at the PT level in FAERS.
(A) Two-column forest plots of RORs and 95% CIs for the top 15 PTs ordered by ROR for P+A and P+L (log scale).
(B-D) Shared PTs found in both regimens (top 15 ranked by ROR):
(B) Combination plot; the left panel shows case numbers, while the right panel shows RORs and 95% CIs between regimens (log scale).
(C) Scatter plot comparing RORs between regimens; the horizontal axis is P+L ROR, and the vertical axis is P+A ROR; bubble size shows average case numbers; a dashed diagonal line helps with comparison.
(D) Heat map of $\log_2$(ROR) signal strength for each PT; darker colors mean stronger signals.

Figure 5. Multi-algorithm disproportionality analysis of immune-related adverse events.
The heat map shows ROR, PRR, and IC values for each irAE in both regimens; cell values show effect sizes; ROR and PRR cells show values and 95% CIs, while IC cells show IC values and IC_{025}; colors show Z scores, and a value above 0 means the signal is above average.

Figure 6. Subgroup analysis of immune-related adverse events in FAERS.
(A) Forest plots show IC and IC_{025} values for both regimens across sex and age subgroups for 10 irAEs; $IC_{025} > 0$ means a positive signal.
(B) Bar charts show report frequency proportions for both regimens across sex and age subgroups for 10 irAEs.

Figure 7. Time distribution of adverse events in FAERS.
(A-B) Stacked bar charts of SOC-level adverse event counts across time intervals for P+A and P+L.
(C) Proportion of cases grouped by time-to-onset for both regimens.
(D-E) Stacked bar charts of the top 5 PTs across time intervals for P+A and P+L.
(F) Box plots of time-to-onset distribution (days) for both regimens.

Figure 8. Temporal risk curves for four immune-related adverse events in FAERS.
Solid line: predicted event probability; shaded area: 95% CI. The horizontal axis shows days from treatment start; because sample sizes were limited, modeling was only done for these four irAEs; the curves are descriptive, and they do not prove cause and effect.

Funding statement

This study was supported by the 2023 Wu Jieping Foundation (no. 320.6750.2023-03-54) and Hubei Chen Xiaoping Science and Technology Development Foundation (no. CXPJJH122011-034).

Author Contributions

Zihan Guo and Dandan Guo analyzed the data and wrote the paper. Bufu Tang, Na Hang,

Chenyu Gao, Xiaonan Rong, Jiadi Han, Peng Luo, Ruiqing Gao, Zhijun Shen, Qing Li, Chun Wang, Lei Li, Zhao Xie helped to analyze the data. Bin Zhang and Ling Wang revised the paper and supervised the project.

Ethics approval and consent to participate

NO applicable.

Declaration of competing interest

This research was conducted without any competing financial interests on the part of any author.

Consent for publication

Each contributing author has explicitly confirmed their awareness of the paper's content and approved their co-authorship designation.

Data sharing statement

Interested parties may access the study data by submitting a request to either the first author or the corresponding author.

Provenance and peer review

Not commissioned, externally peer-reviewed.

References

[1] Argentiero A, Solimando A G, Krebs M, et al. Anti-angiogenesis and immunotherapy: novel paradigms to envision tailored approaches in renal cell-carcinoma[J]. Journal of Clinical Medicine, 2020, 9 (5): E1594. DOI:10.3390/jcm9051594.

[2] Finke J H, Rini B, Ireland J, et al. Sunitinib reverses type-1 immune suppression and decreases T-regulatory cells in renal cell carcinoma patients[J]. Clinical Cancer Research, 2008, 14 (20): 6674-6682. DOI:10.1158/1078-0432.CCR-07-5212.

[3] Plimack E R, Powles T, Stus V, et al. Pembrolizumab plus axitinib versus sunitinib as first-line treatment of advanced renal cell carcinoma: 43-month follow-up of the phase 3 KEYNOTE-426 study[J]. European Urology, 2023, 84 (5): 449-454. DOI:10.1016/j.eururo.2023.06.006.

[4] Powles T, Plimack E R, Soulières D, et al. Pembrolizumab plus axitinib versus sunitinib monotherapy as first-line treatment of advanced renal cell carcinoma (KEYNOTE-426): extended follow-up from a randomised, open-label, phase 3 trial[J]. The Lancet. Oncology, 2020, 21 (12): 1563-1573. DOI:10.1016/S1470-2045(20)30436-8.

[5] Bex A, Ghanem Y A, Albiges L, et al. European association of urology guidelines on renal cell carcinoma: the 2025 update[J]. European Urology, 2025, 87 (6): 683-696.

DOI:10.1016/j.eururo.2025.02.020.
[6] Motzer R, Alekseev B, Rha S Y, et al. Lenvatinib plus Pembrolizumab or Everolimus for Advanced Renal Cell Carcinoma[J]. The New England Journal of Medicine, 2021, 384 (14): 1289-1300. DOI:10.1056/NEJMoa2035716.
[7] Rini Brian I., Plimack Elizabeth R, Plimack Elizabeth R, et al. Pembrolizumab plus axitinib versus sunitinib as first-line therapy for advanced clear cell renal cell carcinoma: 5-year analysis of KEYNOTE-426[J]. Journal of Clinical Oncology, [2026][2026-09-03]. https://ascopubs.org/doi/10.1200/JCO.2023.41.17_suppl.LBA4501. DOI:10.1200/jco.2023.41.17_suppl.lba4501.
[8] Motzer R, George S, Merchan J R, et al. Characterization and management of adverse reactions from the CLEAR study in advanced renal cell carcinoma treated with lenvatinib plus pembrolizumab[J]. Oncologist, 2023, 28 (6): 501-509. DOI:10.1093/oncolo/oyac269.
[9] Motzer R J, Porta C, Masatoshi Eto, et al. Lenvatinib plus pembrolizumab versus sunitinib in first-line treatment of advanced renal cell carcinoma: final prespecified overall survival analysis of CLEAR, a phase III study[J]. Journal of Clinical Oncology, 2024, 42 (11): 1222-1228. DOI:10.1200/JCO.23.01569.
[10] Adachi Y, Kamiyama H, Ichikawa K, et al. Inhibition of FGFR reactivates IFNγ signaling in tumor cells to enhance the combined antitumor activity of lenvatinib with anti-PD-1 antibodies[J]. Cancer Research, 2022, 82 (2): 292-306. DOI:10.1158/0008-5472.CAN-20-2426.
[11] Fusaroli M, Salvo F, Begaud B, et al. The reporting of a disproportionality analysis for drug safety signal detection using individual case safety reports in PharmacoVigilance (READUS-PV): development and statement[J]. Drug Safety, 2024, 47 (6): 575-584. DOI:10.1007/s40264-024-01421-9.
[12] CIOMS. CIOMS-Cumulative-PV-Glossary-v1.0.pdf[R]. 2021. https://cioms.ch/wp-content/uploads/2021/03/CIOMS-Cumulative-PV-Glossary-v1.0.pdf.
[13] Fusaroli M, Salvo F, Khouri C, et al. The reporting of disproportionality analysis in pharmacovigilance: spotlight on the READUS-PV guideline[J]. Frontiers in Pharmacology, 2024, 15[2026-09-03]. https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2024.1488725/full. DOI:10.3389/fphar.2024.1488725.
[14] Fusaroli M, Salvo F, Begaud B, et al. The REporting of a disproportionality analysis for DrUg safety signal detection using individual case safety reports in PharmacoVigilance (READUS-PV): explanation and elaboration[J]. Drug Safety, 2024, 47 (6): 585-599. DOI:10.1007/s40264-024-01423-7.
[15] Jun Matsumoto, Naohiro Iwata, Shogo Watari, et al. Adverse events of axitinib plus pembrolizumab versus lenvatinib plus pembrolizumab: a pharmacovigilance study in food and drug administration adverse event reporting system[J]. European Urology Focus, 2023, 9 (1): 141-144. DOI:10.1016/j.euf.2022.07.003.
[16] van Puijenbroek E P, Bate A, Leufkens H G M, et al. A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions[J]. Pharmacoepidemiology and Drug Safety, 2002, 11 (1): 3-10.

DOI:10.1002/pds.668.
[17]Szarfman A，Machado S G，O'Neill R T. Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database[J]. Drug Safety，2002，25（6）：381-392. DOI:10.2165/00002018-200225060-00001.
[18]FDA adverse event reporting system (FAERS) database[J]. Fda，2024[2026-08-30]. https://www.fda.gov/drugs/drug-approvals-and-databases/fda-adverse-event-reporting-system-faers-database.
[19]Li H，Sun X，Sun D，et al. Thromboembolic events associated with immune checkpoint inhibitors: a real-world study of data from the food and drug administration adverse event reporting system (FAERS) database[J]. International Immunopharmacology，2021，98：107818. DOI:10.1016/j.intimp.2021.107818.
[20]What is a serious adverse event?[J]. Fda，2024[2026-08-30]. https://www.fda.gov/safety/reporting-serious-problems-fda/what-serious-adverse-event.
[21]Li H，Wang C，Deng A，et al. A real-world disproportionality analysis of mepolizumab based on the FDA adverse event reporting system[J]. Frontiers in Pharmacology，2023，14[2026-09-03]. https://www.frontiersin.org/journals/pharmacology/articles/10.3389/fphar.2023.1280490/full. DOI:10.3389/fphar.2023.1280490.
[22]Pham M，Cheng F，Ramachandran K. A comparison study of algorithms to detect drug-adverse event associations: frequentist, bayesian, and machine-learning approaches[J]. Drug Safety，2019，42（6）：743-750. DOI:10.1007/s40264-018-00792-0.
[23]Leroy F，Dauxois J Y，Théophile H，et al. Estimating time-to-onset of adverse drug reactions from spontaneous reporting databases[J]. BMC Medical Research Methodology，2014，14（1）：17. DOI:10.1186/1471-2288-14-17.
[24]Li B，Song R，Tang L，et al. Pharmacovigilance analysis of severe cutaneous adverse reactions associated with antiseizure medications: a FAERS database study with time-to-onset evaluation[J]. Frontiers in Pharmacology，2026，17：1848600. DOI:10.3389/fphar.2026.1848600.
[25]Eto M，Takagi T，Kimura G，et al. Lenvatinib plus pembrolizumab versus sunitinib for advanced renal cell carcinoma: japanese patients from the CLEAR study[J]. Cancer Medicine，2023，12（6）：6902-6912. DOI:10.1002/cam4.5483.
[26]Sharaf M，Shah P. Lenvatinib plus pembrolizumab–associated delayed-onset hypothyroidism[J]. AACE Endocrinology and Diabetes，2025，12（1）：54-56. DOI:10.1016/j.aed.2025.02.007.
[27]Isoda B，Shiga M，Kandori S，et al. Prognostic impact of immune-related adverse events on combination immune checkpoint/tyrosine kinase inhibition for metastatic renal cancer[J]. Cancer Research Communications，2025，5（9）：1681-1689. DOI:10.1158/2767-9764.CRC-25-0337.
[28]Ishihara H，Nemoto Y，Nakamura K，et al. Comparison of the impact of immune-related adverse events due to immune checkpoint inhibitor dual combination therapy and

immune checkpoint inhibitor plus tyrosine kinase inhibitor combination therapy in patients with advanced renal cell carcinoma[J]. Targeted Oncology，2023，18（1）：159-168. DOI:10.1007/s11523-022-00940-8.
[29]Tannir N，Hammers H，Amin A. First-line vascular endothelial growth factor targeted therapy in renal cell carcinoma: priming the tumor microenvironment for immunotherapy[J]. Current Medical Research and Opinion，2018，34（5）：825-831. DOI:10.1080/03007995.2018.1423960.
[30]Xu H，Tu B，Li H，et al. From chronic inflammation to immune escape: mapping the tumor microenvironment evolution in renal cell carcinoma[J]. Frontiers in Immunology，2025，16[2026-09-03]. https://www.frontiersin.org/journals/immunology/articles/10.3389/fimmu.2025.1683844/full. DOI:10.3389/fimmu.2025.1683844.
[31]Gu L，Zhang Q，Liang Q，et al. PD-1 blockade plus tyrosine kinase inhibitor remodels the tumor microenvironment in advanced renal cell carcinoma[J]. Nature Communications，2026，17（1）：4626. DOI:10.1038/s41467-026-70978-z.
[32]Kato T，Mizuno R，Miyake H. Prevalence and management of proteinuria associated with vascular endothelial growth factor receptor-targeted tyrosine kinase inhibitor treatment in advanced renal cell carcinoma, hepatocellular carcinoma, and thyroid cancer[J]. International Journal of Urology，2024，31（5）：465-474. DOI:10.1111/iju.15409.
[33]Rini B I，Atkins M B，Plimack E R，et al. Characterization and management of treatment-emergent hepatic toxicity in patients with advanced renal cell carcinoma receiving first-line pembrolizumab plus axitinib. Results from the KEYNOTE-426 trial[J]. European Urology Oncology，2022，5（2）：225-234. DOI:10.1016/j.euo.2021.05.007.
[34]Nishimoto K，Shirotake S，Miyama Y，et al. Case study: a japanese patient with metastatic renal cell carcinoma who achieved long-term treatment-free survival with pembrolizumab and axitinib in the KEYNOTE-426 phase III trial of pembrolizumab and axitinib versus sunitinib[J]. IJU Case Reports，2022，5（3）：149-152. DOI:10.1002/iju5.12422.
[35]Kobatake K，Goto K，Fujita A，et al. Real-world effectiveness and safety of pembrolizumab–lenvatinib versus axitinib-based regimens in metastatic clear cell renal cell carcinoma: a multicenter retrospective study[J]. International Journal of Urology，2026，33（4）：e70461. DOI:10.1111/iju.70461.