Journal for ImmunoTherapy of Cancer

Identification of conserved immune-related adverse event risk factors and clinical outcomes in a pan-immunotherapy data mart

David F Lamparter,[1] Virginia C Schmid,[1] Rajat Mohindra ,[2] Vaios Karanikas ,[3] Tony Kam-Thong ,[1] Petar Scepanovic,[1] Guillemette Duchateau-Nguyen,[1] Andreas Roller ,[1] Dominik Heinzmann,[2] Cameron Adams,[4] Sarah L Mycroft,[5] Benjamin P Fairfax ,[6] Nicolas Staedler [1]

To cite: Lamparter DF, Schmid VC, Mohindra R, *et al*. Identification of conserved immune-related adverse event risk factors and clinical outcomes in a pan-immunotherapy data mart. *Journal for ImmunoTherapy of Cancer* 2025;**13**:e012891. doi:10.1136/jitc-2025-012891

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jitc-2025-012891).

DFL, VCS and NS contributed equally.

BPF and NS are joint senior authors.

Accepted 22 October 2025

For numbered affiliations see end of article.

Correspondence to
Dr Nicolas Staedler; nicolas.staedler@roche.com

Professor Benjamin P Fairfax; benjamin.fairfax@oncology.ox.ac.uk

ABSTRACT

Background Cancer immunotherapy (CIT) often triggers immune-related adverse events (irAEs). Analysis of irAEs in large checkpoint inhibitor (CPI) trials has enhanced their management and demonstrated their prognostic value for treatment outcome. However, data on irAEs in non-standard CITs are limited, and systematic exploration is lacking. Identifying predictive biomarkers for irAEs in these therapies is still emerging and essential for improving patient care.

Methods We established a harmonized data mart from 27 early-phase CIT trials, encompassing 14 molecules with diverse mechanisms across various cancer indications. This dataset includes 3,608 patients, both CPI-naïve and CPI-experienced, with detailed information on clinical data, tumor characteristics, soluble biomarkers, and genome-wide genotyping. We examined the occurrence of different irAEs and CIT molecules concerning incidence, severity, and onset. A meta-analysis was conducted to assess the association between risk factors and the time to onset of irAEs. Finally, we explored the predictive value of irAEs for clinical outcomes, specifically measured by progression-free survival (PFS).

Results Our analysis reveals significant variation in irAE incidence and kinetics across CIT molecules. Common irAEs include hepatitis, rash, acute kidney injuries, and hypothyroidism, with hepatitis often severe and others mild. Hepatitis is frequently associated with immunocytokine treatment, while T-cell bispecifics are linked to organ-specific toxicities. Hepatic metastases correlate with hepatitis but inversely with rash; elevated liver enzymes are associated with hepatitis, and high ferritin levels with acute kidney injury risk. Higher myeloid cell counts are associated with reduced rash likelihood. No tumor microenvironment associations were found, and polygenic risk scores show limited utility in our setting. Rash correlates with improved outcomes, whereas hepatitis is associated with a poorer prognosis, independent of baseline prognostic state assessed by the Real World Prognostic score.

Conclusions These findings highlight the complexity of immune toxicities in early-phase trials, emphasizing the importance of the CIT class, as well as the roles of tumor burden, metastasis sites, and systemic immune state in the development of irAEs. Additionally, the observed association between skin toxicities and improved PFS suggests that skin toxicity could serve as a marker of systemic immune activation across immunotherapy contexts.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Cancer immunotherapy can induce immune-related adverse events (irAEs); their management and prognostic significance have advanced thanks to data from large checkpoint inhibitor trials.

WHAT THIS STUDY ADDS

⇒ This study reveals the complexity of irAEs in early-phase pan-immunotherapy trials, highlighting the impact of tumor burden, metastasis sites, and systemic immune state, while identifying skin toxicity as a potential surrogate marker for improved patient outcomes.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our study lays a foundation for pan-immunotherapy irAE research, offering insights for clinicians and drug developers to assess risk profiles and guide the design of future trials for new immunotherapies.

INTRODUCTION

Checkpoint inhibitors (CPIs) have significantly enhanced cancer treatment and are now the standard of care for many cancer indications. However, CPI treatment is often associated with the development of immune-related adverse events (irAEs), which vary in type and severity depending on the context and the specific CPI drug administered.[1] Generally, the occurrence patterns of irAEs are known to be highly agent specific. Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) inhibitors are associated with

high-grade irAEs (Grade 3–4) in about 30% of patients, more commonly manifesting as colitis, hypophysitis, and skin rash. In contrast, treatment with an anti-programmed death-ligand 1 (PD-L1) antibody or an anti-programmed cell death protein 1 (PD-1) antibody results in fewer severe irAEs, affecting about 10% of patients, with pneumonitis, myalgia, hypothyroidism, arthralgia, and vitiligo being more frequently observed.[1 2] IrAEs typically occur within the first 12 weeks of treatment. They generally occur earlier with anti-CTLA-4 treatment, and distinct timing patterns can be observed across different irAE categories, with cutaneous and gastrointestinal events occurring earlier than endocrine and renal events.[2 3] Additionally, there is evidence suggesting that the occurrence pattern of irAEs depends on the tumor type. It is hypothesized that variations in the tumor microenvironment (TME) may partially explain these tumor-specific differences.[1 2] In a recent study of the PD-L1 inhibitor atezolizumab, the most frequent irAEs of all grades were rash (22.8%), followed by hepatitis (12.4%), hypothyroidism (9.0%), pneumonitis (3.0%), and hyperthyroidism (2.4%). Hepatitis and pneumonitis were the most common high-grade irAEs observed.[4] Although high-grade irAEs are undesirable and may lead to the discontinuation of treatment or, in rare severe cases, death, low-grade irAEs are predictive of treatment success.[3 5]

Biomarkers that predict the occurrence and severity of irAEs would be invaluable in clinical settings. Despite the extensive data available on CPIs to date, general biomarkers remain elusive and appear to be context-specific.[6 7] Nevertheless, several emerging categories of biomarkers have demonstrated predictive power in various contexts. For example, a larger T-cell receptor repertoire at baseline, along with increased on-treatment clonality and richness, has been associated with the development of irAEs in patients treated with CTLA-4 and PD-1/PD-L1 inhibitors.[8–10] Additionally, elevated levels of effector memory CD4+ T cells at baseline have been observed as predictive markers for irAEs in these treatments. A meta-analysis of 15 Roche-sponsored clinical trials involving atezolizumab assessed baseline risk factors and identified organ-specific risk factors, such as elevated liver enzyme levels for hepatitis and elevated thyroid-stimulating hormone (TSH) for thyroid toxicities. Asian origin was also identified as a risk factor for developing rash, hepatitis, and pneumonitis.[4] Germline genetic biomarkers have been explored as predictors of irAEs. A recent genome-wide association study involving 1,751 patients receiving CPI treatment identified a variant in the *IL7* gene that predisposes individuals to irAEs, which was validated in an independent cohort.[11 12] Similarly, a polygenic risk score (PGS) for hypothyroidism has been associated with the development of hypothyroidism following atezolizumab treatment.[13] The impact of pre-existing autoantibodies on irAEs has also been investigated. Several studies report an association between increased antithyroid antibodies and thyroid dysfunction.[14–16] In conclusion, while general biomarkers for predicting irAEs are still lacking, several promising context-specific biomarkers have been identified for CPI.

In addition to CPIs, T cell engaging bispecific antibodies (TCBs), immunocytokines, and immune modulators have also emerged as promising strategies for the treatment of solid tumors, with several agents rapidly advancing towards clinical application. These innovative cancer immunotherapies (CITs) are associated with unique adverse event profiles. For instance, two currently approved TCBs for the treatment of solid tumors, tarlatamab and tebentafusp, are primarily linked to cytokine release syndrome (CRS) and skin-related events.[17 18] More broadly, TCBs often encounter challenges such as on-target off-tumor toxicity and CRS.[19 20] Furthermore, immunocytokines are another promising class of CITs, with high-dose interleukin-2 (IL-2) being the first Food and Drug Administration-approved treatment for renal cell carcinoma and melanoma. While these first-generation immunocytokines were associated with a substantial side effect burden, the development of second-generation immunocytokines has focused on enhancing their efficacy while reducing or controlling toxicity.[21]

Although new patterns of irAEs have been identified in individual clinical trials of non-standard CITs, comprehensive aggregated clinical data remain scarce. The exploration of biomarkers predictive of irAEs is still in its early stages. Therefore, understanding the context and patterns of irAE occurrence, along with identifying biomarkers that can predict these events in novel agents, is crucial for enhancing patient care. To further our understanding of factors influencing irAE risk secondary to non-standard CITs, we have examined both the relationship between tumor characteristics, specific soluble biomarkers, germline genetics and irAEs, as well as the relationship of irAEs with patient outcomes. This analysis is based on harmonized data from 27 Roche-sponsored early-phase clinical trials encompassing 14 molecules in 3,608 patients.

METHODS

Patient cohort

A harmonized data mart was developed from 27 early-phase CIT trials (26 phase I, 1 phase II) sponsored by F Hoffmann–La Roche to assess the occurrence of irAEs, identify associated risk factors, and examine their relationship with progression-free survival (PFS). This comprehensive dataset includes information from 3,608 patients across trials involving 14 distinct therapeutic molecules. These molecules encompass a range of modalities, including agents targeting the PD-1/PD-L1 axis (PDL1 Ab, PD1/TIM3 bsAb); T-cell bispecific antibodies (CEACAM5-TCB, CEA-TCB, TYRP1-TCB); immunocytokines (CEA-IL2v, FAP-IL2v); a costimulatory bispecific antibody (FAP-41BBL); and immunomodulators (TLR7 agonist prodrug, CSF1R Ab, agonistic CD40 Ab, CD25 Ab, BET inhibitor, agonistic OX40 Ab),

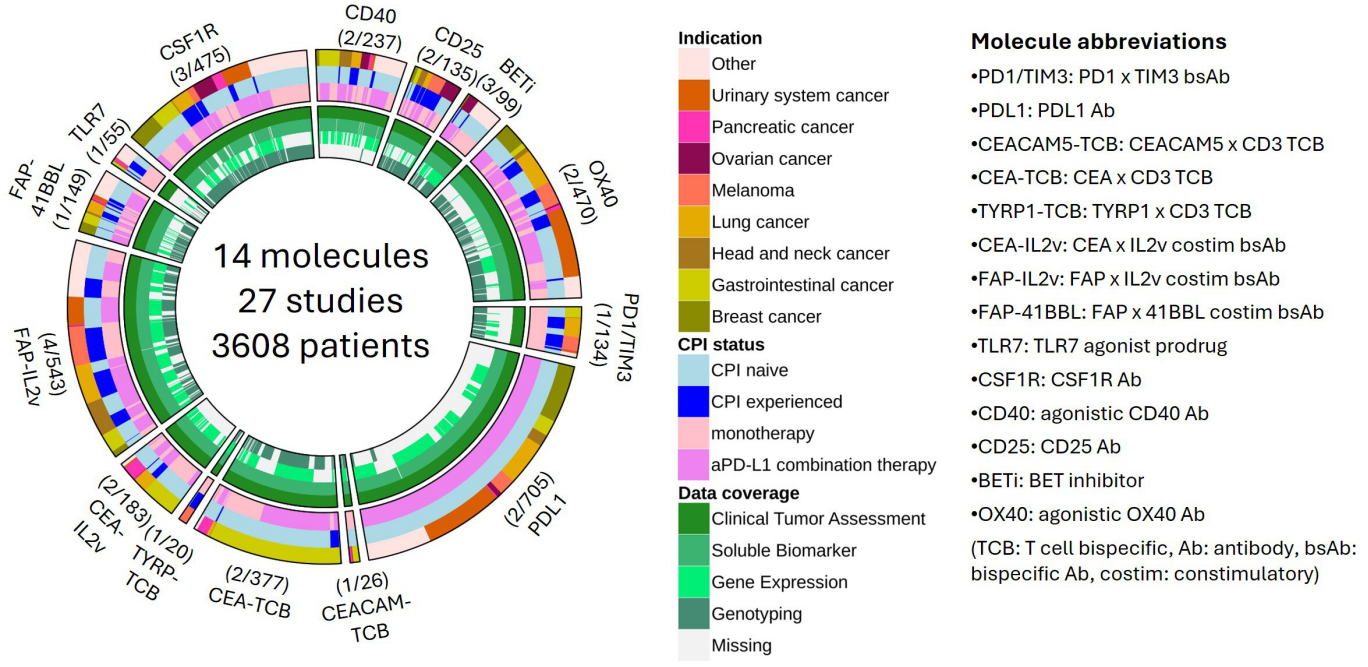

Figure 1 Overview of the harmonized data mart from early-phase CIT trials. The figure illustrates the heterogeneity in CIT treatments and cancer indications, highlighting the depth of data coverage per patient. Note: For each molecule, (X/XX) indicates X number of studies and XX number of patients. CIT, cancer immunotherapy; CPI, checkpoint inhibitor; IL2v, interleukin-2 variant; PD1, programmed cell death protein-1; PDL1, programmed death-ligand 1.

where Ab denotes antibody and bsAb denotes bispecific antibody. The trials span various cancer indications, with a primary focus on gastrointestinal cancers (n=725), urinary system cancers (n=504), and lung cancers (n=498). The patient cohort comprises both CPI-naïve and CPI-experienced individuals (n=756), as well as 2,176 patients receiving anti-PD-L1 combination therapies. The dataset offers detailed information on tumor characteristics, including clinical assessments and gene expression profiles, along with data on soluble biomarkers and genome-wide genotyping (see figure 1 and online supplemental table 1).

Immune-related adverse events

irAEs are summarized by medical concept using a comprehensive set of definitions, including Standardized MedDRA Queries, High Level Terms, and Sponsor-defined Adverse Event Grouped Terms.[4] Time to the first adverse event was considered for each irAE category and patient. Safety follow-up times varied across studies (see online supplemental table 3). Competing risks cumulative incidences compared irAE occurrences across studies and molecules, accounting for censoring and varying follow-up times. The National Cancer Institute's Common Terminology Criteria for Adverse Events were used to grade irAE severity. We recorded the most severe event per patient for each irAE category and compared high-grade (grades 3–5) to low-grade (grades 1–2) events both on an individual molecule basis and in aggregate, descriptively without statistical comparisons.

Risk factors

We evaluated tumor metrics, soluble biomarkers, previous CPI usage, and PGS as potential risk factors for irAEs. Among the tumor metrics, we assessed liver and lung metastasis status, the number of metastases, the sum of the largest diameter of target lesions (SLD), and TME features, such as the level of CD8+ T cell infiltration, and immune-related patterns characterized by specific gene signatures (see online supplemental figure 2 and table 4). TME features were derived from genome-wide RNA expression data (n=1,900), as previously described.[22] The soluble biomarkers measured include various liver enzymes, blood cell counts, lymphocyte subsets, C-reactive protein (CRP), ferritin, and TSH. All were treated as continuous variables. Detailed information on their ranges and measurement units is provided in the online supplemental figures 6-9. PGS were selected from the EBI PGS Catalog[23] through a systematic search. This process involved mapping irAEs to related traits and filtering scores based on quality criteria, source study, and development method, as detailed in the online supplemental material. This yielded a total of 16 scores: 7 for hepatitis, 5 for rash, 1 for acute kidney injury, and 3 for hypothyroidism. The selected scores were then calculated for all 1,929 patients with available genotyping data.

Statistical analysis

The median follow-up time was estimated using the reverse Kaplan-Meier method.[24] Cumulative

incidences of competing risks were calculated, and incidence proportions were compared descriptively without performing statistical comparisons. A two-stage meta-analysis, selected for its robustness and simplicity, was used to identify irAE risk factors for the four most frequent irAEs. At the study level, Cox regression adjusted for age and sex was applied to assess associations with time-to-event endpoints. A random-effects meta-analysis of log hazard ratios (HRs) was then used to evaluate the combined effect across all studies. Time-dependent Cox regression was used for time-varying soluble biomarker measurements up to 1 week from treatment start. To ensure robustness, our analyses required at least six events for Cox regression and three log HRs for meta-analysis. Continuous risk factors were log-transformed and standardized at the study level, except for immune-cell proportions, which were logit-transformed. Similarly, two-stage meta-analysis was used to evaluate the association between irAEs and PFS. Cox regression used PFS as the endpoint and irAE as a time-dependent covariate ("event", "no event", "censored"). Log HRs contrasting "event" and "no event" were summarized with a random-effects meta-analysis. To adjust the effect of irAEs for the general prognostic state at baseline, a multivariate two-stage meta-analysis was performed, treating PFS as the endpoint and including irAE and the Real World Prognostic score (ROPRO)[25] as covariates in the Cox regression. Finally, to determine the "variance explained" by polygenic scores (PGS) developed for liver enzymes[26] in relation to corresponding protein measurements, linear regression was performed on standardized baseline enzyme measurements against PGS, stratified by liver metastasis. The regression slopes were combined across studies via meta-analysis and then squared. Only studies with at least 10 samples available in each stratum and at least 40 in total were included. All analyses were performed using R Statistical Software.[27]

RESULTS

Occurrence of irAEs

Among a total of n=3,573 safety-evaluable patients, n=973 experienced at least one instance of hepatitis, n=967 had rash, n=366 had acute kidney injury, n=159 had hypothyroidism, n=81 had pneumonitis, n=59 had colitis, n=55 had pancreatitis, and n=43 had hyperthyroidism (of whom 16 also experienced a hypothyroidism event). A large percentage of hepatitis (45.2%), pneumonitis (39.0%), colitis (48.3%), and pancreatitis (50.9%) cases were severe (grades 3–5). In contrast, rash (7.6%), acute kidney injury (14.0%), hypothyroidism (2.6%), and hyperthyroidism (4.9%) were rarely severe (see table 1). The severity of irAEs appears to be influenced mainly by the irAE category and, to a lesser extent, by the CIT treatment (see online supplemental figure 1).

Next, we assessed the incidences of irAEs for different molecules. To account for varying study follow-up times, we examined the incidence proportion over a 6-month period (figure 2A). The incidence proportion substantially varies across different CIT molecules, with the most prevalent irAE being hepatitis (21.4–34.9%), rash (19.2–42.6%), acute kidney injury (7.9–13.5%), and hypothyroidism (0.5–7.9%), all reported as 25th–75th percentiles. Hepatitis occurs more frequently with systemic immunocytokines CEA-IL2v and FAP-IL2v (46.2% and 60.2%, respectively) and related immune stimulators FAP-41BBL and TLR7 (38.8% and 35.0%, respectively). Targeted treatments with TCBs are linked to organ-specific toxicities, such as rash and colitis for CEA-TCB (42.7% and 4.4%, respectively), rash for TYRP1-TCB (61.4%), and pancreatitis for CEACAM5-TCB (29.2%). CD25 antibody treatment is associated with high incidences of rash (59.5%), while BET inhibitors are linked to higher incidences of acute kidney injury (18.6%).

Finally, we assessed the kinetics of different irAEs and CIT molecules by categorizing them into instant (<1 week), rapid (1 week to 1 month), moderate (1–3 months), and

Table 1 Summary of irAEs: *n* represents the number of patients experiencing at least one irAE; *N* denotes the total number of patients evaluated for safety; *n* (low grade) indicates the number of patients with at least one low-grade event (grades 1–2); *n* (high grade) refers to the number of patients with at least one high-grade event (grades 3–5)

	Total events			High versus low grade events		
	n	N	Incidence (%)	n (low grade)	n (high grade)	Proportion Severe (%)
Hepatitis	973	3,573	27.2	527	435	45.2
Rash	967	3,573	27.1	865	71	7.6
Acute kidney injury	366	3,573	10.2	313	51	14
Hypothyroidism	159	3,573	4.5	151	4	2.6
Pneumonitis	81	3,573	2.3	47	30	39
Colitis	59	3,573	1.7	30	28	48.3
Pancreatitis	55	3,573	1.5	26	27	50.9
Hyperthyroidism	43	3,573	1.2	39	2	4.9

irAEs, immune-related adverse events.

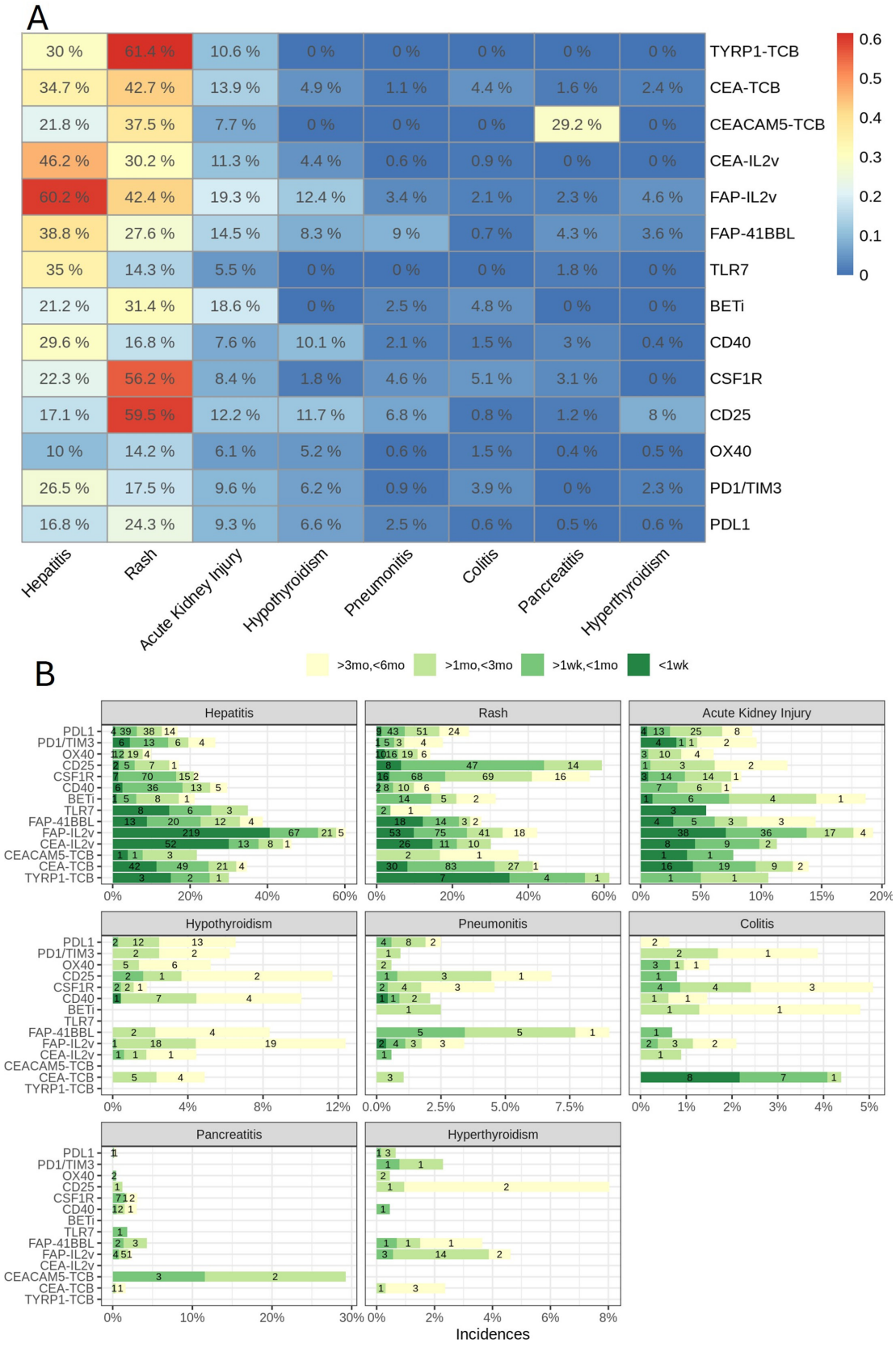

Figure 2 Incidence and onset of irAEs. (A) 6-month incidence proportions of irAEs for various CIT molecules, with cumulative incidences represented by numbers and colors. (B) Onset timing of irAEs categorized by CIT molecules, showing cumulative incidences up to 6 months: instant (<1 week), rapid (1 week to 1 month), moderate (1–3 months), and delayed (3–6 months), along with the number of events in each category. CIT, cancer immunotherapy; IL2v, interleukin-2 variant; irAEs, immune-related adverse events; PD1, programmed cell death protein-1; PDL1, programmed death-ligand 1; TCB, T cell engaging bispecific antibody.

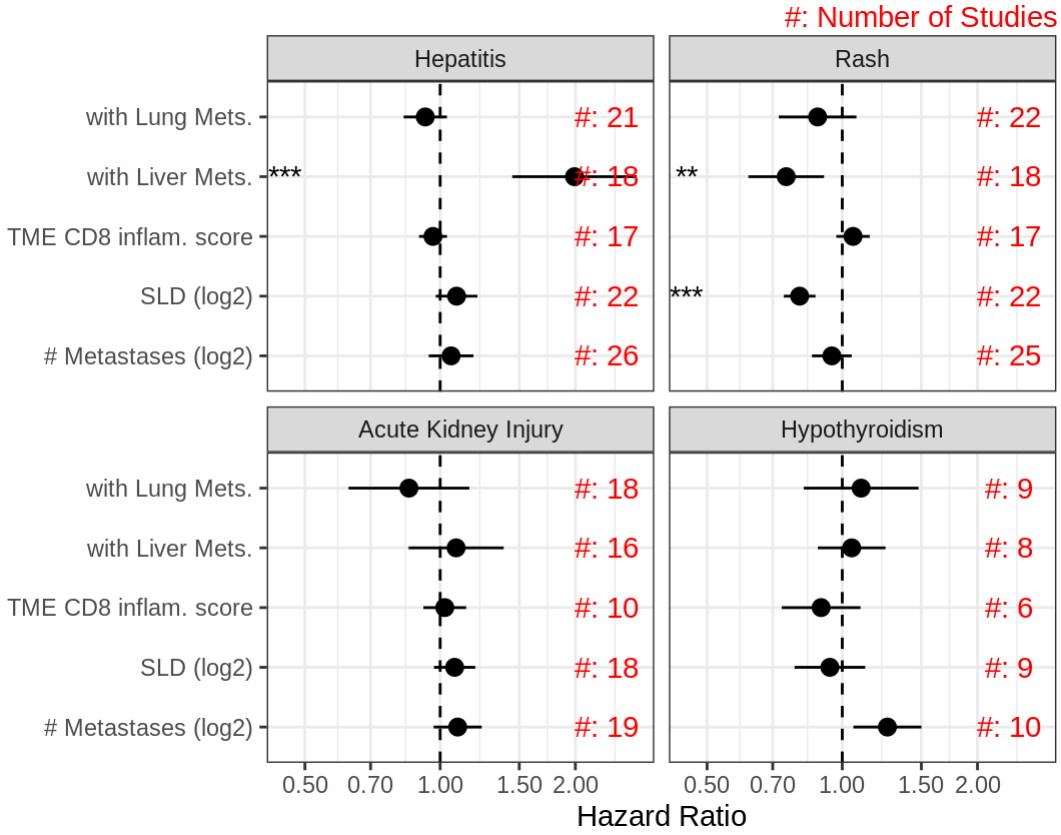

Figure 3 Association of tumor metrics with irAEs. Tumor metrics include the presence of Liver/Lung Mets, the number of metastases (log2-transformed), the SLD (log2-transformed), and the level of CD8+ T cell infiltration (TME CD8 inflammation score as described[22]). HRs and 95% CIs were obtained through a confounder-adjusted two-stage meta-analysis. #(number) indicates the number of studies included in the meta-analysis. ** indicates results with a p value<0.01, and *** indicates results with a p value<0.001. irAEs, immune-related adverse events; Liver/Lung Mets, liver and lung metastases; SLD, sum of the largest diameters of target lesions; TME, tumor microenvironment.

delayed onset (3–6 months). The onset of irAEs varies greatly across treatments and irAE types. Notably, early onset appears to be characteristic of hepatitis, specifically for IL-2 immunocytokines where the onset occurs instantly in the majority of cases, whereas late onset seems to be associated with hypothyroidism (see figure 2B).

In the following sections, we will focus on the most frequent irAEs: hepatitis, rash, acute kidney injury, and hypothyroidism.

Identification of irAE risk factors

Next, we assessed tumor metrics, soluble biomarkers, previous CPI usage, and PGSs as potential risk factors for irAEs. Each factor was tested for association with irAEs using a confounder-adjusted two-stage meta-analysis. Specifically, for each irAE category and potential risk factor, the association was assessed at the study level using Cox regression adjusted for age and sex. The obtained HRs and 95% CIs were then summarized using a random-effects meta-analysis, displayed in figures 3–5.

Tumor metrics

Among the tumor metrics, we assessed liver and lung metastasis status, the number of metastases, the SLD, and TME features, such as the level of CD8+ T cell infiltration and immune-related patterns characterized by specific gene signatures.[22] Meta-analysis reveals that patients with liver metastases have an increased risk of hepatitis (HR 1.99, 95% CI (1.45 to 2.74)). The same group exhibits a reduced risk of rash (HR 0.75, 95% CI (0.62 to 0.91)), as do patients with larger SLD at baseline (HR 0.80, 95% CI (0.74 to 0.87)) (see figure 3). We found no evidence suggesting that tumor metrics influence other irAE categories, such as acute kidney injury and hypothyroidism. Additionally, after analyzing nine gene expression-based TME features, we found no convincing association with irAE development, suggesting that irAE development is relatively independent of the TME. The strongest association was observed between the B cell signature and rash (unadjusted p value: 0.01; HR 1.14; 95% CI (1.03 to 1.26)). For detailed gene signature results, please refer to online supplemental table 9.

Soluble biomarkers

For soluble biomarkers, data were available both at baseline and during treatment, although sampling frequency and data coverage varied significantly across studies (see online supplemental figure 6). To distinguish baseline effects from on-treatment effects, we conducted

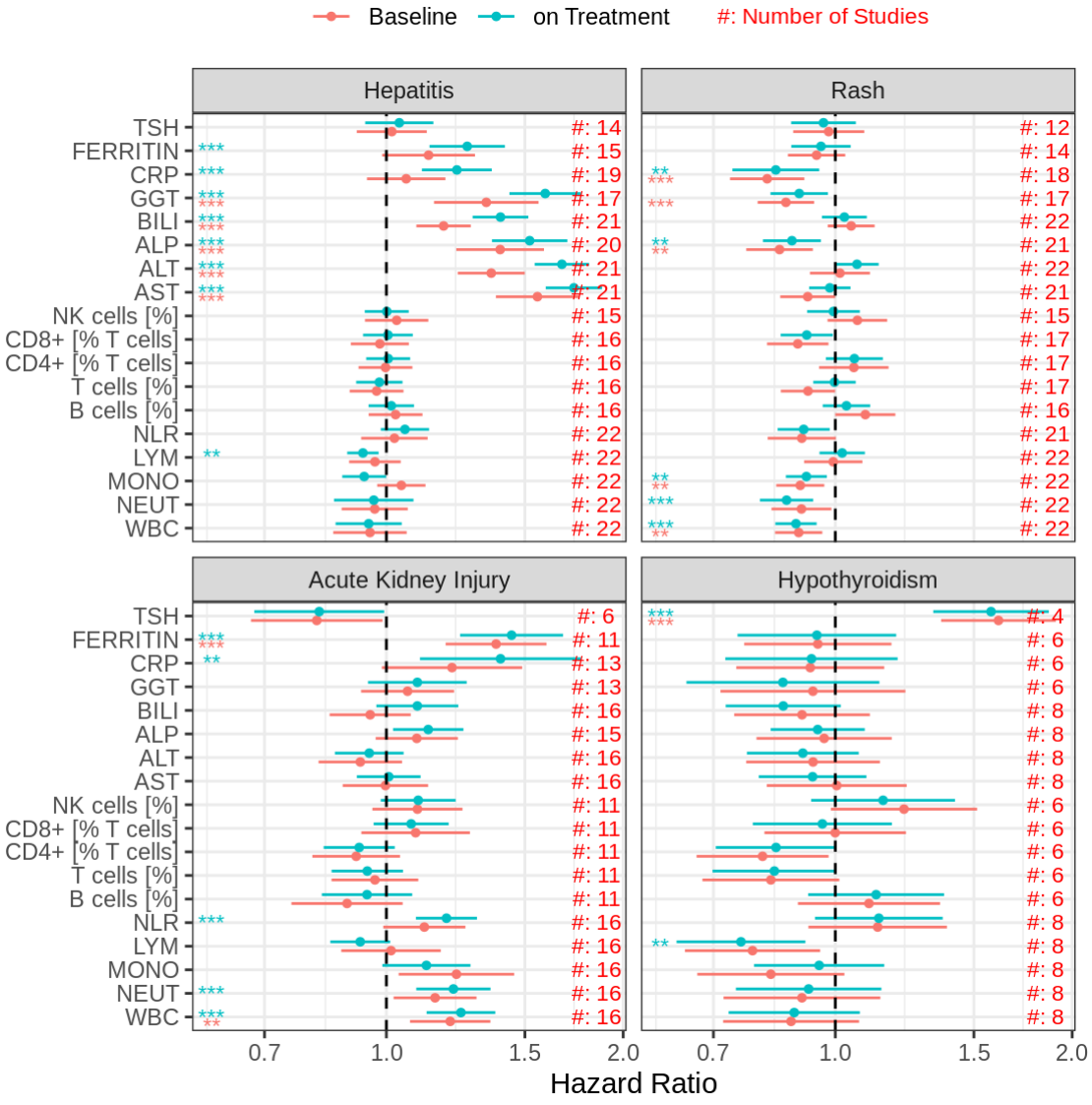

Figure 4 Association of soluble biomarkers with irAEs. HRs and 95% CIs were obtained through a confounder-adjusted two-stage meta-analysis. Red lines represent results obtained from baseline samples only, while blue lines represent results including on-treatment samples. #(number) indicates the number of studies included in the meta-analysis. ** indicates results with a p value<0.01, and *** indicates results with a p value<0.001. ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; BILI, bilirubin; CRP, C-reactive protein; GGT, gamma-glutamyl transpeptidase; irAEs, immune-related adverse events; LYM, lymphocytes; MONO, monocytes; NEUT, neutrophils; NK, natural killer; NLR, neutrophil-to-lymphocyte ratio; TSH, thyroid-stimulating hormone; WBC, white blood cells.

the meta-analysis twice: once using baseline data only and once incorporating on-treatment data as a time-dependent covariate. Our results on baseline data align with expectations and are consistent with previous findings (see figure 4). Elevated levels of liver enzymes are associated with an increased risk of hepatitis diagnosis, as indicated by an HR of 1.56 for aspartate aminotransferase (AST) with a 95% CI of (1.38 to 1.76). Similarly, TSH shows an increased risk of hypothyroidism, with an HR of 1.61 and a 95% CI of (1.36 to 1.91). Higher levels of myeloid cells and CRP (eg, HR for CRP 0.82, 95% CI (0.73 to 0.91)) are associated with a reduced risk of rash, whereas elevated ferritin levels (HR 1.38, 95% CI (1.19 to 1.60)) correspond to a higher risk of acute kidney injury. Including on-treatment data particularly strengthens the observed association for hepatitis, as exemplified by the increase in the HR for AST from 1.56 to 1.73 (see figure 4). Given liver enzymes and TSH are used to diagnose hepatitis and thyroid toxicities, respectively, the consistent associations we observe are expected. Since liver enzymes and liver metastases were associated with hepatitis risk, we investigated their interplay. First, as expected, patients with liver metastases had higher baseline liver enzymes (see online supplemental figure 3). Adjusting for AST (as a representative of liver enzymes) in our analysis attenuated the association between liver metastases and hepatitis (effect size decreased from 1.99 to 1.57), though it remained significant (see online supplemental table 6). This finding suggests that elevated baseline AST only partially mediates the increased hepatitis risk in this

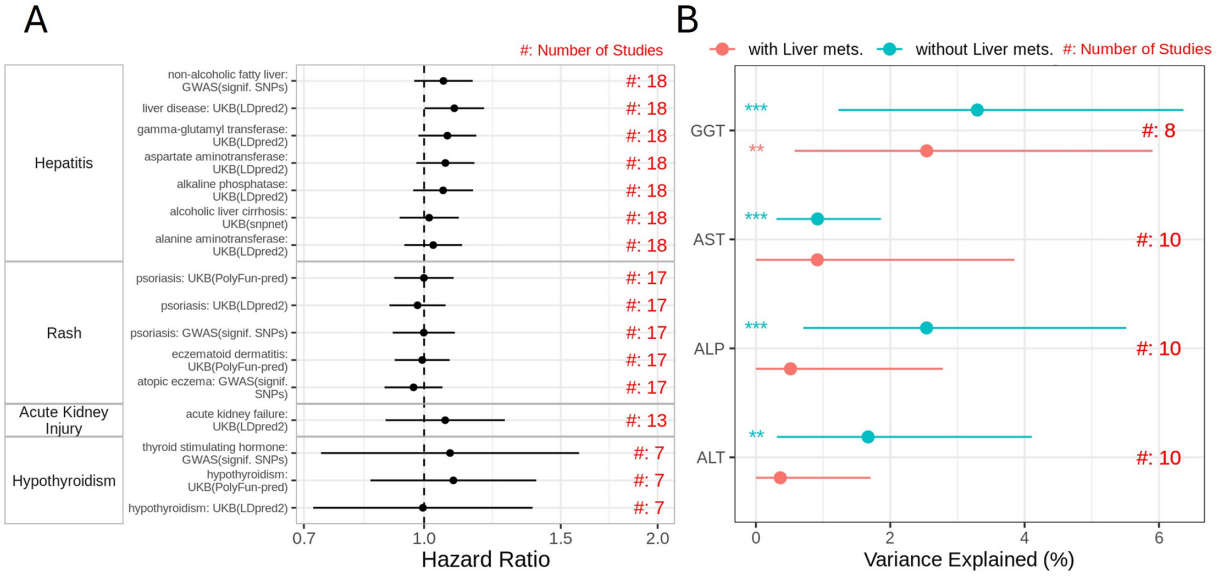

Figure 5 Association with PGS linked to irAEs. (A) HRs and 95% CIs from a confounder-adjusted, two-stage meta-analysis are shown for each PGS and corresponding irAE category. The PGS were selected through a systematic search of the EBI Polygenic Risk Score Catalog, which involved mapping irAEs to related traits and filtering scores based on quality criteria, source study, and development method. Each PGS is identified by its trait, data source (UKB or GWAS), and development method (in parentheses). #(number): the number of studies included in the meta-analysis. (B) Variance explained by PGS developed for liver enzymes with corresponding protein measurements, stratified by liver metastasis status. Variance explained and 95% CI were obtained through univariate two-stage meta-analysis. #(number) indicates the number of studies included in the meta-analysis. ** indicates results with a p value<0.01, and *** indicates results with a p value<0.001. ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; GGT, gamma-glutamyl transpeptidase; GWAS: Genome-wide association study; irAEs, immune-related adverse events; PGS, polygenic risk score; SNPs, single-nucleotide polymorphisms; UKB: UK Biobank.

population. Furthermore, an additional analysis suggests an interaction between liver metastases and treatment type on hepatitis risk. Specifically, immunocytokines appeared to elevate AST particularly in patients without liver metastases, whereas for CEA-TCB, elevations were more pronounced in patients with liver metastases (see online supplemental figure 4).

Previous CPI treatment

We assessed whether previous treatment with CPI impacted the risk of developing irAEs. Out of 3,608 patients, 756 were CPI-experienced. The meta-analysis highlighted moderate evidence for a reduced risk of hepatitis (HR 0.71, 95% CI (0.55 to 0.91)) and hypothyroidism (HR 0.38, 95% CI (0.19 to 0.76)) associated with prior CPI exposure (see online supplemental table 10). However, these findings must be interpreted with caution, as they may be driven by the selection bias detailed in the Discussion.

Polygenic risk scores

To assess the presence of germline genetic components predisposing to irAEs, PGS matching irAE categories (see Methods section for details) were calculated for patients with available genotyping data from blood (n=1,929). We tested the association between irAEs and relevant PGS using the aforementioned two-stage meta-analysis approach. The observed associations were generally weak, with only one liver disease score demonstrating a positive association with hepatitis (unadjusted p value: 0.05; HR 1.09, 95% CI (1.00 to 1.19)) (see figure 5A and online supplemental table 19 for details). Notably, the PGS developed for liver enzymes—alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, and gamma-glutamyl transpeptidase—did not show an association with hepatitis, despite strong signals observed at the protein level. We therefore sought to directly assess the association of PGS with soluble biomarkers and the impact of cancer-induced variability by conducting the analysis for patients with and without liver metastasis. As expected, liver enzymes showed a significant association with their corresponding PGS. Interestingly, the variance explained tended to be larger in patients without liver metastasis, exemplifying the impact of cancer-induced variability. Overall, the maximal observed variance explained was no greater than 4% (see figure 5B).

Relationship of irAEs with clinical outcome

To assess the relationship between irAEs and PFS, we first employed a two-stage meta-analysis. This initial analysis revealed that rash was associated with slower progression (HR 0.78, 95% CI (0.70 to 0.87)), while hepatitis was associated with faster progression (HR 1.43, 95% CI (1.26 to 1.62)). We noted that known prognostic markers (like the neutrophil-to-lymphocyte ratio) were associated with the risk of rash, and we hypothesized that the poor prognosis with hepatitis was driven by liver metastases. To

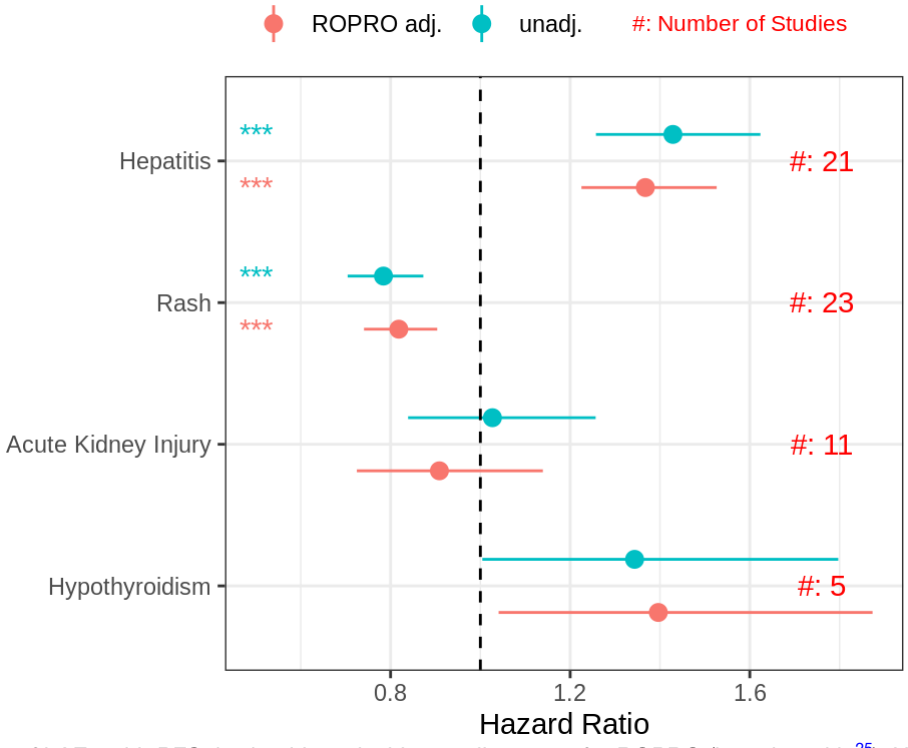

Figure 6 Association of irAEs with PFS, both with and without adjustment for ROPRO (introduced in[25]). HRs and 95% CIs were obtained through multivariate two-stage meta-analysis treating PFS as the time-to-event endpoint and irAE as covariate, with ROPRO as a potential control covariate. #(number) indicates the number of studies included in the meta-analysis. ** indicates results with a p value<0.01, and *** indicates results with a p value<0.001. irAEs, immune-related adverse events; PFS, progression-free survival; ROPRO, Real World Prognostic score.

disentangle these effects, we conducted a more rigorous bivariate meta-analysis that included both the irAE and the ROPRO[25] as covariates. ROPRO is a validated prognostic score that predicts overall survival in patients with cancer by integrating routine clinical variables related to lifestyle (body mass index, smoking), host factors (eg, albumin, Eastern Cooperative Oncology Group, age), and tumor characteristics (stage, number of metastatic sites). Interestingly, our analysis reveals that irAEs and ROPRO are independent predictors of PFS. While higher ROPRO values are associated with an increased hazard of progression, rash and hepatitis remain linked to slower and faster progression, respectively (see figure 6). As a further supplemental analysis, we reevaluated the hepatitis association after excluding patients with liver metastases. The link to poorer PFS remained significant, although with a smaller effect (HR 1.28 vs 1.43), with the caveat that occult hepatic metastases may still be present in these patients (see online supplemental table 7).

DISCUSSION

To the best of our knowledge, this analysis of 27 Roche-sponsored trials represents the first systematic investigation of irAEs associated with non-standard CITs in early-phase trials for advanced solid tumors. Understanding these adverse events is crucial for optimizing treatment strategies and improving patient outcomes. As expected, the incidence patterns of irAEs differ among CIT molecules and often align with anticipated trends based on their modes of action. Reassuringly, our incidence rates for treatments targeting the PD(L)1 axis are consistent with those reported in large anti-PD(L)1 trials. The findings on incidence proportions, severity, and onset can serve as a valuable resource for clinicians and drug developers to assess and compare the risk profiles of newly investigated immunotherapies, as well as to refine treatment regimens and the design of future trials.

Our analysis of risk factors is complicated by the significant heterogeneity among the included studies, particularly regarding molecule classes, study designs (eg, varying dosing regimens or schedule of assessment), included indications, concomitant treatments, and sample sizes (ranging from 20 to 660). This stands in stark contrast to similar investigations conducted in large phase 3 trials for standard CPI treatments. To manage differences across studies and avoid reporting associations valid in only a single study, we used a robust statistical approach based on meta-analysis. This method appropriately weights each study's contribution and includes quality criteria that protect against statistical instabilities arising from small sample sizes. This approach helped us find reliable associations that are valid pan-immunotherapy. Liver metastases correlate positively with hepatitis but conversely were associated with reduced risk of rash. Elevated liver enzyme levels are linked to hepatitis, higher ferritin levels are associated with an increased risk of acute kidney

injury, and higher myeloid cell counts and elevated CRP levels are associated with lower risk of rash. Notably, several of these markers were also identified as the strongest risk factors in a recent systematic review of atezolizumab trials.[4] Interestingly, we found moderate evidence suggesting that prior CPI exposure is associated with a reduced risk of hepatitis and hypothyroidism. Although this unexpected finding lacks a complete explanation, one likely reason is selection bias; indeed, the protocols for several trials excluded patients with a history of severe irAEs from prior CPI treatment (see online supplemental table 2). Furthermore, patients with pre-existing hypothyroidism are likely receiving long-term thyroxine therapy and so could not experience a subsequent hypothyroidism event, even if they were not explicitly excluded from a trial.

Several recent studies involving large trials of standard CPIs have reported promising associations between genetic features and irAE risk. Consequently, we explored polygenic scores as potential risk factors. However, the observed effect sizes were small, and their predictive utility was limited in this early-phase trial setting, likely due to the heterogeneity of our study population, influenced by varying diseases and cancer types, as exemplified by our analysis of liver enzymes and corresponding PGSs. In our analysis, we deliberately adopted a stringent approach to PGS selection. We focused on traits specific to our irAEs and chose scores developed with the most relevant data and methods. While this strategy effectively reduced the burden of multiple testing, it may have excluded some potentially promising, recently developed scores.[28 29]

It was hypothesized that variations in the TME might explain the risk of irAEs. However, our analysis of the TME using RNA sequencing data did not reveal any such association. This is in keeping with irAE development being more linked to features of systemic immunity relative to those of the TME.

We further evaluated irAEs as predictors of patient outcomes. Despite the heterogeneity of our cohort, we observed a consistent association between rash and improved PFS. Interestingly, our analysis suggests that this effect is independent of other baseline prognostic factors. These findings align with observations in the CPI setting,[30 31] supporting the notion that skin toxicities may reflect a response at the tumor site and suggesting a broad immunotherapy benefit given the distinct mechanisms of action of CPIs and the drugs in this cohort. Conversely, other toxicities show no association with PFS or are prognostically unfavorable. Understanding the common features of skin toxicity and treatment response may offer insights for improved therapeutics.

Our study has several limitations that should be considered. First, the safety follow-up duration across studies was both limited and variable, with a median follow-up time ranging from 2 to 15 months. This constraint restricted our investigation of the dynamics of irAEs to a 6-month period, thereby excluding the assessment of long-term and chronic events. Additionally, our analysis was confined to evaluating the time to the first occurrence of irAEs. Analyzing recurrent events, as well as the duration and resolution of these events, could have provided deeper insights into the overall safety profile of irAEs.

The studies included in our data mart exhibit significant heterogeneity across various aspects. For instance, exploring the impact of specific indications on the pattern of irAEs would have been insightful. However, the uneven distribution of indications across the studies could lead to misleading conclusions if irAE rates were compared naively by indication. Similarly, limited sample sizes restricted our ability to identify context-specific risk factors. Further, despite our attempt to focus our investigation on the most relevant factors, the analysis still involved a large number of statistical tests. Therefore, the results should be interpreted with care, taking into account what is known from other studies and the biological understanding of the mechanisms.

Our study included biomarkers of the main immune-cell populations, such as CD4 and CD8 T cells, B cells, and natural killer cells. Evaluating more fine-grained immune-cell subsets in the periphery would be of high interest; for example, CD4+ T effector memory cells have been reported to be predictive of irAEs. Similarly, a standardized cytokine panel would have been another interesting biomarker to investigate.

An interesting aspect of our evaluation of soluble biomarkers as predictors of irAE risk was the impact of including on-treatment samples. This is important because such data can guide patient monitoring. Our results indicate that on-treatment data significantly strengthens the association between liver enzymes and hepatitis. This finding is not surprising, as liver enzymes are used to diagnose hepatitis in the first place. However, we did not observe similar associations for other biomarkers. A notable limitation of our data mart was the substantial variation in on-treatment sampling and, consequently, data coverage across studies and biomarkers. Our analysis investigated the association of one risk factor at a time, with the aim of identifying a robust association. This can be seen as the first step in developing a prognostic or predictive score.[32 33] Development of such a score involves several steps of validation and might ultimately involve a combination of more than one factor. To our knowledge, no such validated score has been developed for irAEs, and admittedly, our work does not address this gap.[7]

Finally, irAEs are not the only “unwanted” immune responses associated with immunotherapies. Several of the molecules we investigated also encountered issues such as CRS and the development of antidrug antibodies.[34] Our data mart provides an excellent starting point for exploring these topics further, though we note that consistently captured CRS data is limited. A systematic investigation of antidrug antibodies in relation to non-standard CITs, with a particular focus on the role of

the human leukocyte antigen region, offers an especially interesting avenue for future research.[35 36]

Overall, our findings highlight the complexity of immune toxicities in the early-phase setting, emphasizing the importance of the CIT class, tumor burden, involved sites of metastases, and systemic immune state in the development of irAEs. These insights could guide future research and clinical strategies to mitigate risks and enhance therapeutic efficacy.

Author affiliations
[1]Roche Pharma Research and Early Development, F Hoffmann-La Roche, Basel, BS, Switzerland
[2]Roche Product Development, F Hoffmann-La Roche, Basel, BS, Switzerland
[3]Roche Pharma Research and Early Development, Roche Glycart AG, Schlieren, ZH, Switzerland
[4]Human Genetics, Genentech, South San Francisco, California, USA
[5]Roche Product Development, F Hoffmann-La Roche, Welwyn Garden City, UK
[6]Oncology Department, University of Oxford, Oxford, UK

Acknowledgements The authors would like to thank the patients, their families, the participating study centers, and the members of the Enhanced Data Insights and Sharing network for their support in the curation and integration of the clinical data.

Contributors NS, DFL and VCS contributed equally. NS, BPF are joint senior authors. NS is the guarantor. NS conceptualized and oversaw the study, with support from BPF, DFL, and VCS. Primary data analyses were conducted by NS, DFL, and VCS. The interpretation of results was carried out by NS, BPF, DFL, VCS, RM, VK, DH, and SLM. NS wrote the manuscript with input from BPF, DFL, and VCS. TK-T, PS, and CA provided support with the analysis of genotyping data. The data mart effort was co-led by NS, GD-N, and AR. All authors reviewed and approved the final manuscript. NotebookLM was used to extract protocol-defined exclusion criteria for immune-related adverse events from prior checkpoint inhibitor therapy (online supplemental table 2). The results were confirmed and complemented by a manual review.

Funding This study is a meta-analysis of completed clinical trials funded by F Hoffmann-La Roche. NS, DFL, VCS, RM, VK, TK-T, PS, GD-N, AR, DH, CA, and SLM are employees of F Hoffmann-La Roche and received support for the preparation of this manuscript. BF received consultancy fees from F Hoffmann-La Roche. The funder had no influence on the study design, data collection, analysis, interpretation of the results or the decision to submit the manuscript for publication, despite author affiliations with the funder.

Competing interests NS, DFL, VCS, RM, VK, AR, DH, TK-T, GD-N, SLM, and PS are employees of F Hoffmann-La Roche. RM is a coinventor on patents filed by Genentech/Roche that are related to atezolizumab use. CA is an employee of Genentech. NS, DFL, RM, VK, DH, TK-T, GD-N, AR, and PS are shareholders of F Hoffmann-La Roche. CA is a shareholder of Genentech. BF has performed consultancy for NICE Consultancy, F Hoffmann-La Roche, Pathios, UCB and TCypher and has received speaker fees from GSK and BMS.

Patient consent for publication Not applicable.

Ethics approval This retrospective meta-analysis was conducted using individual patient data from previously completed phase I and II trials, in accordance with Roche's responsible data reuse processes and in compliance with the master informed consent form (ICF) for each study. Ethics Committee/Institutional Review Board (IRB) approvals were obtained individually for all 27 studies prior to their conduct. Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer-reviewed.

Data availability statement Data are available upon reasonable request. Data are available on reasonable request. Due to Roche company policy and patient privacy reasons, access to individual patient-level data is restricted. For up-to-date details on Roche's Global Policy on the Sharing of Clinical Information and how to request access to related clinical study documents, see here: https://go.roche.com/data_sharing.

ORCID iDs
Rajat Mohindra https://orcid.org/0009-0000-4233-5210
Vaios Karanikas https://orcid.org/0009-0009-9523-7425
Tony Kam-Thong https://orcid.org/0000-0002-5300-9573
Andreas Roller https://orcid.org/0000-0001-6335-0962
Benjamin P Fairfax https://orcid.org/0000-0001-7413-5002
Nicolas Staedler https://orcid.org/0000-0001-8212-7459

REFERENCES

1 Khoja L, Day D, Wei-Wu Chen T, *et al*. Tumour- and class-specific patterns of immune-related adverse events of immune checkpoint inhibitors: a systematic review. *Ann Oncol* 2017;28:2377–85.
2 Suijkerbuijk KPM, Eijs MJM, Wijk F, *et al*. Clinical and translational attributes of immune-related adverse events. *Nat Cancer* 2024;5:557–71.
3 Weber JS, Hodi FS, Wolchok JD, *et al*. Safety Profile of Nivolumab Monotherapy: A Pooled Analysis of Patients With Advanced Melanoma. *J Clin Oncol* 2017;35:785–92.
4 Madjar K, Mohindra R, Durán-Pacheco G, *et al*. Baseline risk factors associated with immune related adverse events and atezolizumab. *Front Oncol* 2023;13:1138305.
5 Das S, Johnson DB. Immune-related adverse events and anti-tumor efficacy of immune checkpoint inhibitors. *J Immunother Cancer* 2019;7:306.
6 Manson G, Norwood J, Marabelle A, *et al*. Biomarkers associated with checkpoint inhibitors. *Ann Oncol* 2016;27:1199–206.
7 Hommes JW, Verheijden RJ, Suijkerbuijk KPM, *et al*. Biomarkers of Checkpoint Inhibitor Induced Immune-Related Adverse Events-A Comprehensive Review. *Front Oncol* 2020;10:585311.
8 Earland N, Zhang W, Usmani A, *et al*. CD4 T cells and toxicity from immune checkpoint blockade. *Immunol Rev* 2023;318:96–109.
9 Lozano AX, Chaudhuri AA, Nene A, *et al*. T cell characteristics associated with toxicity to immune checkpoint blockade in patients with melanoma. *Nat Med* 2022;28:353–62.
10 Hutchinson JA, Kronenberg K, Riquelme P, *et al*. Virus-specific memory T cell responses unmasked by immune checkpoint blockade cause hepatitis. *Nat Commun* 2021;12:1439.
11 Groha S, Alaiwi SA, Xu W, *et al*. Germline variants associated with toxicity to immune checkpoint blockade. *Nat Med* 2022;28:2584–91.
12 Taylor CA, Watson RA, Tong O, *et al*. IL7 genetic variation and toxicity to immune checkpoint blockade in patients with melanoma. *Nat Med* 2022;28:2592–600.
13 Khan Z, Hammer C, Carroll J, *et al*. Genetic variation associated with thyroid autoimmunity shapes the systemic immune response to PD-1 checkpoint blockade. *Nat Commun* 2021;12:3355.
14 Osorio JC, Ni A, Chaft JE, *et al*. Antibody-mediated thyroid dysfunction during T-cell checkpoint blockade in patients with non-small-cell lung cancer. *Ann Oncol* 2017;28:583–9.
15 Maekura T, Naito M, Tahara M, *et al*. Predictive factors of nivolumab-induced hypothyroidism in patients with non-small cell lung cancer. *In Vivo* 2017;31:1035–9.
16 Kimbara S, Fujiwara Y, Iwama S, *et al*. Association of antithyroglobulin antibodies with the development of thyroid dysfunction induced by nivolumab. *Cancer Sci* 2018;109:3583–90.
17 Hassel JC, Berking C, Forschner A, *et al*. Practical guidelines for the management of adverse events of the T cell engager bispecific tebentafusp. *Eur J Cancer* 2023;191:112986.
18 Paz-Ares LG, Felip E, Ahn M-J, *et al*. Randomized phase 3 study of tarlatamab, a DLL3-targeting bispecific T-cell engager (BiTE),compared to standard of care in patients with relapsed small cell lung cancer (DeLLphi-304). *JCO* 2023;41:TPS8611.

19 Shimabukuro-Vornhagen A, Gödel P, Subklewe M, *et al*. Cytokine release syndrome. *J Immunother Cancer* 2018;6:56.
20 Liu J, Zhu J. Progresses of T-cell-engaging bispecific antibodies in treatment of solid tumors. *Int Immunopharmacol* 2024;138:112609.
21 Rokade S, Damani AM, Oft M, *et al*. IL-2 based cancer immunotherapies: an evolving paradigm. *Front Immunol* 2024;15:1433989.
22 Roller A, Davydov II, Schwalie PC, *et al*. Tumor-agnostic transcriptome-based classifier identifies spatial infiltration patterns of CD8+T cells in the tumor microenvironment and predicts clinical outcome in early-phase and late-phase clinical trials. *J Immunother Cancer* 2024;12:e008185.
23 Lambert SA, Wingfield B, Gibson JT, *et al*. Enhancing the Polygenic Score Catalog with tools for score calculation and ancestry normalization. *Nat Genet* 2024;56:1989–94.
24 Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. *Control Clin Trials* 1996;17:343–6.
25 Becker T, Weberpals J, Jegg AM, *et al*. An enhanced prognostic score for overall survival of patients with cancer derived from a large real-world cohort. *Ann Oncol* 2020;31:1561–8.
26 Privé F, Aschard H, Carmi S, *et al*. Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *Am J Hum Genet* 2022;109:12–23.
27 R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2023.
28 Middha P, Thummalapalli R, Quandt Z, *et al*. Germline prediction of immune checkpoint inhibitor discontinuation for immune-related adverse events. *J Immunother Cancer* 2025;13:e011273.
29 Middha P, Thummalapalli R, Betti MJ, *et al*. Polygenic risk score for ulcerative colitis predicts immune checkpoint inhibitor-mediated colitis. *Nat Commun* 2024;15:2568.
30 Cortellini A, Chiari R, Ricciuti B, *et al*. Correlations Between the Immune-related Adverse Events Spectrum and Efficacy of Anti-PD1 Immunotherapy in NSCLC Patients. *Clin Lung Cancer* 2019;20:237–47.
31 Milotay G, Little M, Watson RA, *et al*. CMV serostatus is associated with improved survival and delayed toxicity onset following anti-PD-1 checkpoint blockade. *Nat Med* 2025;31:2350–64.
32 Altman DG, Royston P. What do we mean by validating a prognostic model. *Stat Med* 2000;19:453–73.
33 Dejardin D, Kraxner A, Schindler E, *et al*. An overview of statistical methods for biomarkers relevant to early clinical development of cancer immunotherapies. *Front Immunol* 2024;15:1351584.
34 van Brummelen EMJ, Ros W, Wolbink G, *et al*. Antidrug Antibody Formation in Oncology: Clinical Relevance and Challenges. *Oncologist* 2016;21:1260–8.
35 Hammer C, Ruppel J, Kamen L, *et al*. Allelic variation in HLA-DRB1 is associated with development of antidrug antibodies in cancer patients treated with atezolizumab that are neutralizing in vitro. *Clin Transl Sci* 2022;15:1393–9.
36 Kam-Thong T, Scepanovic P, Schmid V, *et al*. 611 Elucidating the role of human leukocyte antigens in antidrug antibody formation from a large cohort study of early-phase non-standard cancer immunotherapy trials. *J Immunother Cancer* 2024;12:A703.